Elucidating the dark side of personality

The psychometric properties of the Dark Core Inventory: Norwegian version

Linn Kristin Dedekam Øwre

A thesis for the professional program at the Department of Psychology

UNIVERSITY OF OSLO

Autumn 2021

Summary

Title of the thesis: Elucidating the dark side of personality. The psychometric properties of the Dark Core Inventory: Norwegian version

Name of supervisors: Cato Grønnerød (main supervisor) and Kristin Gustavson (co-supervisor)

Name of author: Linn Kristin Dedekam Øwre

Background: While dark traits have been widely studied, little is known about their hypothesized common "core". The Dark Core Inventory (DCI) is a 70 item personality inventory measuring the Dark factor of personality (*D*), said to cover this core. The DCI has not yet been validated on a Norwegian sample. The thesis is part of a research project (*Personality and Distress: A Normative and Experimental Study of Response Modes and Robustness Against Faking in Self Report Methods*) at the Department of Psychology, University of Oslo. The present thesis is based on a pilot study for the project.

Objectives: The aim of the pilot study was to examine the psychometric properties of the newly translated Dark Core Inventory (DCI) with reference to the wording of the Norwegian translation, internal consistency and factor structure.

Method: Along with two other students, I recruited a convenience sample through online advertising (n = 295). Nettskjema, an online survey tool developed by the University of Oslo, was used to administer a questionnaire. The questionnaire consisted of 262 items from different personality inventories, among them the DCI items. To determine the factor structure of the DCI I performed an Exploratory Factor Analysis.

Findings: One main factor was identified in the analysis, consistent with the theory on the Dark Factor of personality (*D*). This factor explained 18,73% of the covariance between the DCI items.

Main conclusions: The findings suggest that one common factor, the dark factor of personality (*D*) was present in the material. Age and gender differences were found in the Norwegian sample consistent with the theory on the Dark Factor of personality.

Acknowledgements

As one of my last endeavors for the professional program, conducting a pilot study has been an important learning experience. It both deepened my respect for researchers and renewed my appreciation for the personality psychology domain.

A sincere *thank you* goes to my supervisors, Cato Grønnerød and Kristin Gustavson, who have been invaluable in this process. I have benefited greatly from your sharp thinking and warm support.

Thank you to Simon Rekkedal Rolfsnes and Pia Pran, who have been by my side during my time at the Department of Psychology, as well as on this pilot.

Thank you to my former supervisor Johan Berg at Akershus University Hospital for your mini lectures on Robert D. Hare.

I also wish to thank the pilot study sample. For obvious reasons, I will never know your names, but I am so grateful for your participation. There would be no thesis without you.

Lastly, to my family:

thank you for always believing in me and for cheering me on.

Oslo, October 17th 2021

Table of Contents

Introduction

So you're a tough guy
Like it really rough guy
Just can't get enough guy
Chest always so puffed guy
I'm that bad type

(Bad guy by Billie Eilish)

If you were to quickly convey an informative and balanced picture of what someone's personality is like, you are likely to describe the person using the following five constructs: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness.

Since the 1980s, research on personality has been centered on the Five-Factor Model (FFM) (McCrae & John, 1992), elaborated on by Five-Factor Theory (McCrae & Costa, 2003). Critical views and findings conflicting with the Big Five perspective notwithstanding, the Five-Factor Model is the predominant model of personality today (Widiger, 2017).

After a nearly century long debate concerning the nature and number of personality traits, most have landed on a modest five – found through lexical language research and factor analysis (Goldberg, 1993). First articulated in the 1880s by Francis Galton, the lexical hypothesis states that all the meaningful ways in which people's personalities differ have been encoded in the language (Ashton & Lee, 2005). The logic is simple: if it's important, there's a word for it. The implication for the study of personality is that a taxonomy of personality traits can be derived by sampling language.

In 1936, Gordon Allport and Henry Odbert did the painstaking work of visually scanning dictionaries, retaining around 4500 adjectives that describe personality (Allport & Odbert, 1936). The next major milestone can be dated to the 1940s, when Raymond Cattell reduced the extensive list to 16 traits using factor analysis (Cattell, 1943). Factor analysis is a set of statistical methods used to identify underlying factors based on correlations between variables. The goal is to reduce a large set of variables to smaller clusters, ultimately finding the most fundamental, underlying dimensions. Factor analysis, paired with computer science, has been critical for the development of modern theories on personality.

According to Five-Factor Theory, personality is made up of broad and overarching traits (factors). The five major factors are summaries of their respective facets (six subcategories for

each factor), making it a hierarchical model of personality. The Five-Factor model of personality covers the most important traits. Analogous to the primary colours' ability to reproduce all the colours of visible light, personality is seen as different "blends" of a handful primary traits (Widiger, 2017). Traits are commonly defined as relatively stable patterns of emotion, thought and behavior that emerge early in life. They are not all-or-none phenomena, but placed on continua; something you have "more or less of" – high or low levels. The Five-Factor model states that everyone, across cultures, shares the same five universal traits, but differ in the degree of their manifestation. According to Five-Factor Theory, traits are more than just descriptive summaries – they have internal causal properties. Rooted in biology, the trait *causes* certain behavioural tendencies, including tendencies to think and feel a certain way.

The Five-Factor model has been operationalized in a 240 item questionnaire known as The Neuroticism, Extraversion, Openness - Personality Inventory Revised (NEO-PI-R). The NEO-PI-R assesses 30 specific traits, six for each factor. Results have been replicated cross culturally, (McCrae & Terracciano, 2005) adding to its status in the field.

Independence of the factors

According to Five-Factor theory, the factors are independent (uncorrelated, orthogonal) and irreducible (Costa Jr & McCrae, 1992), meaning that they cannot be meaningfully summarized by higher-order factors. Digman (among others) has challenged this view. Digman's view is that the Big Five can be reduced to a Big Two model consisting of a socialization ("Alpha") factor of personality (Agreeableness, Conscientiousness and Emotional Stability), and a self-actualization ("Beta") factor of personality (Extraversion and Openness) (Digman, 1997). Costa and McCrae (1992) reject this view, offering a two-argument explanation for the scale intercorrelations. First, they claim that the correlations are spurious – an artefact of "biases in implicative meaning" (p. 862). In their view, the socialization factor might as well (and just as easily) be understood as an evaluation factor. People wish to see themselves as agreeable, conscientious and emotionally stable, thus creating a spurious correlation between these scales. What's more, many personality traits overlap two or more of the factors; they are not "pure elements", but analogous to chemical compounds or composite materials, made up of several elements to create a uniquely colored trait. Interpersonal warmth is often used as an example – a trait that is related to both extraversion and agreeableness. The argument is that any instrument that samples traits

broadly are likely to include a variety of such mixed traits, creating intercorrelations of the scales (Costa Jr & McCrae, 1992).

Coverage of the Five-Factor Model

Assume the person you're asked to describe has a mean streak. Will you be able to fully capture him by using the Big Five traits of the Five-Factor Model? In more recent years, the Five-Factor Model has been challenged by a framework that proposes an additional dimension of personality. It is conceptually similar to the Five-Factor model, but the variance is reorganized in a six-factor structure. The Extraversion (X), Conscientiousness (C) and Openness to Experience (O) factors are highly similar in both frameworks, but the Agreeableness (A) and Emotionality (E) factors (Neuroticism in the Five-Factor model) are conceptualized differently. For example, items related to anger (being tolerant and forgiving versus volatile and short-tempered) load on the Agreeableness factor in the HEXACO model, not the Emotionality factor (as one would expect if simply mirroring the Five Factor Model). Within the HEXACO framework, anger is seen through a relational or social "lens" (an interpersonal trait), whereas the Five-Factor Model views anger as intersubjective.

In addition to reorganization of the facets, the HEXACO model nominates a sixth factor, targeting sincerity, fairness, and modesty: Honesty-Humility (H) (Ashton & Lee, 2005). One way of putting it, is that with the addition of this factor, you can make a statement about how "light" (genuine, honest and modest) or "dark" (cunning, rule breaking, arrogant) someone is. In contrast to the Big Five constructs, this "light versus dark" statement has a clear normative ring to it. It seems more than strictly descriptive. Could the advantage of the HEXACO model lie in its apparent ability to incorporate the moral aspects of our personalities?

If given the option to know your blind date's score on a single factor, which would you choose? According to Lee, Ashton, Pozzebon, Visser, Bourdage and Ogunfowora (2009), you should opt for the H factor. It turns out that friends and romantic partners are similar in H, but not in A, all the while overestimating similarity among themselves on that dimension (Lee et al., 2009). As pointed out by Ashton, Lee & de Vries (2014), the sharp division between H and A in both assumed and actual similarity suggests that it is meaningful to disentangle the constructs, rather than treat them as a single one. We may gravitate towards all sorts of people, for all sorts of reasons, but the ones we choose to keep in our lives tend to share our interpersonal values; our "relationship rules of conduct". A tabloid newspaper headline could read something along the line of *we seek those who share our values, not our "personalities"*.

According to Ashton & Lee (2019), using the Big Five model in place of the HEXACO model involves a considerable loss of information – roughly equal to what would be lost by abandoning one of the Big Five traits. As emphasized by Ashton & Lee (2014), the H factor was not "invented" by elevating the Straightforwardness and Modesty facets of the Big Five to a factor-level variable. The sixth factor emerged as a dimension "through the same objective research strategy that led to the discovery of the five-dimensional model" (p. 139) – namely, through the lexical studies of personality (including the study of non-Indo-European languages such as Filipino, Hungarian and Korean). In their view, the embrace of the five-dimensional personality model during the 1990s was premature and the "near-consensus" favoring the five-factor model a "historical accident" (p. 141) – the accident being the underrepresentation of H-related traits in the early lexical studies. Ashton & Lee (2008) state that "the largest factor space that is widely recovered across the lexical personality studies in various languages consists of six rather than five dimensions" (p. 140). The factor axes locations vary across studies "due to lack of simple structure in the personality domain" (p. 140), but the dimensions that surface when rotating to simple structure correspond to the HEXACO factors, not the Big Five factors (Michael C. Ashton & Lee, 2008).

Let's say the person you are asked to describe scores on the very low end of the Honesty-Humility scale. He has a marked tendency to manipulate others (Sincerity scale), is a committed rule breaker (Fairness scale), enjoys lavish displays of wealth and social status (Greed Avoidance scale) and clearly sees himself as someone entitled to privileges and special treatment (Modesty scale). To further elucidate the dark side of his personality, you could reach for an inventory specifically designed to capture variance linked to dark personality traits.

The Dark Triad

In his 2001 book *Behaving badly: Aversive behaviors in interpersonal relationships,* Robin Kowalski has made the case that behaving badly; being a "bad guy" – at least some of the time – is an inevitable part of interacting with others (Kowalski, 2001).

In recent years, there has been a growing body of research on dark personality traits in normal populations and several non-clinical self-report measures have been developed. Prior to 2002, prominent dark personality traits had independent research trajectories (Schreiber & Marcus, 2020). In 2002, Paulhus and Williams coined the term the Dark Triad, linking narcissism, Machiavellianism and psychopathy (Paulhus & Williams, 2002).

Narcissism is a trait that is influenced by the myth of Narcissus, whose tragic fate was to fall in love with himself and miss out on the experience of having a meaningful relationship with another person. The research of Paulhus and Williams on subclinical (normal) narcissism builds on the work of Raskin & Terry (1988), who delineated a subclinical version from the DSM personality disorder, retaining the following facets: grandiosity, entitlement, dominance and superiority (Raskin & Terry, 1988). Recently, Miller et al. (2011) made a distinction between vulnerable narcissism (high Neuroticism, low Extraversion) and grandiose narcissism (low Neuroticism, high Extraversion) (Miller et al., 2011). Both types are self-centered, have fragile self-esteems and a strong need to be admired, but they differ in the kinds of attention they crave. A person suffering from grandiose narcissism wants to be at the center of attention. The more vulnerable type, characterized by hypersensitivity and shame, craves reassurance. Other researchers maintain that the most parsimonious model of trait narcissism is a three-factor structure. This structure makes the distinction between antagonistic, agentic and neurotic narcissism – the antagonistic narcissist having the strongest link to "dark" behavior (Crowe, Lynam, Campbell, & Miller, 2019).

Machiavellianism is a psychological construct that refers to an exploiting and manipulative interpersonal style. It originates from the study of Machiavelli's original works, philosopher and political advisor to the House of Medici in Renaissance Florence (Christie, Geis, & Berger, 1970). This study resulted in the Mach-IV, a 20 item inventory that captures the essence of his tactical recommendations in a series of statements (summarized: be cynical, unprincipled and deceiving) (Christie et al., 1970). Subjects who endorse Machiavellian statements are more likely to behave in a cold and manipulative manner (Christie et al., 1970). They are also more likely to lie (Forsyth, Anglim, March, & Bilobrk, 2021).

Psychopathy was brought into a subclinical sphere by Robert D. Hare (Hare, 1985), influenced by Hervey M. Cleckley's 1941 book *The Mask of Sanity*, which defines the psychopath using 16 traits belonging to one of two factors (Factor 1: callousness, Factor 2: the "live fast, die young" antisocial lifestyle) (Hare, 1980). Subclinical psychopathy reflects much of its clinical relative: selfishness, arrogance, superficial charm and impulsiveness (McLarnon & Tarraf, 2021). Although prototypically different (in regards to the sensitivity to the opinions of others, for example, psychopaths lean towards indifference whereas narcissists are highly concerned), psychopathy and narcissism often overlap clinically (Miller et al., 2010). The views on this are many. One view, set forth by Glenn & Sellbom, proposed that psychopathy

is the “main” construct, encompassing both narcissism and Machiavellianism (Glenn & Sellbom, 2015).

In more recent years, the inclusion of “everyday sadism” has turned the triad into a tetrad (Buckels, Jones, & Paulhus, 2013). People who score high on the Dark Tetrad have dark personalities, but are not extreme enough to warrant clinical or forensic attention (Paulhus, 2014). They are by their very definition quite “normal”. After all, ethically, morally, and socially questionable behavior – of which dark traits are to blame (Moshagen, Hilbig, & Zettler, 2018) – is not very hard to find.

How do we understand dark personality traits?

Manipulative and unethical behavior is better predicted by HEXACO measures than five-dimensional measures (Ashton & Lee, 2019). Paulhus & Klaiber have also endorsed the HEXACO framework, stating that it is “clearly superior” in terms of compatibility with the Dark Tetrad model (Paulhus & Klaiber, 2020).

There has been a tendency to view the Dark Triad as the opposite pole of the Honest-humility dimension (McLarnon & Tarraf, 2021), as empirical results clearly show that all three load on that factor (Furnham, Richards, & Paulhus, 2013). Yet according to McLarnon and Tarraf (2021), the Dark Triad offers something above and beyond simply inverting the Honesty-Humility scale. Hodson, Visser, Volk, Ashton & Lee (2018) maintain that the bundle of the three traits taken together reflects something that is meaningful on its own – a core (Hodson et al., 2018). All dark traits have something in common; the very “thing” or essence that makes it dark.

There are strong, positive correlations between the traits in the Dark Triad. The original correlations found by Paulhus & Williams were in the .25-.50 range (Paulhus & Williams, 2002). Correlations of equal and greater magnitude have since been reported. The highest mean correlations are found between psychopathy and Machiavellianism ($r = .61$), and the lowest between narcissism and Machiavellianism ($r = .25$) (Furnham et al., 2013). A meta-analysis by Muris et al (2017) reported the following correlations: $r = .58$ (Machiavellianism and Psychopathy), $r = .34$ (Machiavellianism and Narcissism) and $r = .34$ (Narcissism and Psychopathy) (Muris, Merckelbach, Otgaar, & Meijer, 2017).

This gives rise to a new question: what does the “core” represent? Originally, Paulhus and Williams (2002) proposed:

> "(…) members of the Dark Triad share a common core of disagreeableness. Thus the root of their social destructiveness is disturbingly normal - even banal." (p. 561).

This is not the first time that social destructiveness – evil, if you will – has been linked to banality. In her 1963 book *Eichmann in Jerusalem: A Report on the Banality of Evil,* Hanna Arendt writes about the trial of a Nazi officer in a way that confronts each of us with our own (often frail) morality (Arendt, 2012). Arendt viewed Eichmann as "terrifyingly normal"; not sinister – sane. A dull and thoughtless conformist incapable of critical thought, eager to climb the corporate ladder.

In a review article ten years later, Paulhus and associates maintain that the Tetrad consists of four separate "types" – the narcissist, the Machiavellian, the psychopath and the sadist – characterized, at the end of the day, by their shared *callousness,* rather than disagreeableness (Furnham et al., 2013). This was maintained a year later: "our restricted set might be called the callous constellation" (Paulhus, 2014). Callousness is often defined as being insensitive and hardhearted about other people's feelings (lacking in empathy, a deficit in emotional reactivity), whereas disagreeable is a somewhat more general term.

In the same review they warn against concept creep – "the tendency for researchers focused on a single construct to continually expand its scope" (Furnham et al., 2013). According to Haslam, concept creep follows a consistent trend (Haslam, 2016):

> Concepts that refer to the negative aspects of human experience and behavior have expanded their meanings so that they now encompass a much broader range of phenomena than before. This expansion takes "horizontal" and "vertical" forms: concepts extend outward to capture qualitatively new phenomena and downward to capture quantitatively less extreme phenomena. (p. 1)

In my opinion, the following proposal by Marcus & Zeigler-Hill (2015) illustrates the phenomenon well: "(…) we propose traits be judged to be dark based on their likely consequences." (p. 434). According to this view, traits qualify as dark "if they are regularly associated with problematic outcomes" (Marcus & Zeigler-Hill, 2015). This suggestion seems radical: to propose that dark traits should be defined based on their *likely consequences* rather than some fixed characteristic. It can be seen as pragmatic and relativistic in a utilitarian sense, and perhaps even useless at worst. The article makes an important point, though: the Dark Triad/Tetrad might be overly constricting, causing researchers to overlook important

personality traits that should be included in the tent of dark traits. In the article "A Big Tent of Dark Personality Traits", the authors call for a broader conceptualization of dark traits (Marcus & Zeigler-Hill, 2015). Coincidentally, a word often associated with utilitarianism – utility maximization – is at the core of a relatively new definition of dark traits.

D: The Dark Factor of Personality

Moshagen, Hilbig and Zettler answered the call for conceptual change by launching the Dark Factor of Personality in 2018. Simply referred to as *D*, Moshagen, Hilbig & Zettler made the claim that any dark trait should be regarded as a specific manifestation of a single construct: the Dark Factor of Personality (Moshagen et al., 2018). *D* is defined as "the general tendency to maximize one's individual utility — disregarding, accepting, or malevolently provoking disutility for others —, accompanied by beliefs that serve as justifications" (Moshagen et al., 2018).

According to Moshagen et al., *D* forms a theoretical basis explaining why dark traits are connected: through the common characteristic of inflicted disutility in pursuing one's own goals - whether intentional or not. In my opinion, this definition strikes a good balance between being inclusive and exclusive. It is broad and general enough to capture all the different kinds of dark traits and their severity (there is a large difference between passively accepting others' misery and actively deriving joy from it), without falling prey to concept creep.

As was the case for the H factor, the definition of *D* has a clear normative aspect, evoking the world of ethics and moral agents. The definition of *D* lends itself easily to a comparison with altruism – representing, perhaps, its antonym in both principle and moral practice. Altruism can be defined as promoting or prioritizing someone else's welfare at a cost to oneself. Some have viewed *spite* as the opposite of altruism –"representing two sides of the same coin" (Lehmann, Bargum, & Reuter, 2006), as a spiteful act harms the other with no (obvious) benefit to the self. Unless one derives utility from others' disutility, that is, or utility can be derived from performing the spiteful act itself.

D can be understood analogous to the *g*-factor of intelligence: an "underlying disposition responsible for the emergence of *any* particular dark trait" (Moshagen, Zettler, & Hilbig, 2020, my emphasis). As a fluid construct analogous to the *g*-factor of intelligence, any dark trait inventory will also give an indication of *D*, albeit indirectly.

Where does *D* fit in in the personality structure?

D is scattered throughout multiple traits, but is mainly to be found as low Honesty-Humility (H), low Agreeableness (A) and low Conscientiousness (C) in the HEXACO framework (Moshagen et al., 2018).

The exploitation of others associated with trait narcissism and psychopathy has been found to be much more strongly associated with low H than with low A, however (M. C. Ashton, Lee, & de Vries, 2014). This discovery would not have been made if compared with A in the Five-Factor model, but it is not particularly surprising when seen through the HEXACO lens. The HEXACO model clearly contrasts H and A, predicting different behavioral outcomes. Both are relevant to cooperation and reciprocal altruism, but there are important differences. The following excerpt highlights the characteristic difference between the traits (M. C. Ashton et al., 2014):

> High levels of H represent a tendency to cooperate with another person even when one could successfully exploit that individual, whereas high levels of A represent a tendency to cooperate with another person even when that individual appears to be somewhat exploitive (or not fully cooperative). The low-H form of uncooperativeness involves an approach to interactions whereby one seeks to cheat and defeat others for material and social-status gain, whereas the low-A form of uncooperativeness merely represents a disposition to respond sharply to any perceived exploitation by others. (p 146)

Dark Core Inventory

The Dark Core Inventory (DCI) is a self-report inventory assessing the dark factor of personality (*D*) (Moshagen, Hilbig, & Zettler, 2018). The inventory consists of 70 items (see Appendix A), available at http://darkfactor.org, as well as Open Science Framework. Thirty-five and 16 item short forms are also available.

The DCI was developed using rational item selection techniques to large samples (total N > 165,000) evaluating dark traits. Without going into further detail, rational item-selection uses powerful algorithms to select items (from an item pool) in a way that optimizes the final item set. Twelve dark traits were considered to ensure a broad enough sampling of the construct. Measuring a single dark trait will primarily reflect the trait in question, and secondarily reflect

D. Dark traits have *D* in common, but some other characteristic(s) that lies outside the scope of *D*, as well. For instance, an inventory measuring psychopathy will include some items on poor impulse control; yet poor impulse control on its own gives no indication of *D* (Moshagen et al., 2018).

Moshagen et al. (2018) noted that most existing dark trait inventories relied on positively keyed items and wanted to avoid potential response biases (e.g. passively agreeing) when developing the Dark Core Inventory. As the wording of a phrase "encodes the attitude towards the content" (Boase-Beier, 2006) there is a 50/50 balance between positively and negatively keyed items.

Twelve dark traits were measured using their established scales. The traits included by Moshagen et al. were:

- Psychopathy
- Narcissism
- Machiavellianism
- Sadism
- Greed
- Amoralism-Crudelia (amoralism involving brutality)
- Amoralism-Frustralia (amoralism caused by frustration)
- Egoism (excessive concern for oneself combined with little concern for others)
- Moral Disengagement (an attribution style that enables the individual to avoid blame and deny responsibility)
- Psychological Entitlement (the belief that one is more deserving than others)
- Self-centeredness (indifference to the suffering of others)
- Spitefulness (craving revenge at all costs)

As the items have been taken from a wide range of preexisting scales and inventories, naturally they vary in wording and style; ranging from Shakespearean style one-liners ("sweeter is the revenge that takes a long time to plan") and biblical grandeur ("I tend to forgive the wrongs I have suffered") to the almost puzzlingly straightforward "Actually, I'm kind of greedy".

Bifactor modelling, which hypothesizes one general factor as well as specific factors, supported the idea of a single general factor (see figure 1) (Moshagen et al., 2018). Several

new studies have supported this initial finding (Hartung, Bader, Moshagen, & Wilhelm, 2021; Moshagen, Zettler, & Hilbig, 2020).

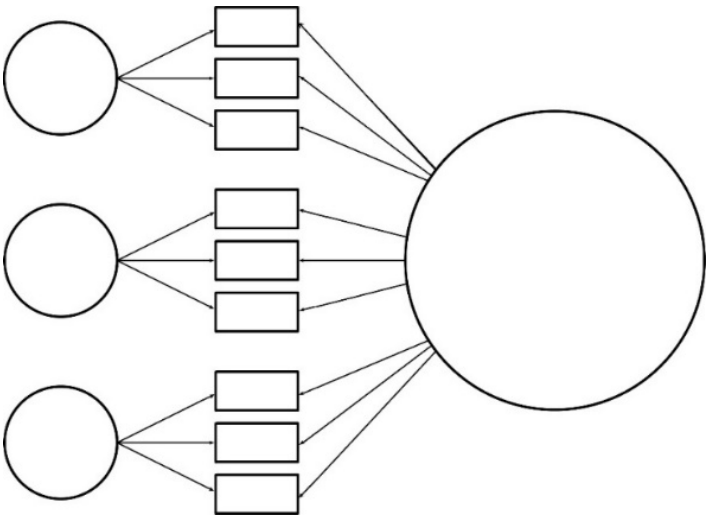

Figure 1. Illustration of a bifactor model displaying one general factor (right) with grouping factors (left). Variables are indicators of both the general factor and their respective grouping factor.

In a bifactor model, variables load onto at least two factors, as the name implies. The general factor is hypothesized to load directly on each of the observed variables (Dunn & McCray, 2020), giving the general factor greater theoretical weight compared to the grouping factors (Markon, 2019). The general factor captures covariance across all variables, whereas the grouping factors represent the remaining covariance among the variables after controlling for the general factor (Moshagen et al., 2018). Some researchers have suggested that a bifactor structure has the potential of providing greater conceptual clarity than other model structures, especially when evaluating subscales (Reise, Bonifay, & Haviland, 2018).

Because of *D*'s conceptualization, it makes theoretical sense to employ bifactor models when examining the D70 (Moshagen et al., 2020). The short versions (D35 and D16) are better suited for single-factor models, however, as the additional (grouping) factors derived from them don't capture much variance beyond *D* (Moshagen et al., 2020).

The results reported by Moshagen et al. indicate that all three versions of the DCI have good psychometric properties in terms of high internal consistencies (D70: $\alpha = .970$, D35: $\alpha = .950$, D16: $\alpha = .906$) and retest reliabilities (D70: $r = .95$, D35: $r = .93$, D16: 0.90) (Moshagen et al., 2020).

Based on criterion measures, the validity of the DCI is strong. It has been able to predict selfish and cheating behavior in famous behavioral economic experiments as the Dictator Game (Moshagen et al., 2018). In the Dictator Game, there really *is* such a thing as a free lunch. Player number one (the dictator) is handed some money and told he can split it with a stranger – if he feels like it. The reason why this "game" is so interesting to psychologists and behavioral economists alike, is that not only is it perfect for measuring altruistic sharing, you can literally count the difference between individuals in willingness to share. Moreover, because it is a zero-sum game (one player's gain is the other player's loss), the dictator's utility is maximized only when keeping it all for himself. Of course, humans are not Homo economicus, and most transactions and exchanges are not one-time anonymous situations like this (or zero-sum games, for that matter). However, the results are interesting all the same. A high *D* dictator might value only money. A low *D* dictator might draw utility from a broader and more "abstract" and prosocial (altruistic) pool of utility – he might derive utility from sharing. Utility can be anything that gives you dopamine, endorphins and serotonin. Some people get that warm and fuzzy feeling from seeing it in others. Perhaps the opposite of *D* is being genuinely happy for your fellow man, or deriving joy from sharing experiment pocket money?

Context for the thesis

The thesis is part of a research project at the Department of Psychology, University of Oslo led by Cato Grønnerød. The project *Personality and Distress: A Normative and Experimental Study of Response Modes and Robustness Against Faking in Self-Report Methods* will examine normative data and how well response distortion in the form of defensiveness or malingering is detected using a set of self-report methods. The methods examined in the research project are the Dark Core Inventory (DCI), the Personality Assessment Inventory (PAI), the Inventory of Problems (IoP-29), and Big Five personality trait measures (IPIP-NEO-120, NEO-PI-3).

The present thesis is based on a pilot study for the project *Personality and Distress*. The aim of the pilot study was to examine the psychometric properties of the newly translated Dark Core Inventory (DCI) with reference to the wording of the Norwegian translation, internal consistency and factor structure.

Method

Translation of the DCI

In 2019-2020, the DCI was translated into Norwegian. See Appendix B for the Norwegian translation. The inventory consists of 70 items available at http://darkfactor.org, as well as Open Science Framework (https://osf.io/...).

Cato Grønnerød did the initial translation. Two other students on the project – Pia Pran and Simon Rekkedal Rolfsnes and I received copies of the translation and added comments and suggestions. When we were in agreeance on all items, items were translated back into English by a professional translator. Reverse translation is advised to reduce loss caused by translation (Van de Vijver & Hambleton, 1996), as subtle shifts in wording can have an unintended effect. Reverse translation gives the added benefit of comparing the texts for quality and accuracy. The back translated version was given to Moshagen, who offered his comments based on his own preliminary translation into Norwegian.

Some items received more attention than others: item 24, 31, 40, 56 and 61.

Item 24: People who mess with me always regret it.
Translating slang is an example of when you have to allow some artistic license, as some slang has absolutely no meaning in another language. The first translation into Norwegian was "*folk som tuller med meg angrer alltid*". We opted to go for a wording that captures the more serious quality to the word "mess": *Folk som kødder med meg angrer alltid.*

Item 31: I would not cheat, even if there was only a small chance of getting caught.
We discussed whether to use the phrase "*bli oppdaget*" or "*bli tatt*", agreeing on the latter. The final item: *Jeg ville ikke jukset, selv om sjansen for å bli tatt var lav.*

Item 40: I make a point of trying not to hurt others in pursuit of my goals.
Hurt can be translated in a number of ways in Norwegian. *Skade* is most commonly used when speaking of physical pain. *Såre* is more commonly used when speaking of emotional pain. *Skade* can also refer to emotional pain, but it implies a more severe offense/transgression (e.g long term psychological abuse). We discussed whether to use *skade* or *såre*, deciding on the former. The initial wording was "*Jeg gjør et poeng ut av å ikke skade noen på jakt etter mine målsettinger*". Based on feedback from Moshagen, we changed the

item so that it more clearly stated that it was about achieving one's goals: *Jeg gjør så godt jeg kan for ikke å skade noen når jeg prøver å oppnå mine mål.*

It is worth noting "I make a point of trying" is different from "I try my best" which quite easily can let the mind wander to a justification ("I try my best, but if it can't be helped…").

Item 56: I would not pursue what I want if this causes problems for other people.
The word "pursue" can be translated to *forfølge* in Norwegian, but it is more often used in the context of stalking (someone). Final item wording: *Jeg ville ikke prøvd å få det jeg vil ha dersom det skaper problemer for andre.*

Item 61: For most things, there is a point of having enough.
This item is meant to tap greed, although it may appear a bit ambiguous. The original wording was "*For de fleste ting kommer det til et punkt der det er nok*". The translation was modified to *For de fleste ting kommer det til et punkt der man har nok.* so that the translation would not imply anything along the lines of being fed up. By mistake, the first version of the translation ended up in the final UFF questionnaire ("*For de fleste ting kommer det til et punkt der det er nok*").

Participants and procedure

The pilot studies shared a common data gathering framework. We used Nettskjema to electronically administer a questionnaire labelled Utvidet femfaktormodell (UFF). Nettskjema is a design tool for online surveys and data collection provided by the University of Oslo. One can reply from a web browser on a computer, a mobile phone or tablet. Nettskjema was developed by the University of Oslo to meet requirements for security and privacy.

UFF consisted of IPIP-NEO-120, D70, 32 items from Inventory of Interpersonal Problems (IIP) and 40 items from the Honesty-Humility scale of the HEXACO inventory – 262 items in total. The questionnaire used a five point likert scale format where participants had to choose from strongly disagree (*stemmer ikke*) to strongly agree (*stemmer svært godt*). The questionnaire was available at https://pop.utredning.info.

Participants were recruited via Facebook and Finn.no, as well as Psychology Departments of major Norwegian universities by two other students on the project and myself. The initial plan was to have psychology students enrolled in two courses at the University of Oslo to complete the questionnaire as a research participation exercise in exchange for course credit. The study

was accepted in the spring 2021 Research Participant Pool, but due to the Covid-19 pandemic, activities were not mandatory.

Psychology students at the University of Tromsø received an e-mail with an invitation to participate in research on personality. Psychology students from NTNU in Trondheim were reached through posting in a psychology student group on Facebook. E-mails were sent to BISO (the student organization of the Norwegian Business School BI), as well as two Norwegian media outlets (Aftenposten and NRK) to recruit participants, but nothing came of these attempts.

Ethical perspectives: informed consent and confidentiality

Testing was based on informed, written consent from the participants using a standard template issued by the Regional Committee for Medical Research Ethics. The consent was signed using Bank ID, a highly secure electronic ID for signing online documents. Participants were informed that they could withdraw their consent at any time without stating a reason for doing so. Participants were told in advance that feedback and monetary rewards would not be given, but they could opt to be eligible for the prize draw of one of 10 gift cards worth NOK 500. Participants were encouraged to avoid long breaks and to complete the entire questionnaire in one sitting in a quiet environment.

The web forms were set up for encrypted delivery to secure storage in Tjenester for Sensitive Data (TSD). TSD is the University of Oslo system for high security storage for sensitive data. Nettskjema provided a secure solution for data collection via the web, allowing for direct storage in TSD, thus meeting General Data Protection Regulation (GDPR) requirements.

Norsk senter for Forskningsdata (NSD) (the Norwegian center for research data) has approved the study (reference number 131597), as well as the Internal Ethics Committee at the Department of Psychology, University of Oslo (9327260).

Statistical Analyses

Statistical analyses were performed in IBM SPSS Statistics (version 27, 2020), except for Exploratory Factor Analysis in RStudio ("RStudio Team," 2020).

Positively phrased items on the DCI were reverse-keyed before analysis, so that a higher score indicated a higher level of the dark core trait for all items. Preliminary analysis revealed that the variables were positively skewed. Most participants responded “don’t agree” (*stemmer*

ikke) and "somewhat agree" (*stemmer litt*) on the items of the DCI indicating a dark core (e.g "make sure your plans benefit you, not others"), and "agree" (*stemmer godt*) or "strongly agree" (*stemmer svært godt*) on the 35 reverse coded items (e.g "most people deserve respect"). The mean value for all items combined was 1.84, which is somewhere between don't agree (*stemmer ikke*) and "somewhat agree" (*stemmer litt* and *stemmer en del*) on the items of the DCI indicating a dark core. The highest mean for an item was 3.68. The lowest mean for an item was 1.07.

This consistent response style translates to an overall low dark core score. This means that most of the respondents have a low *D* score, so the variables were positively skewed, not normally distributed. Log transformations were performed in an attempt to remedy this.

To determine the factor structure, dimension reduction using Exploratory Factor Analysis (EFA) and Parallel Analysis was performed, using the R packages psych (Revelle, 2021) and EFAtools (Steiner, 2020). First, the number of factors was decided using parallel analyses, then the solution was rotated using a bifactor rotation, in line with the theory on *D*.

Group comparisons were made for DCI sum scores across gender, age and region. Power analyses were performed in SPSS. A desireable level of power is 0,8, which gives an 80% chance of detecting an existing difference between groups (Cohen, 2003).

Exploratory Factor Analysis: Principal Axis versus Principal Component Analysis

Factor analysis (FA) is a set of different but related techniques. FA is a set of statistical methods used to identify underlying factors based on correlations between variables. The goal is to reduce a large set of variables to smaller clusters, ultimately finding the latent variables (the most fundamental, underlying dimensions) that explain the correlations between the variables (Gaur & Gaur, 2009). There are two main types of factor analysis: exploratory and confirmatory. A version of Exploratory Factor Analysis is Principal Component Analysis (PCA).

PCA is a method to find the underlying structure – the main (principal) components – of the data set. Although PCA strictly speaking yields components (and does not hypothesize a latent variable *causing* the correlations), many researchers use the word "factor" when reporting the results from a Principal Components analysis (Abdi & Williams, 2010).

The experts disagree on whether to use common FA or PCA. PCA yields a linear combination of variables (data reduction), whereas FA has an assumption of (and seeks to reveal) latent

variables underlying the data set. Although the theoretical assumptions differ, the methods often yield highly similar results. One exception is when there are weak correlations between items (DeVellis, 2012).

The assumption about a latent construct (*D*) made the choice of EFA appropriate for examining the factor structure of this sample. In addition, EFA was the method chosen by Moshagen et al., rendering comparisons more accessible.

When deciding on how many factors to keep in an EFA, the most common methods are the scree test, Kaiser-Guttman criterion (Eigenvalues >1) and Parallel Analysis. Parallel analysis is the preferred method by most researchers today (DeVellis, 2012).

Horn's Parallel analysis

Horn's Parallel Analysis (Horn, 1965) is a way to help guide the decision on how many factors to keep in an EFA (Hayton, Allen, & Scarpello, 2011). A Parallel Analysis generates random samples where there is no real factor structure. When deciding how many factors to retrieve, the researcher compares the Eigenvalues of the data set to the Eigenvalues in the randomly generated data. It differs from the Kaiser-Guttman criterion (KGC), which indicates that Eigenvalues greater than 1 are factors. The KGC can lead to a large overestimation of the number of factors (Zwick & Velicer, 1986).

The scree test has its name from the Old Norse word skriða (*skred* in Norwegian), as the shape of the plot, displaying Eigenvalues in a descending curve, often resembles the characteristic build-up of rock debris at the base of mountains (Douglas, 2021). When using the scree plot to determine the number of factors, one looks for the "elbow" of the graph. This method has received criticism for being too subjective. First, there can be multiple "elbows" in the plot. Different observers may have different opinions on which is the most prominent. Second, there is no single agreed upon standard scaling of the x- and y-axes, meaning that different software can produce different plots for the same data.

Parallel Analysis is regarded a more reliable and sober method for determining the number of factors (Zwick & Velicer, 1986). It has been referred to as "one of the most recommended procedures for the decision by many experts" (Lim & Jahng, 2019).

Still, it is worth noting that a parallel analysis can end up indicating too many factors when using common EFA. This is attributable to the fact that parallel analyses were developed in

the framework of PCA (DeVellis, 2012). Hence, I therefore ran Parallel Analyses using both EFA and PCA.

Validity

Validity takes several different forms that provide complementary evidence that help answer the question: "does my research make any sense?". The answer to which (as is often the case in the social sciences) is more closely related to a tedious "it depends" than a simple, clear cut yes or no.

There isn't a single measure of validity, but several different approaches when trying to establish validity for a test or personality inventory. The question of validity is not an all-or-none phenomena, but a "sliding scale" which can shift in either direction depending on many different parameters. To replicate the factor structure of the DCI, for example, would be one way of demonstrating validity for the DCI. Another approach to demonstrate validity, could be to replicate differences between groups. The published material on the DCI so far suggests that there are predictable differences in age and gender (Hartung et al., 2021). On average, men score higher than women, and *D* declines with age. As previous studies have found a relationship between *D*, age and gender, it may be taken as evidence for validity if I am able to replicate these results (find differences in the same direction as previous studies on *D*).

Reliability

Since this pilot study involved a one-time measurement using only one rater (self-report), reliability was examined by calculating Cronbach's alpha.

Cronbach's alpha

Developed by Lee Cronbach in 1951, Cronbach's alpha (α, coefficient alpha) reliability is a measure of internal consistency – how closely related a set of items are as a group (Cronbach, 1951). A rule of thumb for Likert scale questions is $0.9 \geq \alpha \geq 0.7$., ranging from excellent to acceptable (Tavakol & Dennick, 2011). A high α level may mean that the items are highly correlated, but the measure is also sensitive to the number of items in a test. A high α can indicate that there are redundant items in the test. A low α can indicate more than one latent variable. It is also worth noting that alpha is a characteristic of the sample (Schmitt, 1996). A sample that is biased or small could yield a very different value for alpha than a representative sample (Schmitt, 1996). One way of interpreting this statistic is that alpha is the average value for all possible combinations of a split-half correlation of a test or a measure – and also with increasing values with increasing number of items.

Results

Sample

The total number of respondents in Nettskjema was 380. Eighty-two respondents had not filled out the entire questionnaire and were excluded from the sample.

I also excluded three additional respondents from the analyses. Two respondents had suspiciously short response times (4 minutes 3 seconds, 5 minutes 17 seconds). One of them selected the same response option for all 262 items, indicating a non-serious response style. A third case was duplicated for unknown reasons.

The final sample consisted of 295 participants (75.5% female, 23.2% male, 1.3% preferred not to say). For anonymity purposes, the exact age of each respondent was replaced by six age categories with 10 year intervals (15 through 85) in the data set I received (see table 1).

Table 1

Number of participants in each age category

Age category	Number of participants	Percent
15-25	24	8.1
25-35	113	38.3
35-45	83	28.1
45-55	51	17.3
55-65	20	6.8
65-75	3	1
75-85	1	0.3
Total	295	100

Note. Mean: 33.07 years, *SD*: 11.25, N = 295.

The UFF questionnaire consisted of 262 items in total with the items from D70, IPIP-NEO-120, 32 items from Inventory of Interpersonal Problems (IIP) and 40 items from the Honesty-Humility scale of the HEXACO inventory. The questionnaire took on average 34 minutes to complete (median = 29, *SD* = 17.63). This response time was as expected, and indicates a serious response style (taking the time to read the questions carefully before selecting a response). Five respondents showed a long response time (defined as more than three hours) and were excluded when calculating average time spent on completing the questionnaire.

Exploratory Factor Analysis and reliability analyses

Parallel analysis was performed on the items from DCI, based on the correlation matrix, and using the 95th percentile Eigenvalues in the simulated data as threshold for keeping factors. This result indicated the presence of eight factors (n = 295). A bifactor rotation was then performed on these eight factors.

For comparison reasons, the analysis was run using PCA in addition to EFA. This method also rendered eight factors.

Cronbach's alpha estimate of internal consistency for DCI was $\alpha = 0.933$.

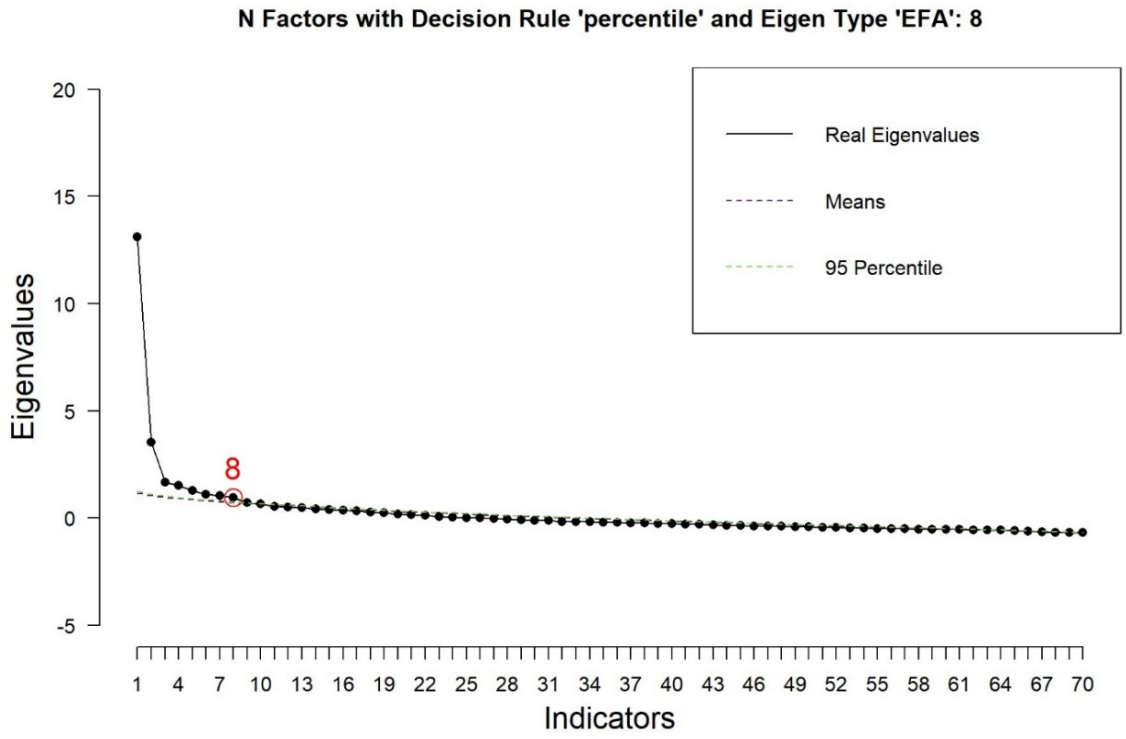

Figure 2. Scree plot based on parallel analysis using EFA. This scree plot indicates eight factors.
The eight latent factors explain 36.86% of the total covariance of the original data.

The eight factors accounted for 36.86% of the total covariance. One large factor (shown in figures 2 and 3), explained 18.73% of the covariance. The remaining factors explained between 2.19% and 3.18% of the covariance. Hence, the first factor explained much of the covariance, and further factors made only limited contributions to explaining further covariance.

In a bifactor rotation, the first factor represents the core, and the other seven factors are the specific factors. The first factor that emerged in the bifactor rotation can be interpreted as *D*. Most of the items show factor loadings onto to this first factor. All items load onto factor 1 except for CRUD1 (*It is hard for me to see someone suffering*), CRUD5 (*I am willing to volunteer for people in need*), GRE4 (*For most things, there is a point of having enough* (*For de fleste ting kommer det til et punkt der det er nok*[1])) and PENT1 (*If I were on the Titanic, I would not deserve to be on the first lifeboat any more than anyone else*). Table 2 shows factors 2-8 with their respective items. See Appendix D to see which item corresponds to each variable. Table 3 shows the results from the EFA with bifactor rotation.

Table 2. Factor overview

	Factor						
	2	3	4	5	6	7	8
Variable	FRUST5	EGO5	CRUD6	EGO4	PATHY4	NARC1	EGO3
	SPITE2	FRUST3	SAD2	FRUST2	PATHY5	NARC2	FRUST1
	SPITE5	GRE1	SAD5	MDIS1			PATHY7
		GRE2	SAD7	MDIS2			
		GRE3	SCTR1				
		MACH2	SCTR4				
		MACH7					
		PATHY2					
		SCTR2					

Note. Displaying positive factor loadings only. The variable names: CRUD refers to Amoralism-Crudelia, EGO to Egoism, FRUST to Amoralism-Frustralia, GRE to GREED, MACH to Machiavellianism, MDIS to Moral Disengagement, NARC to Narcissism, PATHY to Psychopathy, SAD to Sadism, SCTR to Self-Centeredness, and SPITE to Spitefulness, PENT (not displayed) to Psychological Entitlement.

[1] The correct translation was *«For de fleste ting kommer det til et punkt der man har nok.»*

Table 3

Factor loadings for eight factors (bifactor rotation)

VARIABLE	FACTOR							
	1	2	3	4	5	6	7	8
CRUD1								
CRUD2	0.32							
CRUD3	0.41					-0.37		
CRUD4	0.62							
CRUD5		-0.31						-0.30
CRUD6	0.43			0.43				
CRUD7	0.43					-0.48		
EGO1	0.44							
EGO2	0.48							
EGO3	0.39							0.42
EGO4	0.49				0.44			
EGO5	0.44		0.32					
FRUST1	0.32							0.45
FRUST2	0.49				0.50			
FRUST3	0.51		0.38					
FRUST4	0.37							
FRUST5	0.48	0.49						
FRUST6	0.44							
GRE1	0.33		0.35					
GRE2	0.34		0.36					
GRE3	0.34		0.31					
GRE4								
MACH1	0.34							
MACH2	0.48		0.31					
MACH3	0.37							
MACH4	0.40							
MACH5	0.56							
MACH6	0.61							
MACH7	0.35		0.32					
MDIS1	0.40				0.30			
MDIS2	0.37				0.37			
MDIS3	0.37							
MDIS4	0.45					-0.4		
MDIS5	0.55							
NARC1	0.36						0.43	

Table 3 (continued)

VARIABLE	FACTOR							
	1	2	3	4	5	6	7	8
NARC2	0.52							
NARC3	0.42							
NARC4	0.32						0.49	
NARC5	0.49							
NARC6	0.44							
PATHY1	0.39							
PATHY2	0.48		0.35					
PATHY3	0.41							
PATHY4	0.48					0.3		
PATHY5	0.41					0.3		
PATHY6	0.5							
PATHY7	0.43							0.35
PENT1								
PENT2	0.43							
PENT3	0.44							
PENT4	0.45						-0.40	
PENT5	0.38							
SAD1	0.42							
SAD2	0.60			0.49				
SAD3	0.38							
SAD4	0.36							
SAD5	0.47			0.53				
SAD6	0.45							
SAD7	0.56			0.43				
SAD8	0.56							
SCTR1	0.45			0.36				
SCTR2	0.44		0.38					
SCTR3	0.46							
SCTR4	0.41			0.40				
SPITE1	0.49							
SPITE2	0.53	0.35						
SPITE3	0.51							
SPITE4	0.38							
SPITE5	0.34	0.44						
SPITE6	0.31	-0.58						

Note. Factor loadings for eight factors using Bifactor rotation.

Figure 3 shows the bifactor solution. Factor 2, 6, 7 and 8 are *as shown* in figure 3. Additional items load on factors 3-5 (shown in table 2 and 3).

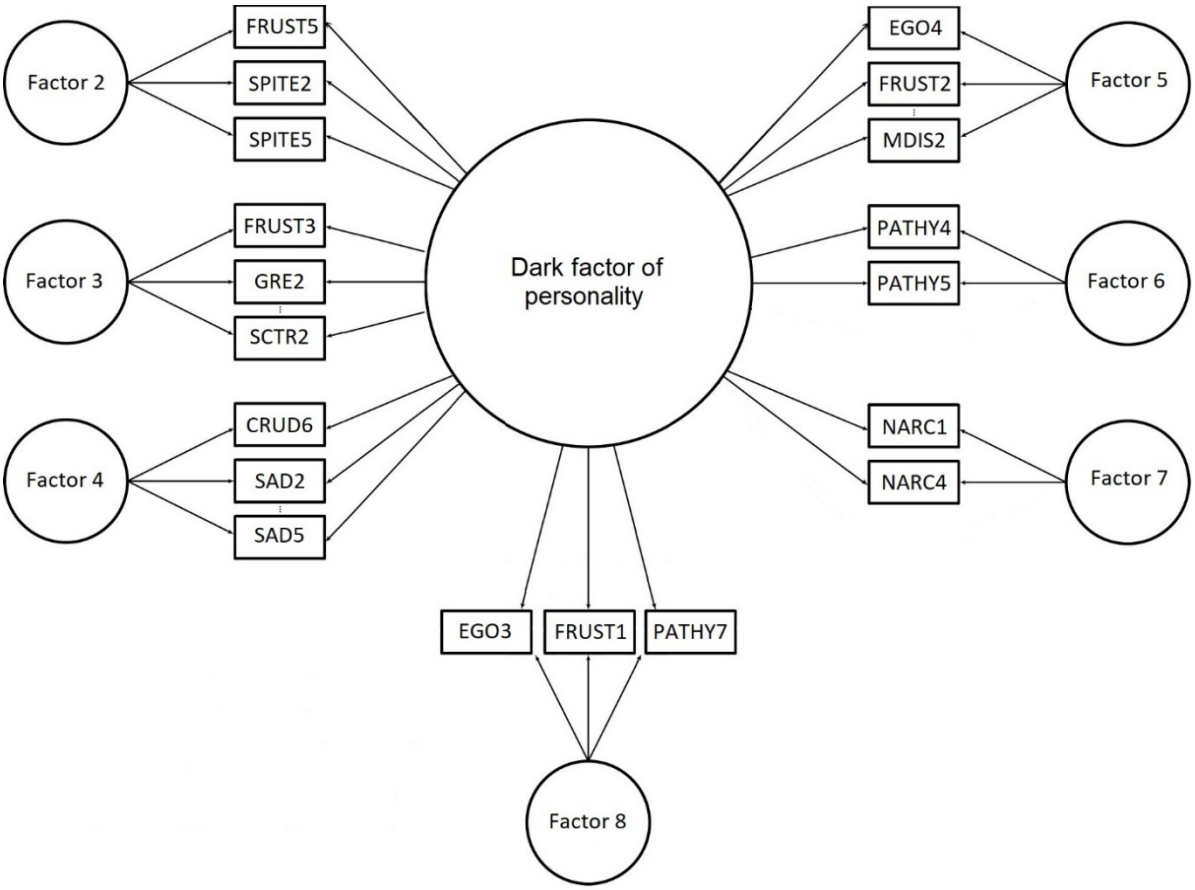

Figure 3. A maximum of three items per factor are displayed to enhance clarity. Item omission is indicated by a dotted line between the middle and last item. EGO5, GRE1, GRE3, MACH2, MACH7 and PATHY2 omitted from factor 3. SAD7, SCTR1 and SCTR4 omitted from factor 4. MDIS1 omitted from factor 5.

D and its relationship with demographic data

Comparisons of the mean DCI scores were performed across gender, location of residence and age (living in a city versus living in a rural area). Differences in mean scores on the DCI were found across all groups.

Gender differences

Four participants did not disclose their gender. These were excluded from analysis. As shown in table 2, on average, men scored higher than women. Power analysis suggested that I would

have 80% chance of detecting a real group difference of Cohen's D = 0.39 in DCI scores between men and women.

Table 4

Group differences in mean DCI score

		N	Mean	SD	SE Mean
DCI70	Female	224	125.45	28.43	1.89
	Male	67	141.15	32.60	3.98

Note. Significant (2-tailed) at the 0,05 level.

Country versus city dwellers

The sample consisted of 256 persons living in a city (defined as having a population of more than 5000 people) and 39 persons living in a rural area (a population of 200 or less). There was a significant difference in mean score on the DCI. People living in a city scored higher (a mean of 130.4) than people living in a rural area (a mean of 119.2) (Cohen's d = 0.37).

Power analysis suggested that I would have 58% chance of detecting a real group difference of Cohen's d = 0.37 in DCI scores between city and country dwellers.

Table 5

Group differences in mean DCI score

		N	Mean	SD	SE Mean
DCI70	City	256	130.37	30,87	1.93
	Rural area	39	119.23	21,81	3.49

Note. Significant (2-tailed) at the 0,05 level.

A linear regression analysis was run to examine gender and region differences in DCI scores, controlled for each other. The DCI score variable was standardized and used as an outcome in this analysis, while gender and region were used as predictors. Using dichotomous predictors and standardized outcomes allows interpreting the regression coefficients as Cohen's ds (J. Cohen, Aiken, Cohen, & West, 2003). The results showed Cohen's d = 0.206 for gender and d = -0.129 for region, both significant at the 0.05 level. Hence, the differences persisted after controlling them for each other.

Age differences

Age was measured using age groups with 10 year intervals. A dichotomous variable was made by splitting the sample in two. To ensure enough participants in each group cut-off was set at 35 years of age. One group consisted of the participants aged 15-35 years (n = 137), with the remaining participants in the other group (n = 158). No significant group mean DCI difference was found.

When treating the variable as a continuous variable however, a significant difference was found. Using the standardized DCI score variable as an outcome in the regression analysis, the DCI score was found to decrease with increasing age (Cohen's d = -0.011), significant at the 0.05 level. For every ten years increase in age, the standardized log transformed DCI score decreases with 0.011 standard deviations.

Items from the translation phase: revisited

Some items were more difficult to translate than others, shown by the lack of initial agreement among us. To see if this could be detected by the way people have responded, I examined each of these items in terms of how they compared to the total mean score, as well as the general distribution of responses. The items we had trouble translating were the following: 24, 31, 40, 56 and 61.

Item 24: *People who mess with me always regret it* (*folk som kødder med meg angrer alltid*). A measure of psychopathy. With a mean of 1.62 it is only slightly lower than the average of 1.84. 24% of the sample responded "somewhat agree" (*stemmer litt*) to this item. 10% responded "*stemmer en del*" (agreeing slightly more than "somewhat") .

Item 31: *I would not cheat, even if there was only a small chance of getting caught* (*jeg ville ikke jukset, selv om sjansen for å bli tatt var lav*.) This is a Machiavellian statement, but not among the most extreme. This is reflected in an even distribution of responses (all of the response categories have many hits) and consequently, a higher than average mean of 2.62. The translation disagreement concerned whether «*bli oppdaget*» or «*bli tatt*», was the better option. Compared to other translation *faux pas*, I view this as a minor issue.

Item 40: *I make a point of trying not to hurt others in pursuit of my goals* (*jeg gjør så godt jeg kan for ikke å skade noen når jeg prøver å oppnå mine mål*). A measure of psychopathy. With a mean of 1.56, it is pretty close to the mean for all items. 91% of the sample responded "strongly agree" and "agree".

Item 56: *I would not pursue what I want if this causes problems for other people* (*jeg ville ikke prøvd å få det jeg vil ha dersom det hadde skapt problemer for andre*). 60% of the sample agree with this statement, 18% only somewhat agree. 22% of the sample disagree, elevating the mean to 2.41.

Item 61: *For most things, there is a point of having enough*. This item was one of the items Moshagen double-checked for accuracy. As described earlier, we decided on the translation *For de fleste ting kommer det til et punkt der man **har** nok*. Unfortunately, the earliest version of the translation was used in the questionnaire. The Norwegian sample answered "*For de fleste ting kommer det til et punkt der det **er** nok"*. With a mean of 3.68, this is the highest item mean on the entire inventory. Close to 82% of the sample have disagreed with this statement. Only 18% have endorsed it. This item is supposed to tap greed, but can be interpreted as being "fed up". People have either not understood the item, or perhaps thought it is about "putting one's foot down". See Appendix D, figure 4, for a graphical display of the distribution. I will return to this result in the discussion section of my thesis.

In comparison, with a mean of 1.07, almost no one endorsed item 55. This was a measure of moral disengagement: *Folk som blir mishandlet har vanligvis fortjent det* (People who get mistreated have usually done something to bring it on themselves). This was the most skewed item on the entire inventory (see Appendix D, figure 5, for the response distribution.). Every person in the sample, with very few exceptions, have responded "strongly disagree". Not a single person selected the "strongly agree" option, but three selected "somewhat agree" and three selected the "agree" option. I will also return to this result in the discussion section.

Discussion

Overall, the sample has responded as expected, with many participants showing their "light" side while rejecting their dark side. Still, I identified one large factor in the Norwegian sample that can be interpreted as the dark core, *D*, in addition to seven small factors (corresponding items shown in Appendix D).

According to the latest research on *D*, the best measurement model is a bifactor structure which models *D* as well as five grouping factors: Callousness, Deceitfulness, Narcissistic Entitlement, Sadism, and Vindictiveness (Bader et al., 2021). I was not able to find these exact factors in my sample. Still, it is possible to interpret Factor 2 as a Vindictiveness factor. In my opinion, vindictiveness is the overarching theme of the items loading onto this factor. Factor 3 is most closely aligned with the deceitfulness theme (although it has items concerned

with greed, as well – “greed driven deceit” could be one interpretation). Factor 4 can be understood as the Sadism factor. Factor 7 is has two items from the narcissism subscale (and can be understood as Narcissistic Entitlement), whereas factors 6 and 8 are both contenders for the Callousness factor, in my opinion. Factor 5 is the factor that doesn’t easily align with any of the grouping factors suggested by Bader et al. (2021). I labeled factor 5 “Integrity”.

If there was more variance in *D* in my sample, perhaps I would have found results more line with the themes proposed by Bader et al. (2021). In sum, I was not able to differentiate between the five “different kinds of dark”. I was, however, able to detect some of the differences predicted by Dark Factor theory. According to Dark Factor theory, on average, men have a higher *D* score than women (Hartung et al., 2021), which was replicated in this study. The results from this pilot study suggests that individuals living in a city have a higher score on the DCI than people living in a smaller community. Trying to make sense of this result, it is possible to imagine that it is easier to “hide” in the city and thus get away with dark behavior. The conditions are not optimally suited for dark behavior in small communities. Smaller communities are more transparent, perhaps increasing solidarity between its members, or at least making it easier to hold people accountable for their behavior. Burning bridges has real consequences when the number of bridges are limited. In the city, bridges are, in fact, almost infinite. Your reputation needs never catch up with you. The direction of causality might also be in the other direction: living in a city might exert a specific influence (bring out dark tendencies). The same individual might return a wallet with its contents if he lived in a small town, but keep the contents after some time spent living in a city.

In a way, how one responds on the Dark Core Inventory can also be said to be an exploration of what we value and how we relate to others. Based on Dark Core theory, *D* decreases with age (Hartung et al., 2021). If you assume that as people age, they become more mature (e.g. more patient, forgiving and level-headed, less petty, small minded and vengeful), this makes sense. A decreasing *D* level across the lifespan is supported by the research that exists on change across the lifespan in Big Five traits. A massive body of research shows how Big Five personality traits change across the lifespan. This research consistently finds that N, E and O decline with age, whereas and A and C increase (McCrae, 2002). In short, we grow less neurotic, more well-rounded.

This pilot study has administered a version of the DCI where its items were scattered throughout a larger questionnaire with 192 additional items (262 items in total). The factor

analysis of the DCI is based on the items in this context. This could have impacted the results. Maybe there are different results depending on whether 1) the DCI is administered on its own, 2) the DCI is sprinkled into another (single) inventory 3) the DCI is scattered throughout multiple inventories (as was the case in this pilot study). I cannot rule out the possibility that the different ways of presenting the questions of the DCI might have impacted how people respond. An advantage of presenting the items this way is that it is very difficult for the responder to make a hypothesis about which specific traits we are studying, making "strategic" responding more difficult.

A questionnaire with 262 questions is a lot to consider. The fact that some people took breaks (shown by a response time of several hours) indicates this. Future studies on Norwegian samples should consider administering the DCI on its own to see whether and/or how this affects the results.

When asking the kinds of questions the DCI asks, one cannot expect to get normally distributed data. Many of the items are quite extreme. Most people will not endorse them (e.g "I think about harassing others for enjoyment" – although 3 respondents did, in fact, agree with this statement). Log transformation on sum scores helps somewhat, but will not be able to change the fact that some items have only a handful of people who have endorsed the item.

One solution for highly skewed data from Likert scales is to treat the variables as categorical (Rhemtulla, Brosseau-Liard, & Savalei, 2012). This implies treating the observed values as manifestations of a normally distributed continuous underlying variable (or alternatively as if a non-normal underlying continuous variable was log-transformed) (Greer, Dunlap, & Beatty, 2003). Associations between the underlying variables can then be estimated with statistical packages (Gustavson, Røysamb, & Borren, 2019), such as the psych package in RStudio. Treating variables as categorical is computationally demanding, particularly with small samples and very few observations on some of the response options (Rhemtulla et al., 2012). Preliminary analyses in RStudio were done to examine the factor structure with the DCI items treated as categorical variables. However, the computational complexity of this did not match the available information in the relatively small sample, and the solution did not converge. This is a pilot study, so in the future the sample size will be larger, possibly allowing running more complex analyses.

The limitations associated with a small sample size will become less of a problem in the future, but had implications for this thesis nonetheless.

Item revisions

I wonder what would happen if some of the items were worded differently – if some minor changes were made (e.g. some slight reservations added) to lessen ambiguity and add "reasonability". Item 14 can serve as an example: *I believe that lying is necessary to maintain a competitive advantage over others.* This is a Machiavellian statement. The point of this item, the way I see it, is to figure out if someone thinks lying can be justified when it makes strategical sense to do so. Because of the wording, however, I suspect that people can disagree with this item for a number of reasons. People who disagree might do so because they think lying is wrong, but disagreeing can also point to something else entirely. Someone can disagree because they think that lying isn't "necessary". If the item was changed to *I believe that lying is sometimes necessary to maintain a competitive advantage over others*, I think the item would better measure what it is intended to. Without "sometimes", some might subconsciously read, or interpret, an *always* between the lines. Someone might disagree not because he is against strategically well-placed lies, but because in his experience, lying is rarely needed (e.g. "I'm no stranger to lying, but I'm usually smart enough to maintain my competitive advantage without resorting to lying."). A person with certain narcissistic tendencies could say "I'm usually so far ahead that lies are redundant". This is exaggerated to enhance my point, but I think the essence of the item (believing that there are circumstances in which lying to keep ahead is okay) is better represented by slightly altering it. In my opinion, adding *sometimes* softens the content of the item, yet an endorsement would still indicate a questionable character.

This is done on item 55: *People who get mistreated have usually done something to bring it on themselves*. The inclusion of the word *usually* will probably make this statement work as a justification and make the reader think of a time they hurt someone "who deserved it". Whereas *People who get mistreated have done something to bring it on themselves* might work differently (e.g. act as a trigger for a memory when the reader was subject to some unfair treatment).

It is worth repeating that the Dark Core Inventory aims to measure dark traits in the subclinical population. This point can be repeated: researchers wish to reveal the dark tendencies in the "normal" population – one is curious about the level of *D* in Jane and Joe, not in *the Joker*. Because of the target population, it is important, in my opinion, that items are "soft" enough to capture variance related to dark *tendencies* (as opposed to their clinical

relatives). Social desirability bias is one thing, but self-serving biases – the subtle and unconscious ways in which we trick ourselves into believing we are better people than we actually are – are even trickier. For this reason, I believe it important that items intended to measure subclinical phenomenons are "mild" enough.

Ambiguous items

The Dark Core Inventory is not a projective measure. There should not be much room for interpretation when reading the items, as we have no way of understanding the reasoning behind the answers; no way to ask follow up questions to hear people's thinking or moral reasoning. I believe some items might come across as a bit too ambiguous. I will shortly discuss these items.

The item that was translated incorrectly ("For most things, there is a point of having enough") stood out in the analysis for two reasons. First, it did not show a factor loading to any of the factors. Secondly, it showed the highest item mean on the entire inventory. When interpreting this result, one must bear in mind that the Norwegian version of the item does not reflect the original one. The sample has answered a different question than intended. I do not think this unusual distribution of responses was a coincidence, and I suspect there were more than one confused participant asking himself what the item was referring to.

Item 44: *I try to look out for myself first, even if it means making things difficult for other people.* What kinds of situations might this refer to? This item might be too general to conjure an example, perhaps thus making it more difficult to answer. When contrasted with item 7, *If I were on the Titanic, I would not deserve to be on the first lifeboat any more than anyone else*, it becomes clear how details impact vividness. Even if you haven't seen the 1997 film by James Cameron, where a man wrongfully secures himself a spot on a lifeboat by kidnapping a young girl, everyone has heard about the sinking of the Titanic.

When translating the DCI into Norwegian and comparing the two versions of the items, one becomes aware of some general issues concerning translations. Translating confronts you with many dilemmas; of choosing between accuracy in a strict sense and accuracy in "ambiance"; the unique mood created by a particular sentence or statement. It is crucial to stay true to the style of the text and its "perceived distinctive manner of expression" (Boase-Beier, 2006), but as phrased by Boase-Beier, "such stylistic loyalty is hardly straightforward". Word by word, the translation may seem fine, but in its totality, something may be off. Items

may simply give off a different “vibe” in Norwegian compared to English. All the while, it is hard to pinpoint exactly what is missing – in short, what has been lost in translation.

Items that seem unproblematic in English, sometimes appear more tricky when read in your mother tongue. The challenges are illustrated in the translation of item D6: *payback needs to be quick and nasty*. Although this was not among the original items the group disagreed on, the item can serve as an example of the general issues we had when translating. This item has been translated to *gjengjeldelse bør være rask og stygg*, which, although it gets the job done, isn’t very sleek or catchy. Some may think it sounds old fashioned because of the word “gjengjeldelse”, which is rarely heard in informal everyday use. In my opinion, *gjengjeldelse* is closer to the word *retaliation*, even if “payback” is closer in a literal sense. “*Gjengjeld*” is originally a Danish word and means to pay back (gjeld meaning debt), but has come to mean “to react with an action or feeling of the same kind”. One could choose the Norwegian word “*hevn*” instead, but then *revenge*, would be the back translation. There is no precise translation of the word “nasty” either, so in this case one could be tempted to make a new item that retains the essence rather than the exact wording. In this particular case, I wonder if it had been best to choose *hevn* and be willing to sacrifice the back translation to gain more appeal in Norwegian.

It seems that something is lost in translation either way – and in the end one simply has to make a choice that reflects the personal opinion or “gut feeling” of the translator. I have some suggestions for minor translation revisions for the following items: item 2, 34, 40, 48, 55, 61 and 66.

Item 2, *All in all, it is better to be humble and honest than important and dishonest*, was translated to “*Når alt kommer til alt er det bedre å være ydmyk og ærlig enn viktig og uærlig*. One could consider “*i det store og hele*” over “*Når alt kommer til alt*”.

Item 34, *I cannot imagine how being mean to others could ever be exciting*, was translated to *Jeg kan ikke forestille meg at det kan være spennende å være ondskapsfull mot andre*. I believe that *slem* would be a better translation of “mean”, as *ondskapsfull* leans more in the direction of “evil”. In my opinion, “mean” and *slem* are closer relatives in terms of severity.

Item 40, *I make a point of trying not to hurt others in pursuit of my goals*, was translated to *Jeg gjør så godt jeg kan for ikke å skade noen når jeg prøver å oppnå mine mål*. Suggestion: “nå målene mine” instead of “oppnå mine mål”.

Item 48, *I would like to make some people suffer, even if it meant that I would go to hell with them* was translated to *Jeg ville nyte å få enkelte mennesker til å lide, selv om det betyr at jeg havner i helvete sammen med dem.* I suggest using the phrase "*jeg kunne tenke meg å få (…)*" instead of "*jeg ville nyte*". *Nyte* is closer to "I would enjoy" or "it would please me". I believe the current Norwegian wording is a nuance darker than its English counterpart.

Item 55, *People who get mistreated have usually done something to bring it on themselves*, was translated to *Folk som blir mishandlet har vanligvis fortjent det.* "Done something" is missing from the translation. One can consider updating this item to *folk som blir dårlig behandlet har vanligvis gjort noe for å fortjene det.* I believe the Norwegian verb "*mishandle*" is more aligned with the English verb "abuse". This pulls the Norwegian item in a darker direction than the original. *Mishandling* will often refer to acts punishable by law and usually involves ill intent. In English, however, "mistreatment" is a broader category (more of an umbrella term; *to treat wrongly*). You can mistreat someone by committing a sin of omission (as in neglect), but abuse more often involves a sin of commission (a prototypical "active" action), as opposed to *refraining* from doing or providing something.

Item 61, *For most things, there is a point of having enough*, was translated to *For de fleste ting kommer det til et punkt der man har nok.* This translation is more correct than *For de fleste ting kommer det til et punkt der det er nok* (the item that ended up in the questionnaire by mistake*).* "There is a point of having enough" – does that imply reaching a level (a point) where one has accumulated enough of something, or does it refer to a moral value (as in "*I make a point out of* reminding myself that I have enough ("greed is *not* good")? This aspect of the original item appears a bit a bit ambiguous to me and should perhaps be examined when updating the Norwegian version of the DCI. Overall, it is my opinion that greed is operationalized too narrowly; three out of four items refer to an amount. Perhaps an item concerned with sharing could be considered? Greed makes people not particularly keen on sharing.

Item 66, *If I ever tormented others, I felt strong remorse*, and item 8, *When I get annoyed, tormenting people makes me feel better*, both contain the verb "torment". In this case, torment has been translated differently in each of the cases (*plage* and *pine*). In my opinion, it's not a point of its own to keep the translation consistent, but in this case I argue that it would be better to use the word *plage* in both cases. *Pine* seems too dark, even if it's closer in a literal sense (*pine* is associated with torture, as is torment).

Some items sound “over-the-top” in both languages (e.g *doing good deeds brings joy to the heart* sounds rather cartoonish and ingenuine – it’s hard to deliver this line and keep a straight face) – overall, it is my impression that the English version seems more informal and has better usage of slang. It uses more available language for the common population, whereas several of the Norwegian items have a heavier or more serious ring to them.

Perhaps this view is merely a reflection of my shortcomings as an English speaker, though; the bluntness of the items appear more stark and vivid *exactly* because Norwegian is my mother tongue.

Concluding thoughts

Do the conclusions extend beyond the sample? General conclusions are usually the goal with research, the sample is merely a means to that end. Do the conclusions ring true for the specific group of people included in the sample, or can the conclusion be used to make a general statement that rings true for a larger group of people?

Individuals were asked to participate without any feedback or reward (other than a chance to win a gift card). Simply appealing to the “goodness of people’s hearts” when recruiting participants for a study on dark personality is not ideal. The “what’s in it for me” types are hard to reach when the research methodology is based on voluntary participation without reward. In comparison, the research on the dark factor of personality by Moshagen et al. on darkfactor.org involves a reward in the form of feedback, which adds to its appeal almost exponentially when compared to having no feedback.

Ideally, the sample should have included more variance on the dark factor. Reaching the target group of “dark individuals” proved quite difficult. Participating in this study can be viewed as an act of altruism (quite ironic given the subject matter), potentially limiting the external validity of the results. Self-selection might have led to a “light” sample.

The sample should have been more balanced with respect to gender and age. On one side, having a quite young sample is an advantage when measuring *D*, as the literature indicates a slow and steady decline in *D* across the life span (Hartung et al., 2021). On the other hand, the sample was imbalanced with respect to gender representation. Men have higher scores on dark traits (Hartung et al., 2021), so there is reason to believe that having a largely female sample resulted in variance reduction.

Some of the analyses in this pilot study had low power (well below 0,8). Results from this pilot study need to be replicated in a bigger sample. Due to limited time and resources, a convenience sample was deemed satisfying for the pilot study. Down the line, however, the main project *Personality and Distress* will have more diverse and strategic samples (including a planned prison population sample), ensuring ample variance in *D*.

Self-report measures can suffer from social desirability bias, the tendency to portray oneself in an exaggerated positive light (Nederhof, 1985). Research shows that when the testing conditions are "low stake" and ensure the anonymity of its participants, there is no incentive to exaggerate levels of socially desirable characteristics and self-report measures can be trusted (M. C. Ashton et al., 2014).

If participation is driven by a desire to gain insight (e.g. being motivated by feedback on the personality inventory), there is nothing to gain by responding untruthfully. Perhaps some of the "rough, tough *just-can't-get-enough"* guys want to understand themselves better. Feedback can be considered in future studies.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Comp Stat, 2*(4), 433-459. doi:10.1002/wics.101

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs, 47*(1), i-171. doi:10.1037/h0093360

Arendt, H. (2012). Eichmann in Jerusalem: A Report on the Banality of Evil. *The Political quarterly (London. 1930), 83*(s1), 349-350. doi:10.1111/1467-923X.02404_14

Ashton, & Lee. (2005). Honesty-Humility, the Big Five, and the Five-Factor Model. *Journal of Personality, 73*(5), 1321-1354. doi:https://doi.org/10.1111/j.1467-6494.2005.00351.x

Ashton, M. C., & Lee, K. (2008). The prediction of Honesty–Humility-related criteria by the HEXACO and Five-Factor Models of personality. *Journal of research in personality, 42*(5), 1216-1228. doi:10.1016/j.jrp.2008.03.006

Ashton, M. C., Lee, K., & de Vries, R. E. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality factors: A review of research and theory. *Pers Soc Psychol Rev, 18*(2), 139-152. doi:10.1177/1088868314523838

Bader, M., Hartung, J., Hilbig, B. E., Zettler, I., Moshagen, M., & Wilhelm, O. (2021). Themes of the dark core of personality. *Psychol Assess, 33*(6), 511-525. doi:10.1037/pas0001006

Boase-Beier, J. (2006). *Stylistic approaches to translation*. Manchester: St. Jerome Pub.

Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral Confirmation of Everyday Sadism. *Psychol Sci, 24*(11), 2201-2209. doi:10.1177/0956797613490749

Cattell, R. B. (1943). The description of personality: basic traits resolved into clusters. *Journal of abnormal and social psychology, 38*(4), 476-506. doi:10.1037/h0054116

Christie, R., Geis, F. L., & Berger, D. (1970). *Studies in Machiavellianism*. New York: Academic Press.

Cohen. (2003). A power primer. In A. E. Kazdin (Ed.), *Methodological issues & strategies in clinical research* (pp. 427–436): American Psychological Association.

Cohen, J., Aiken, L. S., Cohen, P., & West, S. G. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*(Third edition. ed.).

Costa Jr, P. T., & McCrae, R. R. (1992). Reply to Eysenck. *Personality and individual differences, 13*(8), 861-865. doi:10.1016/0191-8869(92)90002-7

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. doi:10.1007/BF02310555

Crowe, M. L., Lynam, D. R., Campbell, W. K., & Miller, J. D. (2019). Exploring the structure of narcissism: Toward an integrated solution. *J Pers, 87*(6), 1151-1169. doi:10.1111/jopy.12464

DeVellis, R. F. (2012). *Scale development : theory and applications* (3rd ed. ed. Vol. 26). Thousand Oaks, Calif: Sage.

Digman, J. M. (1997). Higher-Order Factors of the Big Five. *J Pers Soc Psychol, 73*(6), 1246-1256. doi:10.1037/0022-3514.73.6.1246

Douglas, H. (2021). Online Etymology Dictionary. Retrieved from https://www.etymonline.com/word/scree#etymonline_v_22980

Dunn, K. J., & McCray, G. (2020). The Place of the Bifactor Model in Confirmatory Factor Analysis Investigations Into Construct Dimensionality in Language Testing. *Frontiers in psychology, 11*, 1357-1357. doi:10.3389/fpsyg.2020.01357

Forsyth, L., Anglim, J., March, E., & Bilobrk, B. (2021). Dark Tetrad personality traits and the propensity to lie across multiple contexts. *Personality and individual differences, 177*, 110792. doi:10.1016/j.paid.2021.110792

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of Personality: A 10 Year Review: Dark Triad of Personality. *Social and personality psychology compass, 7*(3), 199-216. doi:10.1111/spc3.12018

Gaur, A. S., & Gaur, S. S. (2009). *Statistical methods for practice and research : a guide to data analysis using SPSS*(2nd ed. ed.).

Glenn, A. L., & Sellbom, M. (2015). Theoretical and empirical concerns regarding the dark triad as a construct. *J Pers Disord, 29*(3), 360-377. doi:10.1521/pedi_2014_28_162

Goldberg, L. R. (1993). The Structure of Phenotypic Personality Traits. *The American psychologist, 48*(1), 26-34. doi:10.1037/0003-066X.48.1.26

Greer, T., Dunlap, W. P., & Beatty, G. O. (2003). A Monte Carlo Evaluation of the Tetrachoric Correlation Coefficient. *Educational and psychological measurement, 63*(6), 931-950. doi:10.1177/0013164403251318

Gustavson, K., Røysamb, E., & Borren, I. (2019). Preventing bias from selective non-response in population-based survey studies: Findings from a Monte Carlo simulation study. *BMC Med Res Methodol, 19*(1), 120-120. doi:10.1186/s12874-019-0757-1

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and individual differences, 1*(2), 111-119. doi:10.1016/0191-8869(80)90028-8

Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. *Journal of consulting and clinical psychology, 53*(1), 7-16. doi:10.1037//0022-006X.53.1.7

Hartung, J., Bader, M., Moshagen, M., & Wilhelm, O. (2021). Age and gender differences in socially aversive ("dark") personality traits. *European journal of personality*, 89020702098843. doi:10.1177/0890207020988435

Haslam, N. (2016). Concept Creep: Psychology's Expanding Concepts of Harm and Pathology. *Psychological inquiry, 27*(1), 1-17. doi:10.1080/1047840X.2016.1082418

Hayton, J. C., Allen, D. G., & Scarpello, V. (2011). Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. *SAGE Quantitative Research Methods, 7*(2), 191-205. doi:10.1177/1094428104263675

Hodson, G., Book, A., Visser, B. A., Volk, A. A., Ashton, M. C., & Lee, K. (2018). Is the Dark Triad common factor distinct from low Honesty-Humility? *Journal of research in personality, 73*, 123-129. doi:10.1016/j.jrp.2017.11.012

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185. doi:10.1007/BF02289447

Kowalski, R. M. (2001). *Behaving badly: Aversive behaviors in interpersonal relationships*. Washington,US,DC: Washington: American Psychological Association.

Lee, K., Ashton, M. C., Pozzebon, J. A., Visser, B. A., Bourdage, J. S., & Ogunfowora, B. (2009). Similarity and Assumed Similarity in Personality Reports of Well-Acquainted Persons. *J Pers Soc Psychol, 96*(2), 460-472. doi:10.1037/a0014059

Lehmann, L., Bargum, K., & Reuter, M. (2006). An evolutionary analysis of the relationship between spite and altruism. *J Evol Biol, 19*(5), 1507-1516. doi:10.1111/j.1420-9101.2006.01128.x

Lim, S., & Jahng, S. (2019). Determining the Number of Factors Using Parallel Analysis and Its Recent Variants. *Psychol Methods, 24*(4), 452-467. doi:10.1037/met0000230

Marcus, D. K., & Zeigler-Hill, V. (2015). A Big Tent of Dark Personality Traits. *Social and Personality Psychology Compass, 9*(8), 434-446. doi:10.1111/spc3.12185

Markon, K. E. (2019). Bifactor and Hierarchical Models: Specification, Inference, and Interpretation. *Annu Rev Clin Psychol, 15*(1), 51-69. doi:10.1146/annurev-clinpsy-050718-095522

McCrae, R. R. (2002). The maturation of personality psychology: Adult personality development and psychological well-being. *Journal of research in personality, 36*(4), 307-317. doi:10.1016/S0092-6566(02)00011-9

McCrae, R. R., & Costa, P. T. (2003). *Personality in adulthood : a five-factor theory perspective* (2nd ed. ed.). New York: Guilford Press.

McCrae, R. R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality, 60*(2), 175-215. doi:https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McCrae, R. R., & Terracciano, A. (2005). Personality Profiles of Cultures: Aggregate Personality Traits. *J Pers Soc Psychol, 89*(3), 407-425. doi:10.1037/0022-3514.89.3.407

McLarnon, M. J. W., & Tarraf, R. C. (2021). Getting to the core: How "(dis)honest" is the core of the Dark Triad? *Personality and individual differences, 171*. doi:10.1016/j.paid.2020.110545

Miller, J. D., Dir, A., Gentile, B., Wilson, L., Pryor, L. R., & Campbell, W. K. (2010). Searching for a Vulnerable Dark Triad: Comparing Factor 2 Psychopathy, Vulnerable Narcissism, and Borderline Personality Disorder. *J Pers, 78*(5), 1529-1564. doi:10.1111/j.1467-6494.2010.00660.x

Miller, J. D., Hoffman, B. J., Gaughan, E. T., Gentile, B., Maples, J., & Keith Campbell, W. (2011). Grandiose and Vulnerable Narcissism: A Nomological Network Analysis. *J Pers, 79*(5), 1013-1042. doi:10.1111/j.1467-6494.2010.00711.x

Moshagen, M., Hilbig, B. E., & Zettler, I. (2018). The Dark Core of Personality. *Psychol Rev, 125*(5), 656-688. doi:10.1037/rev0000111

Moshagen, M., Zettler, I., & Hilbig, B. E. (2020). Measuring the Dark Core of Personality. *Psychol Assess, 32*(2), 182-196. doi:10.1037/pas0000778

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The Malevolent Side of Human Nature: A Meta-Analysis and Critical Review of the Literature on the Dark Triad (Narcissism, Machiavellianism, and Psychopathy). *Perspect Psychol Sci, 12*(2), 183-204. doi:10.1177/1745691616666070

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *Eur. J. Soc. Psychol, 15*(3), 263-280. doi:10.1002/ejsp.2420150303

Paulhus, D. L. (2014). Toward a Taxonomy of Dark Personalities. *Curr Dir Psychol Sci, 23*(6), 421-426. doi:10.1177/0963721414547737

Paulhus, D. L., & Klaiber, P. (2020). HEXACO, Dark Personalities, and Brunswik Symmetry. *European journal of personality, 34*(4), 541-542.

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. doi:10.1016/S0092-6566(02)00505-6

Raskin, R., & Terry, H. (1988). A principal-components analysis of the narcissistic personality inventory and further evidence of its construct validity. *J Pers Soc Psychol, 54*(5), 890-902. doi:10.1037//0022-3514.54.5.890

Reise, S. P., Bonifay, W., & Haviland, M. G. (2018). Bifactor Modelling and the Evaluation of Scale Scores. In (Vol. 2-2, pp. 675-707). Chichester, UK: Chichester, UK: John Wiley & Sons, Ltd.

Revelle, W. (2021). psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 2.1.9. Retrieved from https://CRAN.Rproject.org/package=psych

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When Can Categorical Variables Be Treated as Continuous? A Comparison of Robust Continuous and Categorical SEM Estimation Methods Under Suboptimal Conditions. *Psychol Methods, 17*(3), 354-373. doi:10.1037/a0029315

RStudio Team. (2020). Boston, MA: PBC. Retrieved from http://www.rstudio.com/

Schmitt, N. (1996). Uses and Abuses of Coefficient Alpha. *Psychological assessment, 8*(4), 350-353. doi:10.1037/1040-3590.8.4.350

Schreiber, A., & Marcus, B. (2020). The Place of the "Dark Triad" in General Models of Personality: Some Meta-Analytic Clarification. *Psychol Bull, 146*(11), 1021-1041. doi:10.1037/bul0000299

Steiner, M., Grieder, S. (2020). EFAtools: An R package with fast and flexible implementations of exploratory factor analysis tools. *Journal of Open Source Software, 5(53)*(2521). doi:https://doi.org/10.21105/joss.02521

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *Int J Med Educ, 2*, 53-55. doi:10.5116/ijme.4dfb.8dfd

Van de Vijver, F., & Hambleton, R. K. (1996). Translating Tests: Some Practical Guidelines. *European psychologist, 1*(2), 89-99. doi:10.1027/1016-9040.1.2.89

Widiger, T. A. (2017). *The Oxford handbook of the five factor model*. Oxford University press, Incorporated

Zwick, W. R., & Velicer, W. F. (1986). Comparison of Five Rules for Determining the Number of Components to Retain. *Psychological bulletin, 99*(3), 432-442. doi:10.1037/0033-2909.99.3.432

Appendices

Appendix A: The Dark Core Inventory

Item 1: It is hard for me to see someone suffering.

Item 2: All in all, it is better to be humble and honest than important and dishonest.

Item 3: A smart person always knows when, how, and to whom to say something in order to seal the fate of someone who offended him.

Item 4: Most people are basically good and kind.

Item 5: I do not mind sharing the stage.

Item 6: Payback needs to be quick and nasty.

Item 7: If I were on the Titanic, I would not deserve to be on the first lifeboat any more than anyone else.

Item 8: When I get annoyed, tormenting people makes me feel better.

Item 9: It is not okay to spread rumors, not even to defend those you care about.

Item 10: If my neighbor complained that I was playing my music too loud, I would turn it down, even if this bothers me.

Item 11: My own pleasure is all that matters.

Item 12: If a short-cut to success is illegal, it is not smart to take it.

Item 13: No matter how much I have of something, I always want more.

Item 14: I believe that lying is necessary to maintain a competitive advantage over others.

Item 15: In principle, everyone is worth the same.

Item 16: I'll say anything to get what I want.

Item 17: Hurting people would make me very uncomfortable.

Item 18: If I had the opportunity, then I would gladly pay a small sum of money to see a classmate who I do not like fail his or her final exam.

Item 19: Doing good deeds serves no purpose; it only makes people poor and lazy.

Item 20: Never tell anyone the real reason you did something unless it is useful to do so.

Item 21: A person should use any and all means that are to his advantage, taking care of course, that others do not find out.

Item 22: Actually, I'm kind of greedy.

Item 23: I do not care much for having control over other people.

Item 24: People who mess with me always regret it.

Item 25: Things cannot always go my way.

Item 26: I think about harassing others for enjoyment.

Item 27: Taking credit for someone else's ideas is a no-go.

Item 28: I'm not very sympathetic to other people or their problems.

Item 29: There have been times when I was willing to suffer some small harm so that I could punish someone else who deserved it.

Item 30: Why should I care about other people, when no one cares about me?

Item 31: I would not cheat, even if there was only a small chance of getting caught.

Item 32: Most people are somehow losers.

Item 33: I do not particularly enjoy manipulating other people's feelings.

Item 34: I cannot imagine how being mean to others could ever be exciting.

Item 35: I am willing to volunteer for people in need.

Item 36: To make money there are no right and wrong ways anymore. Only easy and hard ways.

Item 37: There is poor comfort in revenge.

Item 38: Most people deserve respect.

Item 39: I can barely stand it if another person is at the center of events.

Item 40: I make a point of trying not to hurt others in pursuit of my goals.

Item 41: I do not deserve more things in life than others.

Item 42: I hate to see people hurt.

Item 43: If a business makes a billing mistake in your favor, it's okay not to tell them about it because it was their fault.

Item 44: I try to look out for myself first, even if it means making things difficult for other people.

Item 45: If I opposed the election of an official, then I would be glad to see him or her fail even if their failure hurt my community.

Item 46: Doing good deeds brings joy to the heart.

Item 47: A person should obey the law no matter how much it interferes with their personal ambition.

Item 48: I would like to make some people suffer, even if it meant that I would go to hell with them.

Item 49: It doesn't matter how much I have, I'm never completely satisfied.

Item 50: It's wise to keep track of information that you can use against people later.

Item 51: It does not give me much pleasure to see my rivals fail.

Item 52: I don't want people to be afraid of me or my impulses.

Item 53: Someone who hurts me cannot count on my sympathy.

Item 54: I avoid humiliating others.

Item 55: People who get mistreated have usually done something to bring it on themselves.

Item 56: I would not pursue what I want if this causes problems for other people.

Item 57: I would be willing to take a punch if it meant that someone I did not like would receive two punches.

Item 58: Stories about honesty and goodness serve only to make people confused and stupid.

Item 59: It is hard to get ahead without cutting corners here and there.

Item 60: Sweeter is the revenge that takes a long time to prepare.

Item 61: For most things, there is a point of having enough.

Item 62: Make sure your plans benefit you, not others.

Item 63: I do not seek power.

Item 64: Success is based on survival of the fittest; I am not concerned about the losers.

Item 65: I tend to forgive the wrongs I have suffered.

Item 66: If I ever tormented others, I felt strong remorse.

Item 67: It’s okay to treat badly somebody who behaves like scum.

Item 68: I feel sorry if things I do upset people.

Item 69: It is sometimes worth a little suffering on my part to help others in need.

Item 70: Making people feel bad about themselves does not make me feel any better.

Appendix B: Dark Core Inventory: Norsk versjon

Ledd 1: Det er vanskelig for meg å se noen som lider.

Ledd 2: Når alt kommer til alt er det bedre å være ydmyk og ærlig enn viktig og uærlig.

Ledd 3: En smart person vet alltid når, hvordan og til hvem han skal si noe for å avgjøre skjebnen til noen som har fornærmet han.

Ledd 4: De fleste mennesker er i bunn og grunn snille og greie.

Ledd 5: Det gjør meg ikke noe å dele scenen.

Ledd 6: Gjengjeldelse bør være rask og stygg.

Ledd 7: Hvis jeg var på Titanic, ville jeg ikke ha fortjent å komme på den første livbåten noe mer enn alle andre.

Ledd 8: Når jeg blir irritert føler jeg meg bedre av å pine andre.

Ledd 9: Det er ikke greit å spre rykter, selv ikke for å forsvare de du bryr deg om.

Ledd 10: Hvis naboen klaget over at jeg spilte for høy musikk ville jeg dempe den, selv om det plager meg.

Ledd 11: At jeg selv har det bra er alt som betyr noe.

Ledd 12: Hvis en snarvei til suksess er ulovlig er det ikke lurt å ta den.

Ledd 13: Uansett hvor mye jeg har av noe vil jeg alltid ha mer.

Ledd 14: Jeg syns det er nødvendig å lyve for å beholde et forsprang over andre.

Ledd 15: I prinsippet er alle like mye verdt.

Ledd 16: Jeg vil si hva som helst for å få det jeg vil ha.

Ledd 17: Å såre andre ville gjort meg veldig ukomfortabel.

Ledd 18: Hvis jeg hadde hatt muligheten, skulle jeg gjerne betalt et mindre beløp for å se en jeg ikke liker i klassen stryke til eksamen.

Ledd 19: Gode gjerninger fører ikke til noe, de gjør bare folk fattige og late.

Ledd 20: Fortell aldri noen den egentlige grunnen til at du gjorde noe med mindre det lønner seg å gjøre det.

Ledd 21: En person burde bruke alle mulige midler til sin fordel, så lenge man sørger for at andre ikke finner ut av det, selvsagt.

Ledd 22: Jeg er faktisk ganske grådig.

Ledd 23: Jeg bryr meg ikke så mye om å ha kontroll over andre.

Ledd 24: Folk som kødder med meg angrer alltid.

Ledd 25: Jeg kan ikke alltid få det som jeg vil.

Ledd 26: Jeg tenker på å plage andre kun for fornøyelsens skyld.

Ledd 27: Å ta æren for andres idéer er uaktuelt.

Ledd 28: Jeg har ikke mye sympati med andre folk eller deres problemer.

Ledd 29: Det har hendt at jeg har vært villig til å lide litt skade selv, for å kunne straffe noen som fortjente det.

Ledd 30: Hvorfor skal jeg bry meg om andre når ingen bryr seg om meg?

Ledd 31: Jeg ville ikke jukset, selv om sjansen for å bli tatt var lav.

Ledd 32: Folk flest er tapere på en eller annen måte.

Ledd 33: Jeg er ikke spesielt glad i å manipulere andres følelser.

Ledd 34: Jeg kan ikke forestille meg at det kan være spennende å være ondskapsfull mot andre.

Ledd 35: Jeg er villig til å stille opp frivillig for folk i nød.

Ledd 36: Det finnes ikke lenger riktige og gale måter å tjene penger på, bare lette og vanskelige måter.

Ledd 37: Det er liten trøst i hevn.

Ledd 38: Folk flest fortjener respekt.

Ledd 39: Når noen andre enn meg selv er midtpunktet holder jeg det nesten ikke ut.

Ledd 40: Jeg gjør så godt jeg kan for ikke å skade noen når jeg prøver å oppnå mine mål.

Ledd 41: Jeg fortjener ikke flere ting i livet enn andre.

Ledd 42: Jeg hater å se andre ha det vondt.

Ledd 43: Hvis et firma gjør en feil med betalingen i din favør er det greit å ikke fortelle dem om det fordi det var deres feil.

Ledd 44: Jeg prøver å passe på meg selv først, selv om det gjør ting vanskeligere for andre.

Ledd 45: Hvis jeg var imot valget av en representant ville jeg være glad for å se han eller hun feile selv om det ville skade lokalsamfunnet.

Ledd 46: Å gjøre gode gjerninger gleder mitt hjerte.

Ledd 47: Man bør følge loven uansett hvor mye den står i veien for ens egne ambisjoner.

Ledd 48: Jeg ville nyte å få enkelte mennesker til å lide, selv om det betyr at jeg havner i helvete sammen med dem.

Ledd 49: Det spiller ingen rolle hvor mye jeg har, jeg blir aldri helt fornøyd.

Ledd 50: Det er lurt å holde rede på informasjon som du senere kan bruke mot andre.

Ledd 51: Det gir meg ikke mye glede å se rivalene mine mislykkes.

Ledd 52: Jeg ønsker ikke at andre mennesker skal være redde for meg eller mine impulser.

Ledd 53: Den som sårer meg kan ikke regne med min sympati.

Ledd 54: Jeg unngår å ydmyke andre.

Ledd 55: Folk som blir mishandlet har vanligvis fortjent det.

Ledd 56: Jeg ville ikke prøvd å få det jeg vil ha dersom det hadde skapt problemer for andre.

Ledd 57: Jeg ville være villig til å ta imot et slag hvis det betyr at noen jeg ikke liker ville fått to slag.

Ledd 58: Historier om ærlighet og godhet gjør bare folk forvirrede og dumme.

Ledd 59: Det vanskelig å komme foran uten å ta snarveier her og der.

Ledd 60: Den søteste hevnen er den som tar lang tid å planlegge.

Ledd 61: For de fleste ting kommer det til et punkt der man har nok.

Ledd 62: Sørg for at planene dine er bra for deg, ikke for andre.

Ledd 63: Jeg søker ikke makt.

Ledd 64: Suksess er basert på den sterkestes overlevelse; jeg er ikke bekymret for taperne.

Ledd 65: Jeg pleier å tilgi for den urett jeg har lidd.

Ledd 66: Hvis jeg har plaget andre har jeg følt sterk anger.

Ledd 67: Det er greit å behandle de som oppfører seg som drittsekker dårlig.

Ledd 68: Jeg blir lei meg dersom folk tar seg nær av ting jeg gjør.

Ledd 69: Det er noen ganger verdt det å utsette seg selv for litt lidelse for å hjelpe andre i nød.

Ledd 70: Å få andre til å føle seg mindreverdige gjør ikke at jeg føler meg bedre.

Appendix C: Distribution of responses for the highest and lowest mean scores

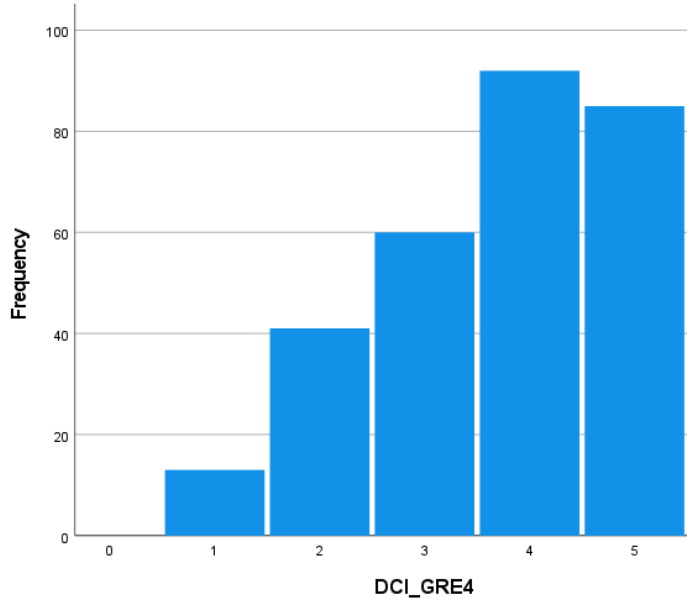

Figure 4. The highest DCI mean score belonged to the incorrectly translated variable GRE4 «*For de fleste ting kommer det til et punkt der det er nok*» (corresponding to the statement «For most things, there is a point of having had enough».)

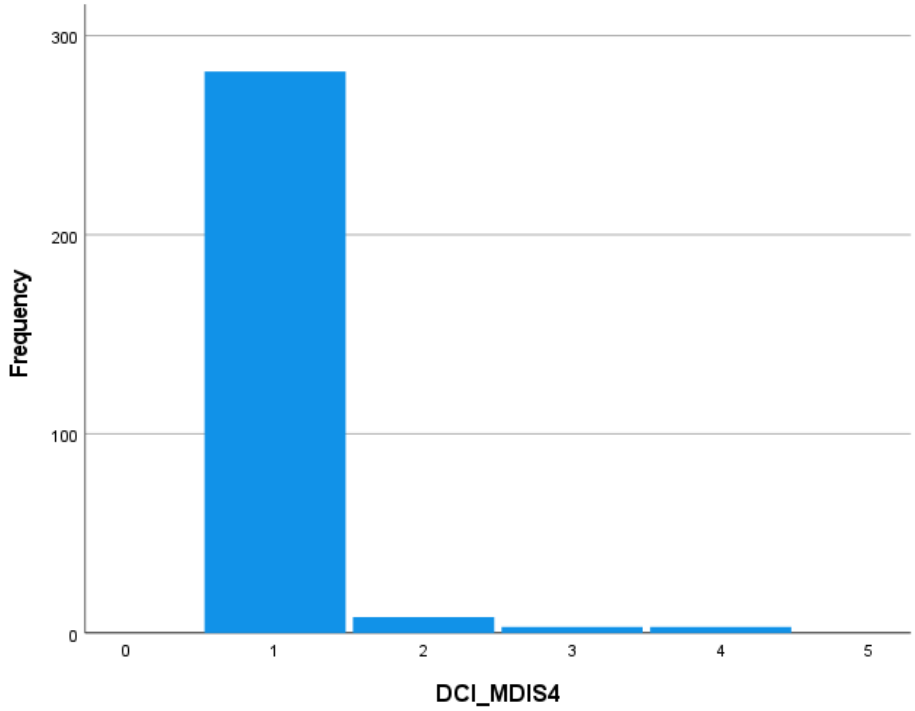

Figure 5. The lowest DCI mean score belonged to a variable corresponding to the statement: «People who get mistreated have usually done something to bring it on themselves» (*Folk som blir mishandlet har vanligvis fortjent det).* In part, this might be attributable to the translation of the word “mistreated”.

Appendix D: Item overview for factors 2 through 8

Factor 2: Vindictiveness
FRUST5: I would like to make some people suffer, even if it meant that I would go to hell with them.
SPITE2: If I had the opportunity, then I would gladly pay a small sum of money to see a classmate who I do not like fail his or her final exam.
SPITE5: I would be willing to take a punch if it meant that someone I did not like would receive two punches.

Factor 3: Deceitfulness ("greed driven" deceit)
EGO5: It is hard to get ahead without cutting corners here and there.
FRUST3: A person should use any and all means that are to his advantage, taking care of course, that others do not find out
GRE1: No matter how much I have of something, I always want more.
GRE2: Actually, I'm kind of greedy.
GRE3: It doesn't matter how much I have, I'm never completely satisfied.
MACH2: I believe that lying is necessary to maintain a competitive advantage over others.
MACH7: Make sure your plans benefit you, not others.
PATHY2: I'll say anything to get what I want.
SCTR2: I try to look out for myself first, even if it means making things difficult for other people.

Factor 4: Sadism
CRUD6: Doing good deeds brings joy to the heart.
SAD2: Hurting people would make me very uncomfortable.
SAD5: I hate to see people hurt.
SAD7: If I ever tormented others, I felt strong remorse.
SCTR1: I'm not very sympathetic to other people or their problems.
SCTR4: I feel sorry if things I do upset people.

Factor 5: Integrity
EGO4: A person should obey the law no matter how much it interferes with their personal ambition.
FRUST2: If a short-cut to success is illegal, it is not smart to take it.
MDIS1: It is not okay to spread rumors, not even to defend those you care about.
MDIS2 Taking credit for someone else's ideas is a no-go.

Factor 6: Callousness
PATHY4: I do not particularly enjoy manipulating other people's feelings.
PATHY5: I make a point of trying not to hurt others in pursuit of my goals.

Factor 7: Narcissistic Entitlement
NARC1: I do not mind sharing the stage.
NARC2: In principle, everyone is worth the same.

Factor 8: Callousness
FRUST1: A smart person always knows when, how, and to whom to say something in order to seal the fate of someone who offended him.
EGO3: To make money there are no right and wrong ways anymore. Only easy and hard ways.
PATHY7: Success is based on survival of the fittest; I am not concerned about the losers.