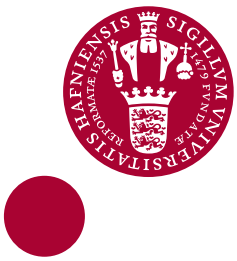

The dark core of personality

Moshagen, Morten; Hilbig, Benjamin E.; Zettler, Ingo

Published in:
Psychological Review

DOI:
10.1037/rev0000111

Publication date:
2018

Document version
Peer reviewed version

This work is scheduled to appear in
Psychological Review

Reference:
Moshagen, M., Hilbig, B. E., & Zettler, I. (2018). The dark core of personality. *Psychological Review, 125*, 656–688. https://doi.org/10.1037/rev0000111

The dark core of personality

Morten Moshagen

Ulm University, Germany

Benjamin E. Hilbig

University of Koblenz-Landau, Germany

and Max-Planck Institute for Research on Collective Goods, Germany

Ingo Zettler

University of Copenhagen, Denmark

Author Note

Correspondence concerning this manuscript should be directed to Morten Moshagen, Research Methods, Institute of Psychology and Education, Ulm University, Albert-Einstein-Allee 47, 89081 Ulm, Germany. E-Mail: morten.moshagen@uni-ulm.de.

All authors contributed equally to this work. The order of authorship was determined by a nonlinear combination of subjectively weighted criteria with unknown ecological validity, mingled with and ultimately outweighed by a nontransparent randomization, which, incidentally, was performed by the first author.

Neither the general idea underlying this manuscript nor the data used in this manuscript have been published, presented at a conference, posted on a listserver, or shared on a website previously, neither in whole nor in part.

Preparation of this manuscript was supported by Grant MO 2830/1-1 to the first, Grants HI 1600/1-2 and HI 1600/6-1 to the second, and Grant ZE 948/1-1 to the final author from the German Research Foundation (DFG).

The dark core of personality

Abstract

Many negatively connoted personality traits (often termed "dark traits") have been introduced to account for ethically, morally, and socially questionable behavior. Herein, we provide a unifying, comprehensive theoretical framework for understanding dark personality in terms of a general dispositional tendency of which dark traits arise as specific manifestations. That is, we theoretically specify the common core of dark traits, which we call the *Dark Factor of Personality* (*D*). The fluid concept of D captures individual differences in the tendency to maximize one's individual utility—disregarding, accepting, or malevolently provoking disutility for others—, accompanied by beliefs that serve as justifications. To critically test D, we unify and extend prior work methodologically and empirically by considering a large number of dark traits simultaneously, using statistical approaches tailored to capture both the common core and the unique content of dark traits, and testing the predictive validity of both D and the unique content of dark traits with respect to diverse criteria including fully consequential and incentive-compatible behavior. In a series of four studies (N > 2,500), we provide evidence in support of the theoretical conceptualization of D, show that dark traits can be understood as specific manifestations of D, demonstrate that D predicts a multitude of criteria in the realm of ethically, morally, and socially questionable behavior, and illustrate that D does not depend on any particular indicator variable included.

Keywords: D factor; dark traits; dark triad; Big Five; HEXACO

The dark core of personality

Ethically, morally, and socially questionable behavior is part of everyday life and instances of ruthless, selfish, unscrupulous, or even downright evil behavior can easily be found across history and cultures. Indeed, both our cultural inheritance (e.g., Gāo Qiú, Vlad the Impaler) and examples of peoples' daily behavior (e.g., corruption, denunciation, white lies in relationships) regularly involve some acts of misconduct. Correspondingly, researchers from various disciplines within and beyond psychology have aimed to describe, explain, and predict such negatively connoted activities that often endanger positive social interactions and thus the very fabric of functioning societies at large. From the viewpoint of psychology, and personality psychology in particular, research has long centered around stable traits that are linked to ethically, morally, and socially questionable behavior—consider Adorno's work on Right-Wing Authoritarianism (e.g., Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950) or Eysenck's work on Psychoticism (e.g., Eysenck & Eysenck, 1976) to name early seminal examples.

Today, many researchers use the umbrella term "dark traits" to subsume subclinical personality traits that are linked to ethically, morally, and socially questionable behavior. Research on dark traits has seen an upsurge of interest recently, as reflected, for instance, in several special issues devoted to dark traits, including outlets typical for personality research (e.g., Miller & Lynam, 2015; Veselka & Vernon, 2014) as well as those typical for applied psychology (Schyns, 2015). Moreover, dark traits have become increasingly prominent concepts in related fields beyond psychology, such as criminology (e.g., Flexon, Meldrum, Young, & Lehmann, 2016) or behavioral economics (e.g., Zhang & Ortmann, 2016). Among other things, high levels in dark traits have been associated with more selfish or unfair decisions in economic and social dilemma paradigms (Seuntjens, Zeelenberg, van de Ven, & Breugelmans, 2015), with counterproductive work behavior (O'Boyle, Forsyth, Banks, & McDaniel, 2012), with instrumental and reactive violence (Blais, Solodukhin, & Forth, 2014),

and with sadistic behavior (Buckels, Jones, & Paulhus, 2013), illustrating the relevance of dark traits for a diverse set of behaviors.

Whereas an increasing number of allegedly different and increasingly narrow dark traits have been introduced (recent examples are, for instance, Sadism, O'Meara, Davies, & Hammond, 2011; or Spitefulness, Marcus, Zeigler-Hill, Mercer, & Norris, 2014), different dark traits are, by definition, necessarily related. Indeed, for any construct to be considered a dark trait it will necessarily reflect some undesirable or problematic tendencies. Correspondingly, both theorizing (e.g., Marcus & Zeigler-Hill, 2015; Paulhus, 2014) and empirical findings revealing considerable overlap across dark traits (e.g., Book et al., 2016; Jones & Figueredo, 2013; O'Boyle et al., 2012) are aligned with the notion that the very basis—and thus common core—of dark traits is a general tendency towards ethically, morally, and/or socially questionable behavior. Herein, we present a theory of this broad, general tendency and thus the content, magnitude, and relevance of the common basis of dark traits, that is, their common core.

Specifically, we present a theoretical conceptualization of a general dispositional tendency underlying different dark traits, which we consider to represent a fluid construct and which we call the *Dark Factor of Personality* or simply *D*. Across four studies we then test our conceptualization of D by investigating the commonalities of several dark traits, by exploring the degree to which these dark traits are subsumed in D and comprise psychologically meaningful characteristics beyond D, and by linking D, the dark traits, and the remaining characteristics of each of the dark traits beyond D to criteria in the realm of ethically, morally, and socially questionable behavior, including "actual behavior" (Baumeister, Vohs, & Funder, 2007; King, 2010). Finally, we test our conceptualization of D as a fluid construct using a simulation approach. Taken together, we thus provide—and critically test—a theoretical framework for understanding dark personality as a whole.

The Dark Factor of Personality

The theoretical idea purported herein is straightforward. We propose that dark traits are specific manifestations of a general, basic dispositional behavioral tendency, which we call the Dark Factor of Personality. Specifically, we conceptualize D 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*. Correspondingly, D encompasses the following main aspects.

First, individuals with high levels in D will generally aim to maximize their individual utility at the cost of others. Here, utility is used in the broad sense as a "measure of extent of goal achievement" (Baron, 2008, p. 233), so that one's individual utility can take the form of visible gains such as a higher status or higher monetary payoffs, but also less tangible ones such as feelings of power, superiority, pleasure, or joy. Crucial for this aspect is that utility maximization is sought despite running contrary to the interest of others or even for the sake of such negative externalities. As such, D refers specifically to those instances of utility maximization that are potentially harmful to others. By contrast, utility-maximizing behaviors that are unlikely to impose costs on others and that would generally be considered to reflect normal psychological functioning—such as doing sports to improve one's health, engaging in consensual sex, or recreational activities—are beyond the scope of D. Note also that utility maximization does not necessarily have to take the form of actively performing some kind of behavior that (directly or indirectly) incurs costs for others, but may come in the form of refraining from behavior that may benefit others at some personal cost (e.g., not warning someone). However, the idea of utility maximization does not necessarily imply that individuals high in D will never behave cooperatively. For example, one may think of a situation in which an individual high in D strategically chooses to cooperate (thus foregoing some immediate gain) rather than defect in a social dilemma situation (Kollock, 1998), e.g., to build a certain reputation (e.g., Cooper, DeJong, Forsythe, & Ross, 1996; Moshagen, Hilbig, & Musch, 2011) or to avoid sanctions (e.g., Fehr & Gächter, 2000; Hilbig, Zettler, &

Heydasch, 2012). Hence, the core idea here is that individuals high in D seek to maximize their own utility whenever it subjectively outweighs their own disutility.

Further, in seeking to maximize their own utility, individuals with high levels in D will generally disregard, accept, or malevolently provoke disutility for others.[1] Corresponding to the use of the term utility, disutility is also understood in a broad sense, representing anything that does not foster or that directly prevents others' goal achievement. One may think of physical (e.g., hurting others), psychological (e.g., cheating on someone), or financial (e.g., exploiting others in negotiations) disutilities for others, including societies at large (e.g., tax fraud). Vice versa, individuals high in D will generally not be motivated to increase the utility of others (e.g., helping others in need) without benefiting themselves and, more generally, will not derive utility for themselves from other's utility per se (e.g., being happy for others). Thus, behaviors that might increase others' utility at some personal costs—that is, altruistic behaviors in a broad sense—are not aligned with D.

Of note, the extent to which individuals high in D are concerned about others' disutility can vary as hinted by the multiple variants "disregarding, accepting, or malevolently provoking disutility". Whereas some high in D may maximize their utility hardly even noticing the negative consequences for other people, others may be aware of—but not held back by—the disutility inflicted on other people, and still others might actually derive immediate utility for themselves (e.g., pleasure) from disutility inflicted on other people. The latter aspect can also involve behaviors in the realm of retaliation and revenge which often not

[1] Note that the definition of D does not apply to contexts characterized by a mutual consensus concerning such actions. For example, ritualized fistfights are based on the common agreement of the combatants regarding the appropriateness of seeking own utility by afflicting disutility on others. The conceptualization of D thus refers to contexts where disutility is afflicted on others in a non-consensual way.

only serve to restore equality but also to regulate emotions such as anger (Bosman & van Winden, 2002; Karagonlar & Kuhlman, 2013; Roberts, Vakirtzis, Kristjánsdóttir, & Havlíček, 2013).

In sum, the utility-based aspects of D specify that individuals high in D, in seeking own utility, pursue behaviors that may negatively affect others and avoid behaviors that unilaterally benefit others, especially if this might incur personal costs. In terms of the economic concept of social preferences or, more specifically, social value orientations (e.g., Murphy & Ackermann, 2014), individuals high in D thus represent those who attach negative or zero weight to others' outcomes but positive non-zero weight to their own (i.e., individualists, competitors, and sadists in the social value orientation terminology). Whereas the utility-based aspects of D thus partially mirror the typology of social value orientations, D encompasses aspects well beyond the narrow and operationally defined concept of social preferences, especially since D subsumes attitudes and beliefs that serve as the basis of and justification for individual utility maximization at others' costs.

Specifically, individuals with high levels of D will hold beliefs that serve as justifications for individual utility maximization at the expense of others. Again, this aspect is understood in broad terms, allowing for a variety of specific beliefs that ultimately share the function of serving as justifications. For example, individuals high in D may tend to consider themselves as superior, to consider others as inferior, to endorse ideologies favoring dominance, or to believe that people generally think about themselves first and that, therefore, there is no injunctive norm to refrain from utility maximization (Ajzen, 1991). These—and other—beliefs offer some kind of justification to strive for own benefits at the cost of others. Note that individuals with high levels in D may not need to construct such justifications actively or consciously. Instead, utility maximization can also be driven or shaped by more implicit processes which serve to justify ethically, morally, and socially questionable

behavior, e.g., by fostering that individuals maintain a positive self-image despite their questionable behavior (e.g., Barkan, Ayal, & Ariely, 2015).

In line with this conceptualization of D as a general tendency to maximize one's utility at the expense of others while adopting justifying beliefs, we propose that any dark trait can be understood as a specific manifestation of D. More precisely, we propose that any dark trait boils down to at least one of the defining features of D, which thus reflects the general basis of different dark traits. Importantly, though, a particular dark trait may entail additional characteristics beyond those comprised in D. For example, the defining features of Narcissism include a sense of entitlement and superiority as well as dominant and aggressive behaviors towards others, but also involve a need for social admiration (Back et al., 2013). Whereas the former characteristics are well aligned with D (reflecting justifying beliefs and utility maximization at the expense of others, respectively), the latter may be a unique feature of Narcissism beyond D. Moreover, various dark traits differ with regard to which of the above aspects is paramount. For instance, for individuals high in Narcissism a justifying belief that one is superior might be more pronounced than the aspect of disregarding, accepting, or provoking disutilities for others, whereas the opposite might be the case for individuals high in Spitefulness. In a similar vein, maximizing one’s utility might be particularly manifested in high levels in Egoism (accepting other’s disutility as a by-product), whereas high levels in Sadism place a strong emphasis on deriving utility from actively provoking disutilities for others.

These differences between various dark traits notwithstanding, we propose that any dark trait relates to at least one (and typically several) of the defining aspects of D. Thus, D represents the underlying general tendency from which dark traits arise as specific manifestations, implying that D is responsible for the commonalities between various dark traits and thereby represents their common core. In more methodological terms, this conceptualization requires that the overlap of indicators reflecting different dark traits can be

represented by one common factor that will mirror the defining aspects of D and that, in turn, this factor is the prime source of variance responsible for commonalities between dark traits.

Importantly, given that dark traits differ with regard to the emphasis placed on the three characteristics of D and that a particular dark trait may comprise aspects beyond D, we conceptualize D as a fluid construct, much like the *g* factor of intelligence in the ability domain (e.g., Jensen, 1998; McGrew, 2009). That is, D is manifested in a very large number of ethically, morally, and socially questionable attitudes and behaviors, and can thus be measured in many different ways. Although a multitude of purportedly different dark traits exists and further, even more specific ones may be defined, the central implication of our proposal is that D constitutes the major source of variance underlying individual differences on dark traits. As a corollary, any particular measurement instrument designed to assess a dark trait will reflect D as well, although to a different degree depending on the nature of the particular trait (and instrument designed to assess this trait). Correspondingly, D is not defined by any one particular set or combination of dark traits. Rather, all *combinations* of dark traits have a common core that, in turn, generally mirrors our conceptualization of D.

Preliminary Empirical Support for a Common Core of Dark Traits

One line of evidence supporting the idea that various dark traits arise from a general underlying disposition and thus share a common core stems from studies that concurrently investigated various dark traits. In practically all these investigations, dark traits are positively—although imperfectly—associated with each other. In a meta-analysis by O'Boyle et al. (2012) on the currently most prominent dark traits Machiavellianism, Narcissism, and Psychopathy (the "Dark Triad", Paulhus & Williams, 2002), the correlation between Machiavellianism and Narcissism was r = .23 (n = 8,423), the correlation between Machiavellianism and Psychopathy was r = .46 (n = 5,762), and the correlation between Narcissism and Psychopathy was r = .42 (n = 8,538). Very similar results were observed in a more recent meta-analysis by Muris, Merckelbach, Otgaar, and Ewout (2017), with

correlations of $r = .34$ between Machiavellianism and Narcissism, $r = .58$ between Machiavellianism and Psychopathy, and $r = .34$ between Narcissism and Psychopathy ($N = 42{,}359$). Examples involving dark traits other than the components of the Dark Triad are more rare, but nonetheless reveal similar degrees of overlap: Campbell, Bonacci, Shelton, Exline, and Bushman (2004) reported a correlation of $r = .50$ ($N = 918$) between Psychological Entitlement and Narcissism, Gerbasi and Prentice (2013) reported a correlation of $r = .46$ ($N = 80$) between Self-Interest and Narcissism, Moore, Detert, Trevino, Baker, and Mayer (2012) reported correlations of $.44 \leq r \leq .56$ ($N = 194$ and $N = 272$, respectively) between Moral Disengagement and Machiavellianism, and Marcus et al. (2014) reported correlations of $.06 \leq r \leq .71$ between Spitefulness and all Dark Triad traits ($297 \leq N \leq 946$).

As these examples illustrate, the correlations between different dark traits typically range between .20 and .60 and are thus substantial in magnitude, though not perfect. This substantial proportion of shared variance among various dark traits, in turn, can be taken as evidence concerning commonalities between various dark traits and consequently makes the existence of a common underlying dispositional tendency plausible.

Theoretical Notions on the Common Core of Dark Traits

In light of the substantial empirical associations between different dark traits, several attempts have been made to explain their commonalities. Given the prominence of Machiavellianism, Narcissism, and Psychopathy, previous theoretical notions were particularly concerned with the links among the components of the Dark Triad or the "Dark Tetrad", which additionally comprises Sadism. Concerning the latter, Paulhus (2014) concluded in a review that callousness (the "lack of empathy toward others", p. 422) is the only shared feature. Callousness is typically understood as a deficit in emotional reactivity characterized by failure to respond to the distress cues of others and, consequently, problems in experiencing appropriate levels of guilt (Frick & White, 2008; Shirtcliff et al., 2009). The relevance of callousness has also been recognized in a description of the common features of

the Dark Triad traits by Jones and Figueredo (2013), who argued that the combination of "interpersonal manipulation and callous affect" (p. 529), and, thus, Hare's (2003) Factor 1 of Psychopathy (characterized by lack of guilt and empathy, shallow affect, glibness, failure to accept responsibility, manipulativeness, grandiosity, conning, lying), represents the shared characteristic of these traits.

A related and arguably broader approach to understanding the commonalities of the Dark Triad components has been proposed by Jonason and colleagues. Building upon life history theory (e.g., Wilson, 1975), they suggested that the Dark Triad traits boil down to a "fast" life history strategy facilitating "the opportunistic and strategic exploitation of one's environment to increase reproductive fitness" (Jonason, Webster, Schmitt, Li, & Crysel, 2012, p. 197). The Dark Triad components are thus seen as indicators of a stable and adaptive strategy directed towards immediate rewards and gratification, which, in turn, are associated with reproductive and survival benefits for an individual. In terms of specific psychological characteristics, the common features underlying the Dark Triad components are argued to reflect "both high levels of self-interest and low levels of empathetic concerns" (Jonason, Li, Webster, & Schmitt, 2009, p. 6), thereby mirroring the conceptualizations sketched above.

The Dark Triad traits have also been viewed through the lens of general models of basic personality structure. In particular, Lee and Ashton (2014) suggested that "the common element of the Dark Triad can […] be viewed as a willingness to exploit others when this is perceived to be advantageous" (p. 3), linking it to low Honesty-Humility, a basic personality trait of the HEXACO Model of Personality comprising aspects such as being deceitful, greedy, hypocritical, and sly (Ashton & Lee, 2007).

Slight differences in the emphasis placed on particular characteristics notwithstanding, previous attempts to describe the common features among the Dark Triad/Tetrad converge with regard to two aspects: low concerns towards others' utilities (callousness, exploitation) and some form of maximizing one's own utilities (self-interest, manipulation/exploitation of

others to promote own goal achievement). As such, the critical characteristics mentioned in recent attempts to describe the commonalities between the Dark Triad/Tetrad traits are well represented in our conceptualization of D as the basic tendency to maximize one's own utility at the expense of others.

However, our theoretical conceptualization of D not only encompasses the theoretical approaches sketched above, but is broader and also differs in several respects from previous accounts. First and foremost, D refers not only to the commonalities of the Dark Triad/Tetrad components, but to a general dispositional tendency underlying *all* dark traits. We argue that dark traits arise as specific manifestations of this more general underlying dispositional tendency, so that the commonalities among any set of dark traits can be traced back to D. Consequently (and unlike any of the attempts sketched above), we explicitly consider D to represent a fluid construct, i.e., that D appears in all combinations of a sufficient number of different indicators of dark traits in a form that mirrors our conceptualization.

D also differs from previous accounts in recognizing that one's own utility can be maximized by the very act of inflicting disutility on others, which may take the form of accepting own (e.g., financial) disadvantages in order to see others suffer (leading to positive feelings such as joy which may subjectively outweigh financial costs). Behaviors such as these are not easily reconciled with either of the accounts mentioned above, as it is neither obvious how this could be considered adaptive (in line with a fast-life strategy) nor plausible that it reflects a mere failure to respond to other's distress (callousness). Similarly, low Honesty-Humility predominantly implies that disutility for others results as a by-product of own utility maximization, rather than being a potential source of utility in and of itself. Relatedly, the concepts of utility and disutility in the definition of D are not limited to tangible benefits (such as money or status), but relate to any goal an individual may strive for, including the desire for psychological need fulfillment such as the need for belongingness and self-actualization. In this sense, actions that are directed towards preventing others' need

fulfillment (such as bullying) are well-aligned with our definition of D, whereas the relation is less clear for any of the alternative accounts mentioned above. Furthermore, the attitudes and beliefs held by individuals high in D that serve as conscious or unconscious justifications are not limited to a sense of entitlement, grandiosity, or own superiority (as, for example, specified in Hare's Factor 1 of Psychopathy or low Honesty-Humility). Rather, high-D individuals may hold a variety of justifying beliefs apart from believing in one's own superiority, for example by adopting a cynical world view, by endorsing ideologies that favor dominance (of individuals or groups), or by rejecting ideas concerning the universality of moral values in favor of a relativist world view.

It should also be noted that the notions of interpersonal manipulation or exploitation inherent in previous accounts concerning the commonalities among some dark traits arguably do not adequately represent the breadth of actions individuals high in D may engage in. For one, individuals high in D may also perform actions that, while negatively affecting others, do not involve any interpersonal manipulation—for example, behaviors aimed at individual utility maximization in a tragedy of the commons situation (Hardin, 1968; e.g., an entrepreneur poisoning the environment for higher profit). Second, both manipulation and exploitation are limited to actively performing certain actions (e.g., deceiving or exploiting someone); however, individuals high in D may provoke disutility for others entirely passively (e.g., refraining from helping someone), which can hardly be considered as manipulation or exploitation, but would nonetheless represent an instance of D.

In sum, the theoretical conceptualization of D as a basic dispositional tendency responsible for the commonalities among the dark traits can be regarded as a unification and extension of previous theoretical approaches as described above. Before critically testing this new and broader conceptualization of D as a fluid construct representing individual differences in maximizing one's utility at the cost of others while adopting justifying beliefs for corresponding behaviors, we next sketch how research on commonalities between dark

traits has typically involved four characteristics that need to be overcome so as to provide a sufficiently direct, comprehensive, and critical test of D.

Limitations of Previous Research on the Common Core of Dark Traits

Studies investigating the common core of dark traits have often been hampered by one or more of the following four limitations. First, corresponding research has mostly focused on a rather small subset of dark traits, especially on the Dark Triad or the Dark Tetrad only, and rarely investigated a larger set of dark traits simultaneously. Clearly, the proposal that D is a general dispositional tendency constituting the driving force behind specific dark traits cannot be strictly tested with such a limited set of dark traits: It is possible that commonalities appear overly large due to the consideration of a few dark traits only, but just as plausible that differences between them are overstated.

Second, it is striking that further dark traits are continuously introduced (for recent examples, see, Krekels & Pandelaere, 2015; Marcus et al., 2014; Moore et al., 2012; Seuntjens et al., 2015). Consequently, the lack of integrative attempts across several dark traits results in a plethora of constructs with few critical tests of whether an introduced trait is sufficiently different from existing ones. Instead, it is arguably essential to test the unique role of any "new" dark trait beyond other dark traits and, more importantly, beyond the commonalities of dark traits in general so as to scrutinize to what extent a newly introduced dark trait captures meaningful characteristics beyond others. This notion has only recently been emphasized in a review on dark traits by Marcus and Zeigler-Hill (2015) who argue that a distinction between Spitefulness (a newly introduced trait) and Psychopathy (a more classic dark trait) partly needs to "rest on whether Spitefulness can predict behavioral outcomes when controlling for Psychopathy" (p. 442).

Third, as is still typical for personality research (Baumeister et al., 2007; King, 2010), studies on dark traits have only sporadically included actual behavioral data (for exceptions, see e.g., Buckels et al., 2013; Campbell et al., 2004). This seems particularly problematic for

dark traits, because a substantial (and unknown) proportion of any association between some dark trait and a mere (self-)reported criterion may be due to response biases (such as socially desirable responding). In addition, beyond such response biases, there may well be particularly noteworthy differences between intended and actual behavior (e.g., Ajzen & Fishbein, 2005; Kaiser, Byrka, & Hartig, 2010). Indeed, it is quite thinkable that one may score high on a dark trait without necessarily showing corresponding behavior. For example, one may be drawn to situations in which harm is inflicted upon others in order to benefit personally, without actually engaging in harmful behavior due to the fear of sanctions or reputational concerns (which have been shown to prevent socially disruptive behavior, e.g., Balliet, Mulder, & van Lange, 2011; Fehr & Gächter, 2000), and/or the ability to channel one's problematic tendencies into a socially acceptable outlet (e.g., Schütte et al., in press).

Finally, research on the common core of dark traits has often been limited by the modeling approach employed. With few exceptions (Jonason, Kaufman, Webster, & Geher, 2013; Kajonius, Persson, Rosenberg, & Garcia, 2016), corresponding research has typically relied on statistical approaches that do not (or only indirectly) allow for (a) testing whether the communalities between different dark traits can be attributed to one single underlying factor (i.e., a common core), (b) whether this single factor captures psychologically meaningful variance, (c) determining the extent to which the dark core absorbs specific dark traits, and (d) scrutinizing the unique role of a particular dark trait beyond this common core. Previously used analysis strategies (considering the pattern of bivariate intercorrelations, Book, Visser, & Volk, 2015; employing higher order factor models and/or composite scores, Jones & Figueredo, 2013; Lee et al., 2013) fall short with regard to at least one or all four of these aims. In what follows, we thus test our conceptualization of D via a series of studies using a modeling approach particularly tailored for these purposes.

The Present Investigation

Comprising four studies, the present investigation aims at demonstrating and testing our theoretical proposal that dark traits are manifestations of a single, general tendency in line with our conceptualization of D. Specifically, we first examine whether the indicators of several dark traits can be modeled as and represented by one factor that mirrors our conceptualization of D, and explore the extent to which each of the several dark traits captures unique content beyond D so as to dissect more clearly the specifics of dark traits vis-à-vis D. Second, we investigate to what extent D, each dark trait, and the unique content of each dark trait beyond D predict a broad array of criteria in the realm of ethically, morally, and socially questionable behavior, including assessments of "actual behavior" (Baumeister et al., 2007; King, 2010). Third, we locate D within the space spanned by models of basic personality structure, that is, both the Big Five (McCrae & Costa, 2008) and the HEXACO traits (Ashton & Lee, 2007). Finally, using a simulation approach, we critically test our assumption that D reflects a fluid concept, that is, that D is invariant in content and predictive validity when including different (sub)sets of underlying dark traits (or dark trait measures) and when measured by a random subset of available items.

A Broader Set of Dark Traits

As reviewed above, previous attempts to determine the common core of dark traits were limited in the number of dark traits considered, to the best of our knowledge rarely exceeding the joint investigation of three or four traits (i.e., the Dark Triad/Tetrad). Thus, we aimed to identify and select a broader and arguably more representative set of dark traits for our investigation, relying on a two-step approach. First, we included the components of the Dark Triad, i.e., Machiavellianism, Narcissism, and Psychopathy, because they are widely accepted as typical examples of dark traits. Second, in early 2014, we searched several journals with links to personality research (*Assessment*, *European Journal of Personality*, *Journal of Personality*, *Journal of Personality and Social Psychology*, *Journal of Personality Assessment*, *Journal of Research in Personality*, *Personality and Individual Differences*,

Personnel Psychology, *Psychological Assessment*, *Psychological Science*) for articles that introduced a "new", negatively connoted trait with proposed links to ethically, morally, and socially questionable behavior. In this search, we focused on constructs that mirror the common understanding of general personality traits in that they are meant to be relatively stable across time (e.g., no state) and situations, reflect typical behavior, thoughts, and/or feelings, and are principally applicable to the entire population (e.g., not limited to clinical populations).

We started the search by reviewing the current issues from said journals (to include the most recently introduced dark traits) and then continued backwards in time up to the year 1999 (covering 15 years in total). Whenever we came across a potential dark trait that was introduced in another journal in this search (e.g., one that was used for investigating convergent validity of a trait that was introduced in one of our selected journals), we considered this trait as well.

This search strategy was pursued for two reasons. First, one would presume that more recently introduced dark traits will have been differentiated theoretically and/or empirically from previously introduced dark traits. That is, one would expect that "newly" introduced traits will have been argued and shown in some way to be unique and independent from existing ones (such as the components of the Dark Triad in particular). As a consequence, the search strategy arguably resulted in a set of diverse dark traits, in turn allowing for a stricter test of the proposal that D represents the common denominator of a variety of different dark traits. Second, this strategy arguably led to the inclusion of traits and corresponding measures that met the most current scientific standards for the theoretical and empirical development of traits and measures, respectively.

This approach ultimately led to the identification and selection of nine dark traits for our studies, namely Egoism, Machiavellianism, Moral Disengagement, Narcissism,

Psychological Entitlement, Psychopathy, Sadism, Self-Interest, and Spitefulness. These traits, their definitions, and corresponding inventories are listed in Table 1.

Modeling the Common Core: A Bifactor Approach

To test our theoretical ideas we relied on structural equation modeling. In particular, to test the notion of one single common core of dark traits, we resorted to a bifactor approach (e.g., Reise, 2012). In such an approach, each observed item is modeled to load both on a general factor (representing D in our case) that captures the commonalities among all items and on a specific factor (representing a specific dark trait in our case) that captures the remaining covariance among the items belonging to the respective scale that is not due to D (see Figure 1). Correspondingly, this approach decomposes the variance of an item into (1) variance that can be explained by the common general factor, (2) variance that is specific to the items of a scale, and (3) item-specific residual variance that is not explained by either the general or the specific factor. Importantly, by definition, the common general factor captures the variance that is shared across *all* items and hence directly represents their commonalities, whereas the specific factors reflect the remaining shared variance of the items of a particular scale, after controlling the general factor. In what follows, we use the terms "specific dark trait" or "residualized dark trait" to denote the dark traits in the bifactor specification, i.e., the specific factors arising when controlling for the common core of dark traits (D). In contrast, we use the term "non-residualized dark trait" to refer to the complete version of a dark trait without estimating or controlling for the effect of D, i.e., when estimating a standard confirmatory factor model for the dark traits without a bifactor structure.

By partitioning the items' variances into three sources, the bifactor approach offers several benefits over alternative modeling strategies (such as higher-order factor models). In particular, the bifactor approach makes it possible to (1) investigate whether a model comprising a dark core provides a superior account of the data as compared to a model without a dark core (and whether this core mirrors our conceptualization of D), (2) determine

the degree to which the dark traits are absorbed by D (versus their unique content beyond D), (3) test the predictive ability of the general factor (D) versus the specific factors (specific dark traits after removing the common core), and (4) investigate the proposed fluid nature of D in terms of whether and how alternative operational definitions affect the meaning and the predictive validity of D. As such, this modeling strategy is well suited to gain meaningful insights into the nature of the common core underlying dark traits as well as the unique content of dark traits over and beyond this dark core.

Nonetheless, it should be noted that both bifactor and higher-order factor models can be reasonably applied in situations where a general construct is to be identified that accounts for the common variance across items in the presence of specific constructs that also contribute to the variance across items (e.g., Brunner, Nagy, & Wilhelm, 2012; Yung, Thissen, & McLeod, 1999). Both models are members of the class of hierarchical factor models, with bifactor models at the more general end of the spectrum and higher-order factor models as nested special cases (subject to the proportionality constraint; Yung et al., 1999). However, despite this nesting relation, statistically distinguishing among these models is highly challenging in practice (Mansolf & Reise, 2017).

As elaborated in the theoretical part of this paper, we argue that D represents the major source of variance underlying individual differences on dark traits, so that D manifests itself in most ethically, morally, and socially questionable attitudes, beliefs, and behaviors. In other words, dark traits arise as specific manifestations of this general tendency, implying that D, as a hypothesized basic structure, exists independently from any one specific dark trait. These theoretical considerations, in turn, are more closely aligned with a bifactor approach than with a higher-order factor model: In the latter, D only exists through its relation to, but not independently of, the dark traits. Also, unlike bifactor models, a higher-order factor model would imply that D exerts its influence on the observed indicators only indirectly through its relation to the dark traits, with the latter directly determining the endorsement of the

associated items. However, the essential implication of D is that any item from any dark trait inventory is an expression of D. Thus, although we primarily relied on the bifactor approach due to its practical advantages over higher-order models, this approach also more closely maps onto the theoretical conceptualization of D.

Study 1

The first study was designed as a preliminary evaluation of the general idea that dark traits arise from a general underlying disposition and thus share a common core that, in turn, mirrors the conceptualization of D as specified above.

Participants and Procedure

The present and the following studies were conducted online to obtain heterogeneous and diverse samples, ensuring that there is sufficient variance in the dark traits. Note that research has consistently shown little to no effects of the medium (e.g., web vs. paper-pencil) on personality measurement (e.g., Weigold, Weigold, & Russell, 2013) and a whole host of other tasks and paradigms (including the behavioral paradigms we use in Studies 2 and 3, see Rand, 2012). In addition, web-based data-collection has the distinctive advantage of increased anonymity (see Joinson, 1999) which is highly relevant for characteristics, beliefs, and behaviors that are socially undesirable (Moshagen, Musch, Ostapczuk, & Zhao, 2010; Postmes & Spears, 1998). In all studies, we closely adhered to established standards of online experimenting (Reips, 2002).

In Study 1, participants were recruited by inviting members of an online panel to participate in a scientific study on personality. Participants completed the study on an anonymous and voluntary basis without any compensation. After providing informed consent and demographic information, participants were asked to complete the various questionnaires, as detailed below. Questionnaires were presented in a randomized order for each participant. The study took about 24 min to complete on average (*SD* = 13.6 min). The final sample

consisted of *N* = 304 participants (64% female, 1 'prefer not to say') with a mean age of 33.4 (*SD* = 11.8; range 18-72) years (14% drop-out rate). None of these participants showed any suspicious response behaviour defined as requiring less than 2 sec for each item on average or always selecting the same response option. All participants reported native or fluent German language skills, except for one participant who still indicated to have good German language skills.

Measures

The inventories to assess the nine dark traits selected for our studies are summarized in Table 1. Aiming to keep the total number of items within a reasonable range and to ensure that the selected dark traits are measured by an approximately equal number of items (to avoid that one trait dominates the remaining traits in terms of the number of indicator variables), we used validated short versions for assessing the traits, whenever available. Specifically, all measures comprised 8-10 items, with the exception of Egoism (12) and Spitefulness (17). Additional analyses (not reported here) using only the 10 highest loading items of Egoism and Spitefulness, respectively, yielded virtually identical results.

With regard to the components of the Dark Triad, we relied on one (Short Dark Triad; Jones & Paulhus, 2014) of two recently introduced instruments offering brief assessment of all three components instead of using a specific inventory for each component (such as the Machiavellian Personality Scale, Dahling, Whitaker, & Levy, 2008; the Narcissistic Admiration and Rivalry Questionnaire, Back et al., 2013; or the Psychopathy Personality Inventory-Revised, Lilienfeld & Widows, 2005). We opted for the Short Dark Triad inventory because it aims to operationalize Machiavellianism, Narcissism, and Psychopathy "without building in redundancy or forcing independence" (Jones & Paulhus, 2014, p. 29) between the constructs, and because it seems to be more reliable and valid (Lee et al., 2013) than another recently introduced brief instrument aiming to assess the Dark Triad, the "Dirty Dozen" (Jonason & Webster, 2010).

All studies were conducted in German language. Consequently, we translated all questionnaires for which there was no German version of the questionnaire available. In this translation process, we followed established standards such as the translation-retranslation technique using independent translators throughout (e.g., Brislin, 1980). Specifically, we translated the scales for Egoism, Moral Disengagement, Psychological Entitlement, Sadism, Self-Interest, and Spitefulness, while we could use the German version (Maaß & Ziegler, 2017) of the Short Dark Triad comprising the scales for Machiavellianism[2], Narcissism, and Psychopathy. A five-point Likert response scale ranging from 1 = strongly disagree to 5 = strongly agree was used for all measures. All negatively worded items were recoded prior to analysis so that higher values are indicative of a stronger manifestation of the underlying dark trait.

Analyses

As detailed above, we used a bifactor approach to investigate our theoretical prediction that the considered dark traits share a common core. An estimate of the extent to which the items indicating a specific construct reflect D can be obtained by comparing the variance of the items of a specific scale that is explained by D relative to the total explained common variance. The proportion of common variance explained by D was measured using

[2] Compared to the English original, the Machiavellianism scale of the German Short Dark Triad differs in three respects. First, item 2 of the English Machiavellianism scale ("I like to use clever manipulation to get my way") is replaced by the item "Generally speaking, people won't work hard unless they have to" (item 3 of the set used in constructing the English Short Dark Triad). Second, item 9 of the English version ("Most people can be manipulated") is replaced by the item "There's a sucker born every minute" (item 4 of the original set of items). Finally, the German scale contains one additional item ("Most people deserve respect"; reverse coded; item 13 of the original set of items).

the *ECV* (explained common variance) index, which is the ratio of variance explained by D to the variance explained by D *and* the respective specific factor (Ten Berge & Sočan, 2004). For example, if the *ECV* for the items of a specific dark trait is found to be equal to 1, the covariances of the items making up the scale are completely due to the D factor. Vice versa, if the D factor cannot explain any of the covariance of the items of a scale (yielding an *ECV* of 0), the conclusion would be that the construct in question has nothing in common with D.

In specifying the bifactor models, the correlations between the specific factors were constrained to 0 for identification purposes (thus enforcing that the shared variance between various specific factors is captured by D). All correlations between the residuals of the manifest indicators were fixed to 0, as were all secondary loadings. Each latent factor was assigned a scale by setting one unstandardized loading to 1. All analyses were based on the raw scores. The models were estimated using Mplus (version 7.11; L. K. Muthén & Muthén, 2015). Incomplete data (occurring in Studies 1 and 2) were addressed by employing full information maximum likelihood estimation, which has been shown to outperform other approaches in the context of structural equation modeling and to provide reasonable results even under conditions of non-normality (e.g., Enders, 2001; Enders & Bandalos, 2001; Yuan, Yang-Wallentin, & Bentler, 2012).

Whereas average levels of univariate skewness and kurtosis indicated a modest degree of non-normality (absolute skewness and kurtosis were on average 0.66 and 1.03, respectively), some items exhibited substantial to severe deviations of normality with a maximum skewness of 3.8 and a maximum kurtosis of 18.8. Thus, Huber-White sandwich estimated standard errors and corrected test-statistics (Yuan & Bentler, 2000) were applied to account for non-normality in the data (called “MLR” estimation in Mplus).

Moreover, as the models comprised a very large number of observed variables, in turn leading to inflated test-statistics (Herzog, Boomsma, & Reinecke, 2007; Moshagen, 2012), we additionally applied the correction proposed by Yuan, Tian, and Yanagihara (2015), yielding

a test-statistic corrected for both the effects of non-normality and model-size (corresponding to $T_{RML}^{(e1)}$ in Yang, Jiang, and Yuan, in press). Similar to the correction for non-normality, the correction for model size by Yuan et al. (2015) involves a linear transformation of the model test-statistic depending on the number of observed variables, the number of free parameters, and the sample size. The effect of this correction increases with both the number of manifest variables and the number of free parameters, but diminishes as the sample size increases (reflecting that the bias resulting from large models disappears asymptotically).

Given the extremely high power of the global model log-likelihood ratio (LLR) test to detect minuscule misspecifications even at moderate sample sizes (with $N = 300$, the power to detect misspecifications of the estimated models compared to the saturated model corresponding to *RMSEA* = .02 with $\alpha = .05$ is $1 - \beta > 99\%$; MacCallum, Browne, & Sugawara, 1996), we relied on two other criteria to evaluate the fit of the imposed models. First, in line with recent recommendations (Moshagen & Auerswald, in press), we considered the point estimates of the *SRMR* and the *RMSEA* (along with 90% confidence intervals) as descriptive indices of model fit referring to the convention that values for the *SRMR* around .08 and for the *RMSEA* around .06 indicate well-fitting models (Browne & Cudeck, 1992). Second, we adopted the recently proposed approach of Neyman-Pearson model testing with balanced error probabilities (Moshagen & Erdfelder, 2016). In this approach, the null hypothesis of a perfectly fitting model is evaluated against a specific alternative hypothesis reflecting an inacceptable degree of model violation, contingent on the requirement of balanced α and β error probabilities. The statistical decision in this testing context refers to whether the data rather support the assumption of a perfectly specified model (the null hypothesis) or the assumption of an unacceptable misspecification (the alternative hypothesis). To this end, a critical value for the LLR test statistic (χ^2_{crit}) is obtained based on the sample size, the model degrees of freedom, and the degree of misspecification considered unacceptable. In the present investigations, we defined an inacceptable degree of model

violation as models associated with an *RMSEA* of .08 or larger (Browne & Cudeck, 1992; Moshagen & Erdfelder, 2016). Since the thereby implied critical value χ^2_{crit} gives the boundary between an acceptable and an unacceptable model, the hypothesized model is considered tenable if the observed value of the LLR test-statistic falls below χ^2_{crit}.

Results and Discussion

First off, we determined the correlations among the nine non-residualized dark traits by fitting a confirmatory factor model positing one latent factor for each of the dark traits and allowing the factors to correlate freely. Although the global LLR model test-statistic was significant, $\chi^2(4149) = 6{,}406.26$, $p < .01$, descriptive indices of fit indicated an acceptable fit to the data, *SRMR* = .073; *RMSEA* = .050 (90% CI: .048 - .052). Further, the observed value of the χ^2 test-statistic fell well below the critical value associated with an inacceptable degree of model violation, $\chi^2_{crit} = 7{,}263$, indicating to retain the model.

The estimated correlations between the non-residualized dark traits, along with descriptive statistics and Cronbach's alpha estimates of internal consistency, are shown in Table 2. All non-residualized dark traits exhibited significantly positive associations with each other. The majority of correlations exceeded $r = .50$ (mean $r = .50$), corresponding to generally strong associations between the non-residualized dark traits and, thus, pointing to a substantial overlap between them. Particularly high correlations were evident between the non-residualized factors of Egoism and Machiavellianism ($r = .78$), Moral Disengagement ($r = .75$), and Psychopathy ($r = .73$), as well as between Psychopathy and Machiavellianism ($r = .80$), Sadism ($r = .79$), and Spitefulness ($r = .80$). In contrast, whereas Narcissism and Self-Interest were substantially correlated with each other ($r = .53$), both traits were less strongly (albeit still moderately) associated with the remaining dark traits, with the exception of a strong correlation with Psychological Entitlement ($r = .59$ and $r = .49$, respectively).

Given the substantial intercorrelations among the non-residualized dark traits, we proceeded by specifying the bifactor model (as described above), positing D as a common,

general factor and one specific factor for each of the dark traits to evaluate the theoretical prediction that the considered dark traits share a common dark core. The model exhibited an adequate fit to the data, $\chi^2(4092) = 6{,}224.72$, $p < .01$; $\chi^2_{crit} = 7{,}137$; *SRMR* = .069; *RMSEA* = .049 (.047 - .051). Beyond improved overall model fit, information criteria (e.g., Bollen, Harden, Ray, & Zavisca, 2014) also suggested that the bifactor model (including D) is to be preferred over the confirmatory factor model. The *AIC*s were 69,757 vs. 69,920, the *BIC*s were 69,960 vs. 70,091, and the normalized evidence ratios (Wagenmakers & Farrell, 2004) of the bifactor model were > .999, indicating that the bifactor model has the highest probability among the models considered in terms of the Kullback–Leibler discrepancy to the true model (*AIC*) and the degree of belief that it reflects the true model (*BIC*), respectively. Thus, the hypothesized bifactor model generally appears tenable and provides a superior description of the data compared to a model omitting D. In other words, the superior model is the one that constrains the *entire* shared variance across all nine dark traits to be due to *one* underlying factor, D.

To further scrutinize this conclusion, we also tested an oblique bifactor model positing two (rather than one) general factors along with the nine specific factors. This model was motivated by the observation that the correlations between the non-residualized dark traits as displayed in Table 2 suggest that Narcissism and Self-Interest are substantially associated with each other, but show comparatively weaker associations to the remaining traits, except for Psychological Entitlement. Thus, we specified one general factor comprising the items related to Narcissism, Self-Interest, and Psychological Entitlement, another general factor for the remaining items, and the nine specific factors for the dark traits. However, the model yielded a psychometrically improper solution. The general factor comprising Narcissism, Self-Interest, and Psychological Entitlement showed a very small variance which did not differ significantly from zero ($p = .21$). This indicates that there is no systematic individual variation on this factor, and thus speaks against the validity of this model. In sum, the model

comparisons confirm that the bifactor model positing D as the only general factor and one specific factor for each of the dark traits provides the most sensible representation of the data. We thus base the following considerations on this solution.[3]

To evaluate the extent to which the nine dark traits share a common dark core, we first considered the estimated (unstandardized) variances of the factors. All factor variances were significantly larger than zero (ranging from .17 to .39), indicating that D and all specific factors explain covariances between at least some items and must thus be kept in the model. We proceeded by evaluating the pattern of loadings (see Table 3 for a summary; detailed loading estimates are available as supplementary materials). Of the 93 items, 85 items exhibited a significantly positive association with D, with the standardized loadings ranging from -.07 to .66. The indicators with the ten strongest loadings on D comprised items of the scales of five different dark traits (Egoism, Machiavellianism, Psychopathy, Sadism, and Spitefulness). Item loadings on the specific factors ranged from -.14 to .68. Sixteen loadings on the specific factors did not differ significantly from zero. The loading estimates are thus in line with the factor variances in suggesting that a common dark core exists, which, however, does not completely obviate any of the specific factors.

We next evaluated whether the emerged general factor mirrors our theoretical conceptualization of D by inspecting the items that exhibited the strongest loadings on D. As expected, these items related to multiple aspects of D or one of the aspects in particular. For example, the utility maximization aspect is reflected in “Never tell anyone the real reason you did something unless it is useful to do so” (one of the items designed to assess Egoism; loading on D was $\lambda_D = .61$; loading on the specific factor for Egoism was $\lambda_S = .03$), “I’ll say

[3]We also tested the model positing two general factors in Studies 2 and 3. However, estimation of the model involving two general factors did not converge successfully in either study.

anything to get what I want" (Psychopathy; λ_D = .57; λ_S = -.14), and "Make sure your plans benefit you, not others"(Machiavellianism; λ_D = .55; λ_S = .24). Provoking disutility for others is reflected in "There have been times when I was willing to suffer some small harm so that I could punish someone else who deserved it" (Spitefulness; λ_D = .60; λ_S = .29), "Hurting people would be exciting" (Sadism; λ_D = .57; λ_S = .62), and "I like to get revenge on authorities." (Psychopathy; λ_D = .55; λ_S = .20). Finally, justifying beliefs are reflected in "There's a sucker born every minute" (Machiavellianism; λ_D = .59; λ_S = .08), "I honestly feel I'm just more deserving than others" (Psychological Entitlement; λ_D = .54; λ_S = .45), and "The biggest difference between most criminals and other people is that criminals are stupid enough to get caught" (Egoism; λ_D = .49; λ_S = .43). As these examples illustrate, the majority of the items strongly loading on D exhibited non-significant or small loadings on their respective specific factor. However, mirroring the *ECV*s discussed in detail below, there were also notable exceptions in particular with regard to items from the scales assessing Sadism and Psychological Entitlement. Nevertheless, beyond supporting our characterization of D, the fact that items referring to such diverse attitudes, beliefs, and behaviors (and being used to measure diverse traits) show similar loadings on D provides evidence for the theoretical idea that a common dispositional tendency is the driving force for individual differences on dark traits.

Of further importance is the question whether D represents all the dark traits considered to an approximately equal extent or whether certain dark traits are more or less completely absorbed by D, whereas others have little in common with D. In order to evaluate the degree to which the dark traits are reflected in D, we considered the proportion of common variance explained by D (*ECV*). Table 4 shows that the magnitude of D saturation substantially varies across the considered dark traits. D is primarily reflected in items related to Psychopathy (*ECV* = .74) and Machiavellianism (*ECV* = .73), explains approximately half

of the common variance in the items indicating Spitefulness (*ECV* = .58), Egoism (*ECV* = .57), Moral Disengagement (ECV = .54), Sadism (*ECV* = .50), and Psychological Entitlement (*ECV* = .38), but only modestly accounts for the items measuring Narcissism and Self-Interest (both *ECV*s = .18). Thus, the *ECV*s suggest that the latter traits capture a substantial proportion of unique variance that is not represented in D. We defer a more detailed discussion of similarities and differences between D and the specific traits to Study 3.

Finally, an important question pertains to whether D may be primarily defined because different scales used to assess the dark traits contain highly similar items. For example, the scales assessing Machiavellianism and Psychopathy, respectively, contain similar items related to revenge, the scales assessing Machiavellianism and Egoism, respectively, items related to disclosing information to others, and the scales assessing Egoism, Machiavellism, and Psychopathy, respectively, items related to the manipulation of others. Clearly, the very theoretical proposal put forward in this manuscript actually requires that certain characteristics are shared across constructs (and, thus, may be represented in the scales used as corresponding measures), so it is hardly surprising to find that different scales contain items assessing similar characteristics. Moreover, one may safely assume that the corresponding scale development proceeded by choosing items in line with the theoretical meaning of the targeted dark trait, entirely independent of the theoretical conceptualization of D. Consequently, the presence of similar items in the questionnaires assessing different dark traits may simply suggest that common characteristics are deemed of relevance for different traits.

Whereas the issue of similar items in different scales is therefore neither surprising nor problematic from a theoretical view, it might still pose an empirical threat in the measurement of D in that the variance of the general factor deemed to reflect D might be spuriously inflated. To investigate this issue, we first identified the items that arguably measure the same characteristic by theoretically inspecting the item content and by closely examining the

modification indices associated with the confirmatory factor model of the non-residualized dark traits. The latter was based on the rationale that a strong content overlap between items of different scales would be evident in non-zero correlations between the item's residuals (i.e., such items exhibit stronger correlations than would be expected based on the correlation of their underlying factors). Using this procedure, we identified six item pairs (or triplets) that potentially address highly similar characteristics (e.g., "Never tell anyone the real reason you did something unless it is useful to do so" from the scale assessing Egoism, and "It's not wise to tell your secrets" from the scale assessing Machiavellianism). We then estimated the bifactor model assuming D as a common, general factor along with the 9 specific factors (as described above) while freely estimating the covariances between the residuals of these six pairs or triplets of items. Results indicated that D actually exhibited an even larger variance in the model including the residual correlations compared to the model without these correlations. The *ECV*s, the item loadings on the D factor, and mean loadings were also highly similar. These results allow for the conclusion that D is not simply the result of different scales sharing identical items.

Taken together, the first study can be interpreted as providing initial evidence for the notion that dark traits are manifestations of a general, underlying tendency and thus share a common dark core that, in turn, is well-aligned with our definition of D. Clearly, however, this study is limited in several ways. First, the sample size was comparatively small, so a larger sample is required to increase the confidence in the findings. Second, the modeling results presented here only show that the considered dark traits share common variance that can be interpreted in line with our conceptualization of D; however, the question whether this common core is able to predict relevant criteria in a psychologically meaningful way remains to be studied. Correspondingly, we conducted a second study that relied on a larger sample and included a criterion measure to replicate and extend these initial findings.

Study 2

The second study (1) served to replicate the results of the first using a larger sample and (2) used behavioral selfishness (versus altruism) as measured by the dictator game (DG; e.g., Forsythe, Horowitz, Savin, & Sefton, 1994) as a criterion to obtain an indicator of the validity of D and to gauge the relative ability of D versus the specific factors to predict an outcome in the realm of ethically, morally, and socially questionable behavior. In the DG, one individual (the dictator) is given an endowment and asked to allocate this endowment between her-/himself and another, unknown person (the recipient) in a one-shot anonymous setting. Thus, the allocation of the endowment is broadly considered a measure of selfishness versus altruism (Andreoni & Miller, 2002) and has been used correspondingly in numerous investigations (for an overview, see Engel, 2011). Importantly, because the DG is a zero-sum game it is perfectly suited to test the utility aspects of D: The dictator's individual utility can only be maximized at the cost of the recipient.

Participants

Participants were recruited via a large web portal that hosts links to all kinds of tests and questionnaires (www.testedich.de). The study was advertised as a scientifically constructed questionnaire assessing peoples' personality. Participants did not receive any monetary compensation for completing the study, but were offered feedback on their individual trait scores at the end. Participation was anonymous and voluntary. Of the 1,845 participants who initially started working on the study, 1,095 participants reached the final page, corresponding to a retention rate of 59%. One participant was excluded based on indicators of suspicious response behavior (see Study 1), yielding a sample of N = 1,094 participants (64% female, 5 "prefer not to say"). Age ranged from 18 to 68 (M = 33.4; SD = 11.8) years. Most of the participants (98%) reported native or fluent German language skills. Excluding participants reporting good (15) or basic language skills (2) did not affect the

results, so we opted to retain them in the sample. All participants provided informed consent. On average, the study took approximately 27 minutes to complete (*SD* = 13.8 min).

Measures and Procedure

We used the same questionnaire measures as in Study 1. After completing the questionnaires (presented in random order), participants were requested to complete a hypothetical dictator game scenario. Specifically, participants were asked to imagine they had been randomly paired with another, unknown individual (whom they would never knowingly meet and who would never knowingly meet them) for a one-shot decision task. In this task, they were going to split an endowment of 100 Euros between themselves and the other person—in whatever way they might want to. Finally, participants specified how much of the endowment they wanted to give to the other person (and, in turn, how much would be kept to themselves).

Results and Discussion

The modeling strategy was identical to the one used in Study 1. We first fitted a confirmatory factor model positing one latent variable for each non-residualized dark trait (without a bifactor structure) to determine the correlations between the nine non-residualized dark traits. Fit-indices indicated to retain the model, $\chi^2(4,149) = 11{,}655.15$, $p < .01$; $\chi^2_{crit} = 14{,}873$; *SRMR* = .059; *RMSEA* = .042 (.041 - .043). The correlations between the non-residualized dark traits obtained from confirmatory factors analysis are shown in Table 5. As can be seen, the dark traits exhibited a highly similar pattern of associations as in the first study. Again, the majority of correlations exceeded $r = .50$ (mean $r = .50$) and the strongest correlations were observed across Egoism, Machiavellianism, Moral Disengagement, Psychopathy, Sadism, and Spitefulness, whereas Narcissism and Self-Interest were less strongly associated with the remaining dark traits; again with the exception of Psychological Entitlement, which, once more, showed comparable associations with all traits.

Next, we proceeded by fitting the bifactor model positing D as a general factor and one latent variable for each of the dark traits as specific factors. Model fit slightly improved compared both to the confirmatory factor model and to the corresponding model in the first study, $\chi^2(4{,}092) = 11{,}360.03$, $p < .01$; $\chi^2_{crit} = 14{,}407$; *SRMR* = .059; *RMSEA* = .041 (.040 - .042), suggesting that the model provides an adequate representation of the data. Likewise, information criteria again indicated to prefer the bifactor model (including D) over the confirmatory factor model (without D), *AIC*s = 268,670 vs. 268,988, *BIC*s = 269,348 vs. 269,562, respectively, and the evidence ratios were also in favor of the bifactor model (> .999). Thus, these results are in line with the findings of Study 1 in suggesting that a bifactor model involving a single general factor representing D provides the most sensible representation of the data. In this model, all factor variances were significantly larger than zero (ranging from .14 to .51). Loadings were comparable to the ones obtained in Study 1 (Table 3). The same was also largely true regarding the extent to which the dark traits were captured by D, as expressed through the *ECV* (Table 4). Compared to the first study, however, D accounted for a slightly smaller proportion of the common variance of the indicators of Narcissism and Self-Interest (*ECV*s were .12 and .13), but explained a substantially larger proportion of the common variance of the items pertaining to Moral Disengagement, Sadism, and Spitefulness (*ECV*s were .61, .58, and .70, respectively). In essence, these results confirm the conclusions put forward in Study 1 in providing evidence for the existence of a common dark core across the considered dark traits.

Beyond replicating the results of the first study in an independent and larger sample, the second purpose of Study 2 was to offer an opportunity to gain further insights into the psychological meaning of D by linking it to a criterion representing selfishness (versus altruism), namely DG choice. On average, participants decided to keep 59.5% (*SD* = 18.1) of the endowment for themselves, thereby indicating an overall trend towards selfishness, closely mirroring typical responses in the DG (in Engel's, 2011, meta-analyses across over

20,000 participants, the mean proportion kept is about 60-70%). All nine non-residualized dark traits were significantly positively associated with selfishness (i.e., the amount of money kept), with the bivariate latent correlations ranging from $r = .17$ for Self-Interest to $r = .34$ for both Machiavellianism and Moral Disengagement (Table 5).

In order to investigate the relation between D and DG choices and whether the specific dark traits are able to predict said choices beyond D, we regressed DG selfishness on D and the specific factors (as defined in the bifactor model). Note that, as dictated by the structure of the model, D and the specific factors are orthogonal and the specific factors are mutually uncorrelated, thereby ruling out potential caveats due to multicollinearity. Moreover, in this situation, the square of a regression coefficient directly corresponds to the unique contribution of the respective variable in explaining variance in the criterion (Cohen, Cohen, West, & Aiken, 2003). The results are reported in Table 6. D was a strong predictor of DG selfishness, accounting for approximately 15% of the variance, $\beta = .38$, $p < .01$. This result can be interpreted as providing further evidence for our conceptualization of D as a psychologically meaningful trait that reflects the tendency to maximize one's individual utility, disregarding, accepting, or malevolently provoking disutility for others. Importantly, of the specific traits, only Psychological Entitlement significantly contributed to the prediction of the DG choice, $\beta = .09$, $p < .05$, albeit with a more modest effect size (accounting for less than 1% of the DG variance). The remaining eight dark traits considered did not predict behavior in the DG once the common dark core, as measured through D, was controlled for. In other words, D absorbed most of the variance relevant for the prediction of selfishness (as measured in the DG).

In sum, Study 2 successfully replicated and extended Study 1 by showing that the nine considered dark traits exhibit a substantial overlap that can be accounted for by a single dark core. This dark core, in turn, predicted a criterion of immediate theoretical relevance, whereas only one of the specific dark traits modestly contributed to the prediction of selfishness. These

results notwithstanding, a number of caveats remain. Most importantly, relying on self-report data only, the results of this study are subject to the rival interpretation that the shared variance among the items assessing the dark traits merely reflects some measurement artifact such as acquiescence, negative self-views, or the tendency to distort responses in a socially desirable manner (Bardsley, 2008; Hilbig, Thielmann, Hepp, Klein, & Zettler, 2015; Paulhus, 2002; Ziegler, MacCann, & Roberts, 2012). Moreover, the dictator game and the dark traits were concurrently assessed in a single session, which might have inflated the observed associations due to consistent responding. Finally, this study was limited in considering only one single criterion that does not cover the breadth of ethically, morally, and socially questionable behavior that D is meant to account for. To gauge the validity of our conceptualization of D more comprehensively and to gain further insights regarding whether (and which) specific traits carry substantial meaning beyond D, it is thus mandatory to investigate the relation of D and the specific factors to a number of further external criteria—most importantly, actual, consequential behavior—in a longitudinal design.

Study 3

The third study had four aims. The first aim was to replicate the results of Studies 1 and 2 using a more diverse sample and a longitudinal design. The second aim was to extend the criteria from a single hypothetical task to two monetarily incentivized measures, one capturing selfishness (again, DG choice) and one capturing unethical behavior (namely, cheating to maximize one's gains), so as to support a substantive interpretation of D. The third aim was to evaluate the theoretical meaning and importance of D as well as the specific dark traits more comprehensively by considering nine additional self-report criteria. These criteria were selected from the studies originally proposing each of the nine dark traits considered. Specifically, for each dark trait, we selected the criterion measure that had been shown to correlate most strongly with said trait. The final aim was to locate D within the

space spanned by broad models of basic personality structure, namely, in terms of both the Five-Factor Model (FFM; McCrae & Costa, 2008) and the HEXACO Model of Personality (Ashton & Lee, 2007). Considering D in relation to basic dimensions of personality allows for further substantiating the theoretical meaning of D and for examining whether D comprises meaningful variance beyond basic models of personality. Relatedly, it is of key interest to determine whether D can be reduced to a single basic personality trait (considering Honesty-Humility of the HEXACO model in particular – see the theory section of this paper) or whether D is best regarded as reflecting a blend of various basic personality traits.

Participants

In Study 3, we relied on a professionally managed online panel to recruit a large sample that would be roughly representative of the general population. Participants were re-invited to complete a total of three measurement occasions (T1 to T3). At T1, participants completed two broad personality inventories. At T2, participants completed the same measures as in Studies 1 and 2 to assess the nine dark traits (and, thus, D). At T3, participants completed nine self-report criteria and, afterwards, two incentivized behavioral tasks.

A total of 2,032 participants completed the measurement at T1 without any indication of suspicious response behavior (see Study 1). Five weeks after the first measurement occasion, these participants were re-invited by the panel provider for the measurement at T2. A total of 1,366 participants responded within 10 days (which was the limit set) and completed all tasks (drop-out-rate from T1 to T2: 33%). The average time lag between the measurements at T1 and T2 was 37 (*SD* = 4.8) days. We then excluded 105 participants (exclusion-rate: 8%) showing suspicious response behavior (see Study 1), resulting in 1,261 participants who were re-invited by the panel provider for the measurement at T3. This re-invitation was sent out six weeks after the end of T2, and we once more set a limit of 10 days for participation. The measurement at T3 was completed by 995 individuals (drop-out-rate from T2 to T3: 21%), on average 43 (*SD* = 4) days after the end of T2. Data matching

between the measurement occasions was achieved through completely anonymous random codes and double-checked using demographics. Upon completion of T3, we had a total of 928 unique data sets that could be conclusively matched across all three occasions (exclusion-rate: 7%).[4] We finally excluded 46 data sets (exclusion-rate: 5%) of participants due to suspicious response behavior at T3, ending up with 882 datasets which constitute our final sample.

Participants in the final sample (46% female) covered a broad range of ages (18 and 65 years, M = 42.5, SD = 13 years). All reported to be native (96%) or fluent (4%) in German. A majority of 65% was in employment, whereas less than 8% were students. As intended, there was also substantial diversity in educational backgrounds with 23% holding a certificate of secondary education (German: Hauptschulabschluss), 20% a general certificate of secondary education (German: mittlere Reife), 25% a vocational diploma or university-entrance diploma (German: Fachabitur or Abitur), and 32% a university/college degree. There was no indication of selective drop-out across measurement occasions for any of these variables.

Every measurement occasion started with participants providing informed consent and demographics. Then followed the specific tasks and measures (see below) and finally participants were thanked, debriefed, and redirected to the panel provider who handled payment. Participants received a flat fee (determined by the panel provider) for every measurement occasion completed and additional incentives depending on their (and others') responses at T3 (see below).

Measures

[4] Conservatively, we excluded 65 data sets that could have been produced by multiple participation (as indicated by identical IP addresses and demographics) or that were implausible upon inspection of demographics across measurement occasions (e.g., if someone indicated a younger age at a later measurement occasion).

The German version (Borkenau & Ostendorf, 1994) of the 60-item NEO-Five-Factor-Inventory (NEO-FFI, Costa & McCrae, 1992) and the German version (Moshagen, Hilbig, & Zettler, 2014) of the 60-item HEXACO Personality Inventory-Revised (HEXACO-60; Ashton & Lee, 2009) were used to measure participants' basic personality traits at T1. The NEO-FFI measures the five dimensions of the widely known FFM (Openness to experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism; see, e.g., Costa & McCrea, 2008). The HEXACO-60 measures the six dimensions of the HEXACO model (Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience). The inventories were presented in random order. Responses were made on a five-point Likert response scale ranging from 1 = strongly disagree to 5 = strongly agree.

At T2, the same dark trait measures as in the preceding studies were assessed, again presented in random order and with the same five-point response scale as before. Finally, at T3 participants completed nine self-report criterion measures (in random order) followed by two incentivized behavioral tasks. Selection of the nine self-report criterion measures was guided by the following rationale: First, we aimed to investigate whether D can predict a broad range of variables in the realm of ethically, morally, and socially questionable behavior. Second, as a more critical test, we aimed to investigate whether the specific dark traits (i.e., the dark traits beyond D) predict relevant criteria over and beyond their common dark core. To this end, we inspected the articles introducing the nine dark traits under scrutiny (see Table 1) and reviewed all reported correlations of the respective dark trait with other variables that served as criterion measures for the trait under consideration. It is reasonable to assume that the criterion measures in the studies introducing the dark traits were chosen so as to be well-aligned with the construct being introduced such that a meaningful nomological net for said construct was obtained. In order to select one particularly well-aligned criterion measure for each of the dark traits in a non-arbitrary manner, we aimed to select the criterion variable (and corresponding measurement instrument) that showed the highest correlation with the

respective dark trait.[5] Both Machiavellianism and Psychopathy showed its highest correlation with the same criterion measure (Nurturance) in the article introducing the Short Dark Triad (Jones & Paulhus, 2014). Therefore, we aimed to assess one further criterion that should differentiate between these dark traits. Based on the notion by Jones and Paulhus (2014) that "the element of impulsivity is key in distinguishing Psychopathy from Machiavellianism" (p. 29), we included a measure of Impulsivity (in addition to a measure of Nurturance). This procedure led to the selection of nine self-report criterion measures as summarized in Table 7.

Aggression. Aggression was selected based on its correlations with Spitefulness ($.52 \leq r \leq .58$) in Marcus et al. (2014), and was assessed via the 40-item Forms and Functions of Aggression Scale (Little, Henrich, Jones, & Hawley, 2003).

Dominance. Dominance was selected based on its correlation of $r = .65$ with Narcissism in Jones and Paulhus (2014), and was assessed via the respective 32-item scale of the IPIP-IPC (Markey & Markey, 2009).

Impulsivity. Impulsivity was selected based on the description by Jones and Paulhus (2014) to be the key distinguishing variable between Machiavellianism and Psychopathy. Impulsivity was not assessed in the article introducing the Short Dark Triad (Jones & Paulhus, 2014), so we selected a corresponding inventory from a different article (Jones & Paulhus, 2011). Specifically, we administered the 12-item Dysfunctional Impulsivity subscale of Dickman's (1990) impulsivity inventory, showing correlations between Psychopathy and overall Impulsivity of $.32 \leq r \leq .52$ (Jones & Paulhus, 2011).

---

[5] We deviated from this rule (and selected the outcome showing the next highest correlation) when the highest correlation was reported for links with other dark traits that are included in our operational definition of D (e.g., among the components of the Dark Triad) or when the highest correlation was reported for a variable that could not be administered in our study (which only occurred concerning the criterion for Egoism).

Insensitivity. Insensitivity was selected based on its correlation of r = -.41 with Sadism in O'Meara et al. (2011), and was assessed via the respective 16-item subscale of the Empathy Quotient developed by Baron-Cohen and Wheelwright (2004).

Internalized Moral Identity. Internalized Moral Identity was selected based on its correlations with Moral Disengagement ($-.55 \leq r \leq -.42$) in Moore et al. (2012), and was assessed via the respective 5-item scale of the measure of Moral Identity described by Aquino and Reed (2002). Therein, a description of a "moral" person (e.g., caring, fair, kind) is presented to participants who are then asked to indicate their level of agreement on a 5-point Likert scale to statements referring to this description.

Nurturance. Nurturance was selected based on its correlations with both Machiavellianism (r = -.43) and Psychopathy (r = -.49) in Jones and Paulhus (2014), and was assessed via the respective 32-item scale of the International Personality Item Pool-Interpersonal Circumplex (IPIP-IPC, Markey & Markey, 2009).

Perspective Taking. Perspective taking was selected based on its correlation of r = -.43 with Psychological Entitlement in Campbell et al. (2004), and was assessed via the respective 7-item subscale of Davis' (1983) Interpersonal Reactivity Index.

Power. Power was selected based on its correlation of r = .65 with Self-Interest in Gerbasi and Prentice (2013) and was assessed via the respective 4-item subscale of Schwartz' Value Survey (e.g., Schwartz, 1992). Therein, respondents are asked to indicate the degree to which different values function as guiding principles for their lives on a 9-point Likert-scale ranging from -1 = opposite to my values to 7 = extremely important.

Self-Centeredness. Self-centeredness was selected based on its correlations with Egoism ($.35 \leq r \leq .47$) in Weigel et al. (1999), and was assessed via four items introduced by Grasmick, Tittle, Bursik, and Arneklev (1993).

Except for the scales assessing Moral Identity and Power, respectively, all measures use a Likert-scale as response format without any particular instruction, so we used a Likert-

based response format ranging from 1 = strongly disagree to 5 = strongly agree for all these measures to maintain consistency. We translated measures for which no German translation was available (Aggression, Impulsivity, Insensitivity, Internalized Moral Identity, Perspective Taking, and Self-Centeredness) using the same approach as described above (i.e., translation-backtranslation technique with independent translators).

Administration of the measures proceeded in the following order: First, Internalized Moral Identity and Power were presented (given their non-standard questionnaire and response format) in randomized order. Second, the remaining seven self-report criterion measures (Aggression, Dominance, Impulsivity, Insensitivity, Nurturance, Perspective Taking, and Self-Centeredness) were presented, again in randomized order. Finally, the two behavioral tasks—first cheating, then DG choice—were presented.

In the first behavioral task, participants completed an incentivized cheating game that has been widely used in behavioral ethics research, the coin-toss-task. This is a structural equivalent of the "die under the cup" paradigm (e.g., Fischbacher & Föllmi-Heusi, 2013; Shalvi, Handgraaf, & De Dreu, 2011), and has been successfully adapted for use in online settings before (Hilbig & Zettler, 2015; Zettler, Hilbig, Moshagen, & de Vries, 2015). In the variant we used, participants were instructed to take a coin, to choose a target side (heads or tails), to toss the coin exactly twice, and to note down or memorize how often the target side came up. Participants were informed that they were going to be asked to report whether the target side came up exactly twice in two tosses, which was equivalent to "winning" and incurred an additional payoff of 5 Euros. They were also told that the probability of this happening by chance is 25% (as by the binomial distribution), and were assured that the additional payoff depended solely on their response. Finally (on a separate page after providing the information about the task), participants selected one of two response options, namely "yes, the coin came up on the target side exactly twice in exactly two tosses (so I will

receive an additional 5.00€ for this task)" or "no, the coin did not turn up on the target side exactly twice in exactly two tosses (so I will not receive an additional 5.00€ for this task)".

The major advantage of this behavioral cheating paradigm is that it both provides an incentive and an opportunity to cheat, that is, to misreport in order to maximize payoffs without any dangers of self-incrimination. Since any one "yes"-response can always stem from an honest individual who got lucky, participants' anonymity is protected (Moshagen, Hilbig, Erdfelder, & Moritz, 2014). At the same time, the aggregate probability of honest "yes"-responses is conclusively known (namely, 25%) and it is thus possible to estimate the extent of cheating on the aggregate (for details, see Moshagen & Hilbig, 2017).

In the second behavioral task, we again used the DG (as in Study 2), but this time in a fully incentivized version: Participants were provided with an endowment of 5 Euros that they were asked to allocate at will between themselves and another, anonymous person who was randomly chosen among the other participants of this study, thereby creating a dyadic-like situation. It was emphasized to participants that only their response mattered and that they would receive an additional payoff corresponding exactly to the amount they decided to keep. Finally, it was guaranteed to participants that the allocation was completely "real" not only for themselves, but also for the other person. Participants then selected the amount they wanted to keep versus give from a drop-down menu (specifying every amount from 0.00€ to 5.00€ in 0.50€-increments).

Note that both behavioral tasks were fully consequential and incentive-compatible, thus countering strong influences of response distortions such as social desirable or consistent responding—as probably did the time delay between the measurement occasions as well as the increased anonymity provided by the fact that payment was handled by the panel provider who was unaware of the task content (and merely informed of participants' final additional payoffs by us). Anonymity (especially to the extent that even the response given cannot be in any way incriminating—as in the cheating task used) will render any motivation to appear

fair, honest, kind etc. towards others essentially irrelevant. Providing monetary incentives, in turn, will ensure that any remaining desire to retain a socially desirable self-image will at least have to compete with the desire to maximize one's monetary payoff. Therefore, anonymous, incentivized tasks are commonly considered to be far less prone to socially desirable responding if not even completely unaffected (Thielmann, Heck, & Hilbig, 2016). Consequently, these tasks essentially measure "actual" or "real" behavior (Baumeister et al., 2007; Hilbig, Glöckner, & Zettler, 2014; King, 2010) in the sense that people's responses had corresponding material consequences without triggering any reputational concerns. In turn, associations between a self-report measure and anonymous, incentivized tasks cannot be a spurious result of socially desirable responding.

Nonetheless, since the behavioral tasks were assessed directly after the self-report criterion measures, effects of consistent reporting may still be present; however, this would primarily apply to the scales assessing the criterion variables, rather than the scales assessing the dark traits and thus D (which were assessed approximately 6 weeks in advance). Also, as argued above, the anonymity and monetary payoffs associated with the behavioral tasks should once more alleviate this concern.

Results and Discussion

The general analytic strategy and model specification were largely identical to Studies 1 and 2. Unlike in the preceding studies, there were no missing data (because participants were technically required to provide a response in order to proceed), so we used Satorra-Bentler (2001) corrected test-statistics and standard errors to address the effects of non-normality in the data, again in conjunction with the correction for model size proposed by Yuan et al. (2015). Table 8 shows the bivariate correlations among the nine non-residualized dark traits obtained from fitting a confirmatory factor model (without bifactor structure), $\chi^2(4,149) = 11{,}625.62$, $p < .01$; $\chi^2_{crit} = 16{,}626$; *SRMR* = .056; *RMSEA* = .039 (.038 - .041). Compared to the previous studies, the intercorrelations were less pronounced in both

directions: previously very strong correlations were slightly weaker, whereas previously rather low correlations were slightly stronger. Overall, however, the traits exhibited a strikingly similar pattern of correlations to the ones observed in the preceding studies with the majority of correlations again exceeding $r = .50$ (mean $r = .49$). Correspondingly, these results provide additional evidence concerning the existence of a common underlying characteristic.

The bifactor model positing D as a general factor along with the nine specific factors for the dark traits exhibited a good fit to the data, $\chi^2(4{,}092) = 11{,}352.90$, $p < .01$; $\chi^2_{crit} = 16{,}377$; *SRMR* = .059; *RMSEA* = .038 (.038 - .039). Likewise, this model was to be preferred over the confirmatory factor model without D in terms of information criteria, *AIC*s = 273,140 vs. 273,365 and *BIC*s = 273,870 vs. 273,983, and the evidence ratios (both > .999). All factor variances differed significantly from zero (range: 0.05 - 0.42) and the estimated loadings (Table 3) were in strong agreement with Studies 1 and 2. Concerning the extent to which D absorbs the dark traits, the *ECV*s (Table 4) were slightly lower compared to the preceding studies, but nonetheless followed a very similar pattern: D explained the majority of the common variance in the items referring to Psychopathy (*ECV* = .75), Spitefulness (*ECV* = .62), Moral Disengagement (*ECV* = .61), Machiavellianism (*ECV* = .60), and Egoism (*ECV* = .57), accounted for a substantial proportion of the common variance in the items related to Sadism (*ECV* = .43) and Psychological Entitlement (*ECV* =.38), but was only modestly reflected in the items of Narcissism and Self-Interest (both *ECV*s = .18). In sum, Study 3 confirms our previous results with a more diverse and arguably more representative sample of participants in showing a substantial overlap between the nine considered dark traits, which can be attributed to the existence of a common dark core.

Testing for Acquiescence

Given that we relied on self-report questionnaires to assess the dark traits used in the measurement model for D and that these questionnaires are primarily (or exclusively) composed of positively-keyed items that describe socially undesirable characteristics (see

Table 1), individual differences in acquiescence as well as in the tendency to endorse socially undesirable items may contribute to the variance of the general factor imposed on the items. Therefore, the general factor arising in the bifactor model (which we interpret as D) may reflect a non-trivial amount of variance related to response styles.

The data obtained in this study offer a means to evaluate the extent to which this issue might affect the measurement of D by considering the average correlations of the general factor arising in the bifactor model to the positively-keyed versus the negatively-keyed items of the personality inventories (the NEO-FFI and the HEXACO-60). On average, the correlation of the general factor to the positively-keyed items of the NEO-FFI was r = -.02 compared to an average correlation to the negatively-keyed items of r = -.13. Concerning the HEXACO-60, the average correlation to the positively-keyed items was r = -.07 as compared to r = -.18 for the negatively-keyed items. Although the differences in correlations between positively and negatively keyed items appear to be rather modest, this result clearly suggests that the general factor in the bifactor specification indeed reflects some variance related to response styles in addition to substantive personality variance (because the general factor correlates more strongly with one group of items, namely, negatively keyed). Whereas the data at hand do not allow for an estimate of the proportion of variance that is due to response-styles, it must be kept in mind that this issue likely inflates associations to other variables that suffer from the same problem, i.e. the self-report criteria that are composed of positively-keyed items and measure undesired characteristics (Aggression, Dominance, Insensitivity, Power, and Self-Centeredness). Nevertheless, the general factor representing D is likely dominated by substantive personality variance as reflected in its empirical relations to variables that either measure a desirable characteristic (see below) or are entirely devoid of acquiescence, i.e., the behavioral outcomes considered next.

Behavioral Criterion Measures

Concerning the behavioral relevance of D, we first considered the cheating task as an indicator of unethical behavior. Of the participants, 37.4% responded "yes" to the question whether the coin had landed target side up in exactly two out of two tosses, which is highly comparable to previous findings with similar baseline probabilities (e.g., Hilbig & Zettler, 2015). Given that the baseline probability of winning (under complete honesty) is 25%, 17% (= (.374 - .25) / (1 - .25)) of the participants most likely cheated by incorrectly responding "yes" (Moshagen & Hilbig, 2017). In the following analyses, we treat the variable indicating participants' "yes" or "no"-response as a proxy for actual cheating behavior. Note that this variable underestimates the true population parameters (Moshagen & Hilbig, 2017), because it conflates both honest respondents (who actually got lucky) and dishonest respondents who cheated.[6] The latent variable models involving behavior in the cheating task were estimated using diagonally (robust) weighted least squares (B. Muthén, duToit, & Spisic, 1997; called wlsmv in Mplus), as this estimator has been shown to outperform other approaches when outcome variables are truly ordinal (Moshagen & Musch, 2014; Yang-Wallentin, Jöreskog, & Luo, 2010). Nested model tests were performed using the procedure outline in Asparouhov and Muthén (2006).

As shown in Table 8, all non-residualized dark traits were significantly positively associated with cheating on a bivariate basis, although the polychoric correlation estimates were small (ranging from r = .08 for Sadism to r = .17 for Egoism and Psychopathy). To evaluate whether the dark traits can predict cheating behavior over and beyond D, we

[6] Moshagen and Hilbig (2017) proposed adapted approaches to estimate correlations and regression coefficients accounting for the contamination of the dependent variable by honest participants who actually won. Unfortunately, however, this approach has yet to be extended to cover latent variables. Note that the approach selected is more conservative in that it will underestimate the effect of D on cheating.

regressed the response variable in the cheating task on D and the specific factors using a probit link function (i.e., probit regression with latent variables). As can be seen from the probit regression coefficients shown in Table 6, D was the only variable that significantly predicted cheating behavior, $b = 0.19$ ($p < .01$).

We next attempted to replicate the finding from Study 2 of a strong association between D and selfishness in terms of DG choice. On average, participants decided to keep 72.6% ($SD = 26.4$) of the available amount of 5 Euros for themselves which is again highly compatible with typical DG behavior (Engel, 2011). Largely in agreement with Study 2, all non-residualized dark traits (except for Narcissism) showed a significantly positive correlation to the amount of money kept (ranging from .04 for Narcissism to .17 for Machiavellianism; Table 8). However, the bivariate associations to DG selfishness were weaker than in Study 2, which might be related to the longitudinal design or to the monetary payoffs involved, both of which reduce spurious covariance.

To evaluate the relative predictive ability of D versus the specific dark traits, we regressed DG selfishness on D and the specific factors (all as defined in the bifactor model) (Table 6). Note again that the specific factors are mutually independent and orthogonal to D. As in Study 2, D was the strongest predictor of selfishness ($\beta = .17$, $p < .01$), whereas of the specific factors only Psychological Entitlement significantly predicted said criterion ($\beta = .09$, $p < .05$) in the theoretically expected direction. Strikingly, the specific factors for Moral Disengagement and Narcissism showed positive (rather than negative) relationships to DG choice, indicating that these specific factors seem to carry a theoretically reversed meaning (as compared to their original meaning) once the common core is controlled for.

Overall, these results are congruent with Study 2 in confirming that maximizing one's own utility at the expense of another's is a psychological core characteristic of D. Also as in Study 2, only Psychological Entitlement positively predicted selfishness beyond D, suggesting that Psychological Entitlement carries additional behaviorally relevant meaning.

Correspondingly, the prediction of selfish behavior in the DG provides results highly congruent with the prediction of unethical behavior in the cheating task, in that D was an important predictor for both measures. By contrast, once their common core is accounted for, none of the specific dark traits predicts behavior in the cheating task, and only one specific dark trait (Psychological Entitlement) predicts behavior in the DG in a theoretically plausible direction.

Self-report Criterion Measures

As additional tests of the validity of D versus the specific factors, we assessed nine self-reported criterion measures that were correlated most strongly with one of the considered dark traits in the articles that originally proposed one (or more) of the considered dark traits. In the following, we refer to this criterion as the "target criterion" of each respective trait.

Instead of fitting one very large model that includes all dark traits and all criteria simultaneously, we adopted a sequential modeling strategy. Specifically, we determined the correlations and performed the regressions by fitting one model for each of the nine criteria (with the exception of the Dominance and Nurturance dimensions of the interpersonal circumplex, which were estimated and related to the relevant traits in a single model). Thus, for each criterion, we computed two models that included (1) the criterion and the nine non-residualized dark traits (but not D) versus (2) the criterion, the specific factors, *and* D. The criteria were always modeled as a single latent variable. The interpersonal circumplex contains eight lower-order scales comprising four items each (so-called octants) that are hypothesized to exist in a circumplex structure with the Dominance and Nurturance dimensions defining the primary axes of the circumplex. Correspondingly, the lower-order scales were modeled as first-order latent factors. Based on these first-order factors, the Dominance and Nurturance dimensions were obtained as orthogonal second-order factors by constraining the loadings of the first-order factors on the two second-order factors to a circumplex structure (for details, see Gaines, Panter, Lyde, Steers, & Al, 1997; Tracey, 2000).

All confirmatory factor models (without bifactor structure) fitted the data satisfactorily, with the *RMSEA*s ranging from .034 to .044 and the *SRMR*s from .057 to .065. Likewise, all bifactor models also exhibited a good fit to the data, *RMSEAs* = .038 - .044; *SRMR*s = .059 - .067.

On a bivariate basis, each non-residualized dark trait was significantly correlated with its target criterion in the expected direction (Table 8). In most cases, the correlations found here were highly similar in magnitude to the ones reported in the original studies (see Table 7), with two exceptions: The correlation between Moral Disengagement and Internalized Moral Identity (r = -.24) was substantially smaller than in the study by Moore et al. (2012; r = -.42 and -.55). Similarly, Psychological Entitlement was only weakly associated with Perspective Taking (r = -.12 as opposed to r = -.43 in Campbell et al., 2004). However, both Moral Disengagement and Psychological Entitlement exhibited moderate to strong associations to the remaining criteria in a psychologically meaningful way, so it seems safe to conclude that both questionnaires provide appropriate measures of the underlying traits.

To critically test whether the dark traits continue to predict their target criterion once the common dark core is controlled for, we regressed each criterion on D and all specific factors. Again, note that D and the specific factors are mutually uncorrelated, allowing for a straightforward interpretation of the regression coefficients. The regression results are presented in Table 9. D significantly predicted all criteria, including those indicating positively connoted attributes (Internalized Moral Identity, Nurturance, Perspective-Taking), and was the strongest predictor for all but one criterion (Dominance). Conversely, most of the specific dark traits (with the exceptions of Psychopathy and Sadism) exhibited significant associations with only one or two of the criteria in a psychologically meaningful way. Indeed, the majority of the specific dark traits (Egoism, Machiavellianism, Moral Disengagement, Psychological Entitlement, and Spitefulness) did not predict their respective target criterion in the theoretical expected direction over and beyond D. Of the remaining specific dark traits,

the incremental contribution of Psychopathy (β = .16) and Sadism (β = .09) in the prediction of Impulsivity and Insensitivity, respectively, was rather small. Only the specific factors of Narcissism (β = .66) and Self-Interest (β = .38) substantially increased the explained proportion of the variance in Dominance and Power, respectively, over the variance accounted for by D (β = .26 and β = .47).

The relations between D and the various criteria also avails further evaluation of our purported definition of D. Reflecting the tendency to maximize one's own utility, D was strongly related to Self-Centeredness, Power, (lack of) Internalized Moral Identity, and Dominance. Further, in agreement with the characteristic of D to disregard or provoke disutility for others, D exhibited strong associations to Aggression, (lack of) Nurturance, (lack of) Perspective Taking, and (lack of) Internalized Moral Identity. Finally—although not explicitly represented by any of the considered criteria—the third characteristic feature of D that individuals hold beliefs that serve as justifications is supported by the positive relation to Dominance in tandem with the negative relation to Nurturance, which is often considered the octant of the interpersonal circumplex reflecting narcissistic personality disorders and thus reflects the belief of one's own superiority (Miller, Price, Gentile, Lynam, & Campbell, 2012; Ruiz, Smith, & Rhodewalt, 2001).

As explained in detail above, selection of a respective dark trait's target criterion was guided by the rationale that the construct validation of a newly proposed dark trait will have relied on criterion variables that are well-aligned with the considered construct. By selecting the one criterion that correlated most strongly in the article introducing a dark trait, we attempted to define an appropriate target criterion for each dark trait in a non-arbitrary manner. However, clearly, this is not meant to imply that the remaining criterion variables considered here would be irrelevant for the other dark traits. Thus, we now consider the full set of correlations to the criterion measures for each specific dark trait to evaluate the meaning of the specific dark traits more comprehensively.

Of the specific dark traits, Psychopathy and Sadism showed various theoretically plausible links to the criteria even after controlling for D. The residualized factor for Psychopathy was significantly negatively related to Internalized Moral Identity and significantly positively related to Self-Centeredness, Dominance, Impulsivity, Insensitivity, and Aggression. Similarly, the specific factor for Sadism exhibited a significantly negative association with Nurturance and significantly positive associations with Self- Centeredness, Insensitivity, and Aggression. These results suggest that both traits, Psychopathy and Sadism, comprise aspects consistent with their respective theoretical conceptualization that are not entirely reflected in D (in particular with respect to a stronger propensity to behave aggressively and a lower sensitivity towards others). This is particularly noteworthy concerning Psychopathy in light of the rather strong D saturation of Psychopathy as indicated by the corresponding *ECV* (Table 4). In tandem with fact that D predicted various criteria (including the behavioral criteria) that were not predicted by the residualized factor for Psychopathy, this indicates that neither Psychopathy can be reduced to D nor that, vice versa, Psychopathy is simply a substitute of D.

In contrast to Psychopathy and Sadism, the remaining specific dark traits showed a less clear pattern, often exhibiting theoretically implausible or even downright theory-inconsistent relations with the criteria. The specific factor for Moral Disengagement, for example, displayed theoretically implausible links by a *positive* relation to Nurturance and a *negative* relation to Power. Mixed results also occurred with respect to the other dark traits. For example, the specific factor for Egoism, while positively related to Insensitivity, *positively* predicted Perspective Taking and *negatively* predicted Dominance and Power. These regression coefficients differ qualitatively from the respective bivariate associations of the non-residualized Egoism trait, which suggests that the residualized factor for Egoism (controlled for D) carries a very different meaning than the non-residualized Egoism factor. Similarly, the residualized factor for Machiavellianism, while associated with higher Power,

showed *positive* regression coefficients in the prediction of Nurturance and Internalized Moral Identity, and *negative* coefficients in the prediction of Impulsivity and Aggression, even though the bivariate associations involving the non-residualized Machiavellianism factor consistently suggested the opposite. Likewise, the residualized factor for Self-Interest, while related to higher Power, was *positively* related to Nurturance, Internalized Moral Identity, and Perspective Taking. As these cases illustrate, remaining similarities between the items of some dark traits that are not accounted for by D may require a specific factor; however, these specific factors can strongly differ from the original theoretical conceptualization of the dark trait in question. The residualized Egoism factor (i.e., the specific factor after controlling for D), for example, primarily seems to reflect high levels of Impulsivity and Insensitivity (which are only partly comprised in D), rather than a lack of Perspective Taking or increased Self-Centeredness. In any case, the key conclusion concerning the role of D is straightforward: D showed a highly meaningful and generally strong pattern of associations to a large number of criteria relevant to the domain of dark traits (indeed, all of the ones investigated herein).

In sum, these results can be interpreted as support for the substantive nature of D in terms of reflecting the tendency to maximize one's own utility at the cost of others, accompanied by beliefs that serve as justification. Overall, the regression results thus indicate that (1) D indeed captures psychologically relevant meaning in agreement with our conceptualization and cannot be reasonably explained as reflecting a mere measurement artifact and (2) most of the dark traits considered do not (or only marginally) contribute to the prediction of relevant criteria, once the contribution of their common dark core is accounted for.

Locating D in Basic Personality Space

The final purpose of Study 3 was to locate D in the personality space as spanned by the Five-Factor-Model (FFM) and the HEXACO Model of Personality, respectively. For these analyses, we used all respondents who successfully completed the second measurement

occasion as a data base (N = 1,261). The FFM and the HEXACO personality dimensions were modeled as (first-order) latent variables. Specifically, each personality dimension was measured through its associated items in the corresponding inventories, while assuming a perfect simple structure. For example, FFM Openness and Conscientiousness were modeled as single, correlated, latent variables as measured by the 12 Openness and the 12 Conscientiousness items of the NEO-FFI, respectively. All secondary loadings were fixed to zero, as were all correlations between the item residuals. The full model thus comprised the bifactor structure for the items measuring the dark traits as described in detail above (comprising D along with the specific factors for each of the dark traits), five factors for the FFM dimensions, and another six factors for the HEXACO dimensions. All latent factors were allowed to correlate, except that D and all specific dark trait factors were constrained to be orthogonal (as dictated by the bifactor structure and described in detail above).

As shown in Table 10, in both models of personality structure, D was characterized by low levels of Conscientiousness (r = -.27 and -.32 in the FFM and the HEXACO, respectively) and low levels of Agreeableness (r = -.69 and -.45). The difference between the latter coefficients is plausible given that, unlike FFM Agreeableness, HEXACO Agreeableness primarily covers reactive aspects such as tolerance, lenience, and forgiveness whereas active aspects such as fairness and honesty (that are also part of FFM Agreeableness) are assigned to the Honesty-Humility factor in the HEXACO model (Ashton & Lee, 2005). In line with this reasoning, D additionally exhibited a strong negative relation to Honesty-Humility (r = -.80). Moreover, D was weakly (but significantly) associated with low Openness (r = -.11 and -.05 in the FFM and the HEXACO, respectively) and low Extraversion (r = -.15 and -.17). Qualitatively different associations occurred for Neuroticism in the FFM (r = .20) and Emotionality in the HEXACO model (r = -.09). This difference might be explained by the fact that FFM Neuroticism contains aspects of angry hostility (as well as impulsiveness and irritability), whereas HEXACO Emotionality does not.

In light of the high correlation between D and Honesty-Humility and bearing in mind that Honesty-Humility in and of itself is a theoretical candidate for what may constitute the common core of dark traits (Lee & Ashton, 2014), we further investigated whether D merely reflects the negative pole of Honesty-Humility or whether D comprises psychologically meaningful variance beyond Honesty-Humility. To evaluate the ability of D to account for incremental variance in relevant criteria, we performed latent regression analyses involving D and a latent factor for Honesty-Humility (measured by the corresponding items of the HEXACO-60). Specifically, we regressed each criterion administered in this study—i.e., the nine self-report criteria and the two behavioral tasks—on Honesty-Humility alone (Model 1) as well as on Honesty-Humility and D (Model 2) to evaluate whether the latter model including D incrementally explains variance in the criterion over the first model involving Honesty-Humility alone. In interpreting these results, it should be kept in mind that the measurement model for D—unlike the one for Honesty-Humility—predominantly comprises positively keyed items assessing socially undesirable characteristics, thus giving rise to the concern that D might also reflect variance related to response styles in addition to substantive trait variance. As a consequence, the associations between D and other self-report measures assessing socially undesirable attributes with unbalanced scales might be inflated.

As shown in Table 11, D improved the prediction of the majority of criteria ($.03 \leq \Delta R^2 \leq .18$), with the exceptions of Dominance, Power, DG choice, and cheating.[7] Note that the latter results are not surprising, given the fact that D, while encompassing both (dis-)honesty

[7] Note that the estimation of the model that included both D and Honesty-Humility led to slightly different correlation estimates concerning the criterion measures compared to a model that considered either D or Honesty-Humility in isolation. Thus, some of the correlation estimates between D and the criterion variables reported in Table 11 slightly differ from those reported in in Table 9.

and (un-)fairness, cannot be considered as a pure measure of honesty or fairness (unlike Honesty-Humility). By implication, as soon as individual differences on these variables are accounted for by Honesty-Humility, little is left to explain by D on criteria that specifically measure modesty (Dominance, Power), selfishness (DG choice), and honesty (cheating). In contrast to this, D improves upon Honesty-Humility with regard to other negatively connoted criteria (Insensitivity, ΔR^2 = .18; Aggression, ΔR^2 = .13; Self-Centeredness, ΔR^2 = .11), but also concerning positively connoted criteria (Nurturance, ΔR^2 = .13; Internalized Moral Identity, ΔR^2 = .06; Perspective Taking, ΔR^2 = .03). Note that variance related to response styles might have inflated the association between D and the negatively connoted criteria, however, this issue would actually suppress the association to positively connoted criteria. In tandem, these results indicate that, although Honesty-Humility clearly provides a good proxy, D reflects a specific mélange of personality characteristics not identical with one single basic personality trait, which is also in line with the findings reported above that D is comprises variance related to Agreeableness and Conscientiousness.

Relatedly, we also investigated whether D incrementally predicts the criterion measures once controlling for all FFM and HEXACO personality dimensions to investigate whether D comprises variances that lies outside contemporary models of personality structure. Again, we performed latent regression analyses predicting the criteria either by the five FFM (or six HEXACO) dimensions (Model 1) or by said basic personality dimensions and D (Model 2) to evaluate whether D may account for additional variance. Results indicated that D significantly improved the prediction of dishonest behavior in the cheating task as well as Self-Centeredness, Dominance, Impulsivity, Insensitivity, Power, and Aggression ($.01 \leq \Delta R^2 \leq .10$) over all five FFM dimensions. Beyond the HEXACO dimensions, D incrementally predicted Self-Centeredness, Internalized Moral Identity, Impulsivity, Insensitivity, and Aggression ($.02 \leq \Delta R^2 \leq .10$). In tandem, these results indicate that D is not entirely reflected

in either the FF or the HEXACO model, but instead comprises substantive personality variance that lies outside the space spanned by these models.

Overall, this study further substantiated and extended the findings of the preceding studies, using a more representative and larger sample, longitudinally investigating links to a diverse set of self-report criterion measures, and considering different measures of actual behavior. It was again shown that the considered dark traits share a common core in line with our conceptualization of D, which captures most of the behaviorally meaningful variance across the dark traits, as evident in the prediction results concerning (incentivized) behavioral and self-reported criteria in the realm of ethically, morally, and socially questionable behavior. Unlike D, the specific dark traits only sporadically exhibited substantial and theoretically plausible relations with the criterion measures, even though the latter were selected based on maximal criterion validity of the dark traits. Finally, locating D in the personality space showed that D is associated with (low) Agreeableness and (low) Conscientiousness in both personality models considered, and (low) Honesty-Humility in the HEXACO model. Although D thus clearly overlaps with basic personality traits and especially Honesty-Humility, prediction of the diverse criteria did show that D comprises aspects beyond Honesty-Humility and indeed both complete models of personality structure.

A remaining open question pertains to the specific operational definition used for D thus far. Our theoretical conceptualization of D does not depend on particular dark traits (let alone on particular items), but is instead deemed to reflect a general behavioral tendency in a fluid sense. In theoretical terms, any dark trait is a specific manifestation of D, so that D represents the underlying driving force for a large number of dark traits, including, but not limited to the ones considered. By implication, the meaning of D must not depend on any particular dark trait (let alone item) included in its measurement model. Using a simulation approach, the fourth study was directed towards critically testing the hypothesis that D is a

meaningful (and predictively valid) factor irrespective of specific dark traits (i.e., dark trait measures) included as indicators of D.

Study 4

The purpose of the final study was to investigate the assumption that D can be understood as a fluid construct, much like the *g* factor of intelligence in the ability domain. Specifically, we proposed that D is not exclusively defined by any particular indicator of a single or a set of dark traits, but rather a general underlying tendency responsible for the commonalities among various dark traits, consequently representing their common core. If this holds, the meaning of D must be largely independent of any one indicator variable, that is, the specific set of dark traits included or items employed.

Methods

We tested the notion of D as a fluid concept by resorting to a simulation (or resampling) approach. More specifically, in each simulation, we omitted a certain proportion or a particular set of items from the measurement model of D. We then evaluated the degree to which this affects the meaning and predictive validity of D by considering the correlation of the reduced version of D (as measured through a subset of the available indicators) with the complete version of D (as measured by all available indicators). As a data base, we pooled the samples of Studies 1–3, thereby obtaining information on a total of 93 items to measure nine different dark traits (as described in detail in Study 1) in one large sample comprising N = 2,659 observations.

Because it is not possible to formulate a structural equation model that jointly estimates a full and a reduced version of an otherwise identical latent variable, we relied on estimated factor scores to compute correlations between the full and a reduced version of D instead. The use of factor scores is subject to a drawback, however. If the latent variable is not well defined through its own indicators, different approaches to obtain factor scores may lead

to highly diverging results (Rodriguez, Reise, & Haviland, 2016). The validity of the approach of using factor scores as an approximation of latent variables thus critically hinges on the adequacy of the factor model. Fortunately, indices of factor determinacy indicated that D is very well defined: the index of construct replicability (Hancock & Mueller, 2001) was *H* = .96 and the index of factor determinacy (the correlation of factor scores with the factor; Beauducel, 2011; Rodriguez et al., 2016) was *FD* = .96, thereby scoring well above the recommended values of .70 (Hancock & Mueller, 2001) and .90 (Gorsuch, 1983), respectively. Thus, the estimated factor scores are sufficiently determinate to conclude that "individual differences on the factor score estimates are good representations of true individual differences on the factor" (Rodriguez et al., 2016, p. 142).

Given that relying on factor scores is a reasonable and reliable approach in the present context, we first estimated the factor scores for the complete version of D (including all available items) using the bifactor specification outlined in detail above. Next, we obtained the factor scores when measuring D through a subset of the available items, again using a bifactor specification analogous to the full model. Finally, we correlated the thereby obtained factor scores with the one obtained from the fully specified D.

We considered two types of reduced item sets. First, we defined measurement models for D that omitted the items of one of the dark traits, thereby obtaining nine factor score estimates for D based on reduced models that omitted the indicators of one particular dark trait. In a similar vein, we also defined a measurement model for D that omitted the items of all Dark Triad traits (i.e., Machiavellianism, Narcissism, and Psychopathy) at once. The second type of reduced item sets was determined by randomly omitting either 50% or 75% percent of the available items, so that D was measured through a random subset of 46 or 23 items, respectively, again using a bifactor specification (whenever at least three items of a specific dark trait were retained; when only two items were included, we freely estimated the correlation between the residuals of these items). This random selection of the items was

repeated 1,000 times each. The second approach thus led to 2,000 factor score estimates, each based on a different set of indicator items for D. If D can be seen as a fluid construct, the particular set of items used to measure D should not matter to a great extent, so that all estimated factor scores should be strongly correlated.

Beyond determining the correlation of these reduced versions of D with each other and the complete version of D, we also replicated the regression results in Study 3 (based on the sample of $N = 882$ respondents who completed Study 3). Specifically, we predicted the target criteria (as defined above) by reduced versions of D that omitted the items of the one dark trait for which the criterion was deemed of core relevance. For instance, we predicted Self-Centeredness as a target criterion for Egoism by the non-residualized Egoism factor and the reduced D omitting all items of the Egoism scale. If D can be regarded as a general underlying disposition that represents the core of dark traits independently of their specific indicators, D must continue to predict said criteria even after omission of the respective dark traits from its measurement model.

Results and Discussion

We first consider the results obtained when omitting the items of one of the dark traits from the measurement model for D. The correlations among the thereby obtained factor scores, as well as the correlations to the factor scores obtained with the full model including all items, are shown in Table 12. Generally, the correlations to the factor scores obtained with the complete model were extremely high (all $r \geq .95$), as were the correlations among differently reduced models (all $r \geq .94$). Even excluding all items pertaining to the Dark Triad traits from the measurement model did not change the meaning of D to a substantial degree, as evidenced by a correlation to the factor scores obtained with the complete measurement model of $r = .95$. These results indicate that the items of any one particular dark trait or the Dark Triad only play a minor role in the measurement of D. Stated differently, D has virtually the same meaning, regardless of whether the items of one particular dark trait are omitted.

This supports the notion of D as a fluid concept and thus the behavioral tendency underlying dark traits in general.

In addition, we also replicated the regressions in Study 3 by predicting the criteria by reduced versions of D. In particular, we were most interested in the ability of D to account for variance in a target criterion when the items of the respective dark trait (for which the criterion had originally served as strong validation) were omitted in the measurement model of D. To this end, we estimated structural equation models that specified a bifactor structure for D (as measured through the items of eight dark traits; otherwise specified the same way as in the preceding studies), eight residualized factors for the specific dark traits, one non-residualized factor for the dark trait omitted from the measurement model of D, and another non-residualized factor for the target criterion itself. For example, Self-Centeredness (modeled as non-residualized factor) was predicted by the non-residualized Egoism factor (as measured through its associated items) and the reduced D factor in a bifactor specification (as measured by all dark trait items except for the items of the Egoism scale). In line with the bifactor specification, this model also comprised eight specific (residualized) factors for the remaining dark traits (i.e., the specific factors for Machiavellianism, Moral Disengagement, Narcissism, Psychological Entitlement, Psychopathy, Sadism, Self-Interest, and Spitefulness). However, the specific factors, while required for a proper specification of the bifactor model, were not included in the regression equation as predictor variables. Because the reduced version of D and the remaining non-residualized dark trait were correlated, we base our considerations on the bivariate associations with the target criterion. Furthermore, we evaluated the change in the R^2 when predicting the criterion by D alone (Model 1) or by D and the respective non-residualized dark trait (Model 2) to test whether the respective non-residualized dark trait can account for variance in its target criterion over the reduced version of D (that does not include the very items of this dark trait). The results are presented in Table 13.

As expected (and congruent with Study 3), the reduced versions of D were significantly and substantially correlated with each criterion in the expected direction ($.24 < |r| < .63$). Of the non-residualized dark traits, Narcissism and Self-interest (and, to a lesser extent, Psychopathy and Sadism) incrementally accounted for a substantial proportion of variance in their respective target criteria over the reduced versions of D, indicating that these traits capture unique variance beyond D and, thus, that neither of these traits can be used as a substitute for D. However, five out of the nine dark traits did not (or only very weakly; $\Delta R^2 \leq .01$) improve the prediction of "their own" target criteria over the reduced version of D. These results are closely aligned with those of Study 3, but offer the vital extension that the items of the very target construct are not even required in the measurement model of D concerning a wide array of target criteria. Even under these conditions, D accounted for most of the explained variance in the majority of criteria, thereby again indicating that D captures the behaviorally relevant dark core of many different dark traits without crucially depending on any one of these as an indicator.

These results give rise to the question how many items can be omitted from the measurement model for D without changing its meaning to a substantial degree. To this end, we randomly omitted 50% or 75% percent of the available items from the measurement model of D and repeated this process 1,000 times each. The obtained distributions of correlations between the randomly reduced versions of D to the complete version of D are summarized in Table 14. It is evident that omitting 50% of items at random still leads to factor scores that are strongly correlated with the full version of D. For example, the 5^{th} percentile of the distributions of correlations was .96, showing that 95% of all reduced versions comprising only a random half of the items correlated at least to .96 with the complete version. Similarly, even randomly omitting 75% of items had a negligible impact on the meaning of D in the vast majority of situations, given a median correlation of .94 in this case. Finally, we also evaluated the distributions of correlations between the factor scores

among the reduced versions of D. With 50% of items omitted, the resulting factor scores were still correlated with each other to at least .93 in almost all situations. Somewhat lower (but still high) intercorrelations were obtained when omitting 75% of items, with a median of .89 even in this scenario.

In tandem, the results of this study show that a very similar D emerges from highly diverse item sets and thereby provide strong support for both (1) the idea that D is a fluid concept that does not depend on a particular set of items, nor a large number of items, nor particular indicator constructs, and (2) the more general theoretical position that D can be understood as the basic tendency underlying dark traits in general.

General Discussion

Ethically, morally, and socially questionable behavior affects the everyday lives of literally billions of people and has thus spurred much interest among researchers. Among all the different approaches to describe, understand, and predict such behavior, personality psychologists have considered stable individual differences as one source of explanation. Dating back to research in the early 1950s, many different personality traits related to ethically, morally, and socially questionable behavior have been proposed, including, for example, Machiavellianism (Christie & Geis, 1970) and Psychoticism (e.g., Eysenck & Eysenck, 1976), or, more recently, Sadism (O'Meara et al., 2011) and Spitefulness (Marcus et al., 2014). As a result, a notable body of knowledge has been accumulated concerning the similarities and dissimilarities between different dark traits (e.g., Jonason et al., 2009; Jones & Figueredo, 2013; Lee & Ashton, 2014; Paulhus, 2014), the psychometric characteristics of widely-used questionnaires aiming to assess dark traits (e.g., Paulhus & Jones, 2015), and, ultimately, the links between dark traits and diverse outcomes (e.g., Blais et al., 2014; Buckels et al., 2013). In particular, there is consistent evidence showing that various dark traits are

empirically associated to a notable extent (e.g., Book et al., 2016; Muris et al., 2017; O'Boyle et al., 2012), in turn suggesting that dark traits share common characteristics.

Based on this substantial body of knowledge, we have herein provided a theoretical conceptualization of the basic and general dispositional tendency underlying dark traits. This common dark core, which we call the *Dark Factor of Personality* (*D*), is understood as a fluid construct, independent of any particular dark trait. Correspondingly, we showed that different dark traits can be understood as specific manifestations of this basic and broad dispositional tendency (Studies 1-3), that D predicts a multitude of criteria in the realm of ethically, morally, and socially questionable behavior (Studies 2-3), that D continues to predict these criteria even after removing relevant indicator variables (Study 4), how D relates to basic traits as defined in models of personality structure (Study 3), and that D generally does not depend on any particular indicator variable included (Study 4).

The Common Core of Dark Traits: The Dark Factor of Personality

In a series of studies, we demonstrated that the commonalities among several dark traits can be attributed to a single core, the *Dark Factor of Personality* (*D*), which reflects the basic dispositional tendency to maximize one's individual utility—disregarding, accepting, or malevolently provoking disutility for others—, accompanied by beliefs that serve as justifications. By implication, D represents the prime source of variance for individual differences on dark traits, so that dark traits constitute specific instances of this more general underlying tendency. In line with this reasoning, in three studies, *D* emerged as a single factor underlying a diverse set of dark traits (including both more traditional dark traits and those that have been introduced recently), namely, Egoism, Machiavellianism, Moral Disengagement, Narcissism, Psychological Entitlement, Psychopathy, Sadism, Self-Interest, and Spitefulness.

Beyond providing support for the notion that D captures the commonalities among a diverse set of traits, the degree to which the dark traits are subsumed in D closely mirrors our

definition of D. That is, we defined D in terms of broad, interrelated characteristics—utility maximization along with disutility infliction, and justifying beliefs—and indeed found a particularly high D saturation for dark traits that place a strong emphasis on several of these aspects (e.g., Machiavellianism, Psychopathy, Spitefulness). By comparison, dark traits primarily focusing on one of these aspects (e.g., Narcissism, Psychological Entitlement, Self-Interest) showed a lower D saturation.

The theoretical characterization of D was also supported by the fact that items directly addressing the broad characteristics defining D exhibited particularly high loadings on D, for example “I’ll say anything to get what I want” (referring to utility maximization), “There have been times when I was willing to suffer some small harm so that I could punish someone else who deserved it” (referring to inflicting disutility for others), or “I honestly feel I’m just more deserving than others” (referring to justifying beliefs). In contrast, items that did not bear a clear relation to any of the characteristics defining D were virtually unrelated to D, e.g. “Hearing others praise me is something I look forward to.” or “Nowadays a person has to live pretty much for today and let tomorrow take care of itself”. In sum, the pattern of item loadings is in close agreement with the theoretical conceptualization of D. In addition, the observation that items referring to such diverse beliefs and behaviors exhibit a similar association with D can be interpreted as immediate support for the proposition that D indeed captures general individual differences on dark traits.

Further support for the proposal that D acts as the prime source of variance responsible for individual differences in dark traits and captures behaviorally relevant meaning was provided by testing for the prediction of various criteria in the realm of ethically, morally, and socially questionable behavior. Indeed, D accounted for a substantial proportion of the variance in selfish behavior (as measured through the dictator game), in unethical behavior (as measured through a cheating task), and strongly predicted nine different self-report criteria (namely, Aggression, Dominance, Impulsivity, Internalized Moral Identity, Insensitivity,

Nurturance, Perspective Taking, Power, and Self-Centeredness) all of which had been selected due to their immediate relevance for one of the specific dark traits considered. By contrast, the specific dark traits only sporadically improved the prediction of these criteria beyond D in a psychologically meaningful way. Overall, D thus functioned as a strong underlying factor subsuming diverse dark personality characteristics and their consequences in terms of ethically, morally, and socially questionable behavior.

D in Relation to Basic Personality Traits

We also located D in the personality space as spanned by both the FFM (McCrae & Costa, 2008) and the HEXACO model (Ashton & Lee, 2007) to substantiate the theoretical meaning of D and to evaluate how D is reflected in the space spanned by contemporary models of basic personality traits. With regard to the FFM, and in line with previous studies linking dark traits to the Big Five (Muris et al., 2017; O'Boyle et al., 2012), D was primarily related to low Agreeableness and low Conscientiousness, indicating that individuals high in D are characterized by a lack of compliance, kindness, and modesty as well as higher impulsivity, lack of law-abiding, rule-following, and self-control, especially when dealing with others. In addition, D also exhibited weak to moderate links to (low) Extraversion and Neuroticism, probably reflecting a less sociable and more anxious, irritable, or moody manner. Also, regression results revealed that D incrementally predicted 7 out of the 11 considered criterion measures over the FFM dimensions, including dishonest behavior in the cheating task. This suggests that D comprises behaviorally meaningful variance that lies outside the personality space spanned by the FFM dimensions.

Highly similar results were obtained with respect to HEXACO Agreeableness (subject to the conceptual differences between Big Five and HEXACO Agreeableness, Ashton, Lee, & de Vries, 2014) and Conscientiousness. Different from FFM Neuroticism, D exhibited a negative —albeit weak—correlation to Emotionality, which is probably due to the fact that Emotionality does not comprise aspects of angry hostility or irritability (unlike FFM-

Neuroticism). Most strikingly, and mirroring previous notions that Honesty-Humility in particular may be the prime basic personality dimension responsible for commonalities between the Dark Triad components (Book et al., 2015; Lee & Ashton, 2014), D exhibited a strong negative correlation to Honesty-Humility indicating that key characteristics of Honesty-Humility are mirrored in D and vice versa.

In the HEXACO model, Honesty-Humility subsumes the aspects of fairness, sincerity, greed avoidance, and modesty (e.g., Ashton et al., 2014; Lee & Ashton, 2006; Zettler & Hilbig, 2015). Clearly, this conceptualization is closely related to the main characteristics defining D, first and foremost utility maximization (greed and lack of sincerity in Honesty-Humility terms) and justifying beliefs (lack of modesty in Honesty-Humility terms). Nevertheless, the results also show that D is more than the negative pole of Honesty-Humility, given that D substantially improved the prediction of the majority of the considered criteria over Honesty-Humility (and all six HEXACO factors, respectively), in particular with respect to Aggression, Insensitivity, (lack of) Nurturance, and Self-Centeredness. Moreover, D is additionally defined by inflicting disutility on others (even to the extent that one may accept costs so as to inflict even greater costs on other) which is not mirrored well in Honesty-Humility. In the latter, disutility for others is typically a by-product of utility maximization, but not considered a source of utility of itself.

Thus, on a conceptual level, D differs from Honesty-Humility in at least three respects. First, justifying beliefs of individuals high in D include, but are not limited to a sense of superiority or entitlement. For instance, individuals may also adopt a cynical world-view, endorse ideologies favoring dominance, or may think that everyone looks for himself first as justifying beliefs. Second, individuals high in D are not only motivated by monetary payoffs or achieving a higher social status, but may also strive for less tangible types of utility such as feelings of joy, pleasure, or self-enhancement in general. Finally, our definition of D recognizes that individuals may derive utility from the very act of inflicting disutility on

others. The latter implies that individuals high in D may engage in behaviors that reduce their own utility in terms of money or status, but increase their utility in terms of pleasure or joy as a result of other's disutility. For example, high D individuals may be willing to pay money in order to see others suffer, which is difficult to reconcile within the conceptualization of Honesty-Humility. Likewise, acts of revenge, which are attributed to Agreeableness rather than Honesty-Humility in the HEXACO model (Ashton et al., 2014; Hilbig, Thielmann, Klein, & Henninger, 2016), can also be understood as deriving utility from inflicting disutility on others.

In sum, Honesty-Humility clearly bears notable resemblance to D and is the basic personality dimension that most closely approximates D. However, D also comprises a different array of behaviors and thus cannot be thought of merely reflecting the negative pole of Honesty-Humility. Finally, it should be noted that D is not well suited for inclusion in a more general model of personality dimensions akin to the FFM or HEXACO model. Since D represents a blend of basic traits, achieving orthogonality to other basic personality dimensions (which is generally desired) would require rotating those dimensions to positions in the personality space that are difficult to interpret. Nevertheless, to more strictly dissociate D from Honesty-Humility, future studies should consider additional (behavioral) criteria with a particular emphasis on the aspect of actively provoking disutility for others (e.g., Buckels et al., 2013), possibly even in the face of individual cost.

The Fluid Nature of D and Implications for its Measurement

D is conceptualized as a general behavioral disposition constituting the major source of variance underlying individual differences in dark traits. Thus, any dark trait also reflects D as a specific manifestation thereof, albeit to a varying degree. D is therefore not defined by a single or a particular set of dark traits, but is proposed to represent a fluid concept that does not depend on particular indicator variables. We tested this notion in a series of simulation studies by excluding specific dark traits, all components of the Dark Triad, or random (even

large) sets of items across all dark trait measures from the measurement model for D. All reduced versions of D were highly correlated with the complete version of D, and the predictive power of D for the diverse outcomes was, with few exceptions, still notably high when reduced variants of D were considered, thereby providing strong support in favor of the proposal that D is to be considered as a fluid concept. In fact, D continued to predict criteria that were selected for their high association with a particular dark trait even after excluding all items of said dark trait from the measurement model for D.

The fluid character of D yields some important implications. Unlike the theoretical characterization of D on the construct level, the operational definition of D obviously depends on the indicators of the dark traits included, so that (slight) shifts in meaning are to be expected when different sets of dark traits are investigated. Much like the *g* factor in the ability domain, any specific operational definition of D is thus "flavored" (Carroll, 1993) by the indicators of the dark traits included in its measurement portion. D is therefore best conceptualized as a fluid concept that captures the common basis and represents the prime source of variance of a large number of different dark traits. Correspondingly, although the indicators of any particular dark trait will—to a certain extent—also reflect D, measuring D itself requires the inclusion of a sufficiently large number of indicators of diverse dark traits in order to capture the full theoretical breadth that D represents. Whenever an operational definition of D only comprises indicators of dark traits that, taken together, do not cover its primary characteristics or are otherwise limited in their variety, the resulting measurement of D may be dominated by specific variance underlying the included traits, and, consequently, constitute a poor representation of D as a whole. To measure D with sufficient breadth, only a sufficiently large number of indicators of diverse dark traits will ensure that D emerges as the factor responsible for the commonalities across the dark traits.

The Unique Role of Dark Traits Beyond D

Whereas our studies provided strong support for the notion that the considered dark traits are specific manifestations of a common core as reflected in D, substantial differences emerged regarding the degree to which the investigated dark traits were subsumed in D. In particular, whereas Machiavellianism, Psychopathy, and Spitefulness were largely absorbed by D, Narcissism and Self-Interest exhibited a comparatively large proportion of unique variance beyond D, to name the extremes. In addition, none of the considered dark traits incrementally predicted unethical behavior in the cheating task and only Psychological Entitlement predicted selfish behavior in the Dictator Game over and beyond D. A similar pattern of results occurred in the prediction of the self-reported criteria revealing that most of the residualized dark traits do not carry dark characteristics beyond D, with the exceptions of Sadism and Psychopathy. Concerning the latter, results involving the self-reported criteria suggest that individuals high on Psychopathy are more impulsive, more aggressive, and less sensitive towards others than would be expected based on their level in D, thereby offering an initial characterization of the unique features Psychopathy may have beyond D.

At first sight, the comparatively low D saturation for Narcissism may seem counterintuitive, given that Narcissism has repeatedly been shown to correlate substantially with the remaining components of the Dark Triad (e.g., Muris et al., 2017; O'Boyle et al., 2012). One explanation for this might be that in our studies Narcissism was assessed by the respective subscale of the Short Dark Triad (Jones & Paulhus, 2014), which predominantly captures tendencies that have been termed "narcissistic admiration" by Back et al. (2013), i.e., striving for uniqueness, having grandiose fantasies, and behaving charmingly. In contrast, the narcissistic rivalry subtype—i.e., striving for supremacy, devaluating others, and behaving aggressively—is poorly represented in the respective scale of the Short Dark Triad (as is immediately evident from inspecting the item content). However, narcissistic rivalry is arguably more closely aligned with our conceptualization of D, since it directly corresponds to the characteristics of D in terms of maximizing one's own utility at the expense of others

(aggressively striving for supremacy) accompanied by justifying beliefs (devaluation of others). Hence, the comparatively low D saturation for Narcissism found in the present studies might be a consequence of the specific operationalization employed, rather than being due to low commonalities between D and Narcissism (on a construct level) as such.

D also explained a relatively modest amount of the variance in the items measuring Self-Interest, suggesting that many features associated with the latter trait may lie outside the domain spanned by D. Self-Interest is defined as a rather mild form of individual utility maximization in terms of pursuing gains in "socially valued domains, including material goods, social status, recognition, academic or occupational achievement, and happiness" (Gerbasi & Prentice, 2013, p. 496). According to this definition, aiming for material goods and a higher status are the only aspects that immediately imply potential disutility for others (which is a defining feature of D). It is quite possible, for example, to strive for occupational achievement or happiness without inflicting negative consequences for others. Indeed, Gerbasi and Prentice (2013) even reported *positive* correlations between their Self-Interest construct and Other-Interest ("the motivation to act in another's interest", p. 495), between .20 and .35. Thus, quite contrary to D, high Self-Interest as defined by Gerbasi and Prentice (2013) and measured herein can be associated with benefits—rather than disutility—for others. This line of reasoning is further supported by our results of Study 3 testing the incremental validity of the specific dark traits (after controlling for D) for the diverse outcomes: whereas the residualized Self-Interest factor (which is independent from D) incrementally predicted values of Power over D, it showed positive (rather than negative) relations to Internalized Moral Identity, Perspective Taking, and Nurturance. Taken together, this suggests that those aspects comprised in Self-Interest that *are* associated with ethically, morally, or socially questionable behavior are indeed mainly due to D, but that Self-Interest additionally comprises aspects positively related to others' interests–aspects, in turn, that cannot be reconciled with the very definition of D.

Importantly, it would be premature to abandon any of the dark traits included herein based on the result that most of them failed to predict relevant criteria in a psychologically meaningful way beyond D. Although we selected the criteria in an attempt to maximize the chance that any dark trait accounts for additional variance beyond D (by relying on criteria chosen by authors originally introducing the dark traits and having demonstrated strong associations with these traits), the criteria only represent a small subset of possible outcomes in the realm of socially aversive behavior. Nevertheless, the present findings strongly imply that any future suggestions of "new" dark traits must show in which way these differs from other dark traits and, in particular, whether they capture behaviorally relevant variance beyond D. Both the aim of theoretical integration and the motivation to counteract further construct inflation can be well-served if dark traits are not considered in isolation, but rather as specific manifestations—along with others—of a general underlying tendency.

Conclusion

Herein, we proposed and critically tested the theoretical idea that dark traits are specific manifestations of a common, broad, and basic disposition. This Dark Factor of Personality, 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. Thereby, we unify and extend previous notions describing certain commonalties among the Dark Triad/Tetrad traits and provide a framework for understanding dark personality in general. In light of the importance of dark personality characteristics and, more so, their behavioral consequences for single individuals and societies at large, we maintain that research will profit from such a broader and more comprehensive view and that our conceptualization of D might function as the common denominator of individual differences in the realm of ethically, morally, and socially questionable behavior.

References

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality.* New York: Harper & Row.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracin, B. T. Johnson, & M. P. Zanna, *Handbook of Attitudes and Attitude Change: Basic Principles* (pp. 173–221). Mahwah: Erlbaum.

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*, 737–753.

Aquino, K., & Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*, 1423–1440.

Ashton, M. C., & Lee, K. (2005). Honesty-Humility, the Big Five, and the Five-Factor Model. *Journal of Personality*, *73*, 1321–1354.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*, 150–166.

Ashton, M. C., & Lee, K. (2009). The HEXACO–60: a short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*, 340–345.

Ashton, M. C., Lee, K., & De Vries, R. E. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality factors: A review of research and theory. *Personality and Social Psychology Review*, *18*, 139–152.

Asparouhov, T., & Muthén, B. (2006). Robust chi square difference testing with mean and variance adjusted test statistic. Retrieved from http://statmodel2.com/download/webnotes/webnote10.pdf

Back, M. D., Küfner, A. C., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology*, *105*, 1013–1037.

Balliet, D., Mulder, L. B., & Van Lange, P. A. (2011). Reward, punishment, and cooperation: a meta-analysis. *Psychological Bulletin*, *137*, 594–615.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, *11*, 122–133.

Barkan, R., Ayal, S., & Ariely, D. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology*, *6*, 157–161.

Baron, J. (2008). *Thinking and deciding* (4th ed.). Cambridge: Cambridge University Press.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*, 163–175.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*, 396–403.

Beauducel, A. (2011). Indeterminacy of Factor Score Estimates In Slightly Misspecified Confirmatory Factor Models. *Journal of Modern Applied Statistical Methods*, *10*, 583–598.

Blais, J., Solodukhin, E., & Forth, A. E. (2014). A meta-analysis exploring the relationship between psychopathy and instrumental versus reactive violence. *Criminal Justice and Behavior*, *41*, 797–821.

Bollen, K. A., Harden, J. J., Ray, S., & Zavisca, J. (2014). BIC and Alternative Bayesian Information Criteria in the Selection of Structural Equation Models. *Structural Equation Modeling*, *21*, 1–19.

Book, A., Visser, B. A., Blais, J., Hosker-Field, A., Methot-Jones, T., Gauthier, N. Y., … D'Agata, M. T. (2016). Unpacking more "evil": What is at the core of the dark tetrad? *Personality and Individual Differences*, *90*, 269–272.

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

Bosman, R., & Van Winden, F. (2002). Emotional hazard in a power-to-take experiment. *Economic Journal*, *112*, 147–169.

Brislin, R. W. (1980). Translation and content analysis of oral and written material. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of Cross-Cultural Psychology* (Vol. 1, pp. 389–444). Boston: Allyn & Bacon.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*, 230–258.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*, 796–846.

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

Campbell, W. K., Bonacci, A. M., Shelton, J., Exline, J. J., & Bushman, B. J. (2004). Psychological entitlement: Interpersonal consequences and validation of a self-report measure. *Journal of Personality Assessment*, *83*, 29–45.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge: Cambridge University Press.

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

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.)*. Mahwah, NJ, US: Erlbaum.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, *12*, 187–218.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.

Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2008). The development and validation of a new Machiavellianism scale. *Journal of Management*, *35*, 219–257.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126.

Dickman, S. J. (1990). Functional and dysfunctional impulsivity: personality and cognitive correlates. *Journal of Personality and Social Psychology*, *58*, 95–102.

Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods*, *6*, 352–370. https://doi.org/10.1037/1082-989X.6.4.352

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*, 430–457.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*, 583–610.

Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. London: Hodder & Stoughton.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*, 980–994.

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525–547. https://doi.org/10.1111/jeea.12014

Flexon, J. L., Meldrum, R. C., Young, J. T., & Lehmann, P. S. (2016). Low self-control and the Dark Triad: Disentangling the predictive power of personality traits on young adult substance use, offending and victimization. *Journal of Criminal Justice*, *46*, 159–169.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, *6*, 347–369.

Frick, P. J., & White, S. F. (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry*, *49*, 359–375.

Gaines, S. O., Jr., Panter, a T., Lyde, M. D., Steers, W. N., & Al, E. (1997). Evaluating the circumplexity of interpersonal traits and the manifestation of interpersonal traits in interpersonal trust. *Journal of Personality and Social Psychology*, *73*, 610–623.

Gerbasi, M. E., & Prentice, D. A. (2013). The Self-and Other-Interest Inventory. *Journal of Personality and Social Psychology*, *105*, 495–514.

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Erlbaum.

Grasmick, H. G., Tittle, C. R., Bursik, R. J., & Arneklev, B. J. (1993). Testing the core empirical implications of Gottfredson and Hirschi's general theory of crime. *Journal of Research in Crime and Delinquency*, *30*, 5–29.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. H. C. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future. A Festschrift in honor of Karl Jöreskog* (pp. 195–216). Chicago: Scientific Software International.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, *162*, 1243–1248.

Hare, R. D. (2003). *Manual for the Hare Psychopathy Checklist-Revised*. Toronto: Multi-Health Systems.

Herzog, W., Boomsma, A., & Reinecke, S. (2007). The model-size effect on traditional and modified tests of covariance structures. *Structural Equation Modeling, 14*, 361–390.

Hilbig, B. E., Glöckner, A., & Zettler, I. (2014). Personality and prosocial behavior: Linking basic traits and social value orientations. *Journal of Personality and Social Psychology, 107*, 529–539.

Hilbig, B. E., Thielmann, I., Hepp, J., Klein, S. A., & Zettler, I. (2015). From personality to altruistic behavior (and back): Evidence from a double-blind dictator game. *Journal of Research in Personality, 55*, 46–50.

Hilbig, B. E., Thielmann, I., Klein, S. A., & Henninger, F. (2016). The two faces of cooperation: On the unique role of HEXACO agreeableness for forgiveness versus retaliation. *Journal of Research in Personality, 64*, 69–78.

Hilbig, B. E., & Zettler, I. (2015). When the cat's away, some mice will play: A basic trait account of dishonest behavior. *Journal of Research in Personality, 57*, 72–88.

Hilbig, B. E., Zettler, I., & Heydasch, T. (2012). Personality, punishment and public goods: Strategic shifts towards cooperation as a matter of dispositional honesty–humility. *European Journal of Personality, 26*, 245–254.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger Publishers.

Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers, 31*, 433–438.

Jonason, P. K., Kaufman, S. B., Webster, G. D., & Geher, G. (2013). What lies beneath the Dark Triad Dirty Dozen: Varied relations with the Big Five. *Individual Differences Research, 11*, 81–90.

Jonason, P. K., Li, N. P., Webster, G. D., & Schmitt, D. P. (2009). The dark triad: Facilitating a short-term mating strategy in men. *European Journal of Personality, 23*, 5–18.

Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A Concise Measure of the Dark Triad. *Psychological Assessment, 22*, 420–432.

Jonason, P. K., Webster, G. D., Schmitt, D. P., Li, N. P., & Crysel, L. (2012). The antihero in popular culture: Life history theory and the dark triad personality traits. *Review of General Psychology, 16*, 192–199.

Jones, D. N., & Figueredo, A. J. (2013). The Core of Darkness: Uncovering the Heart of the Dark Triad. *European Journal of Personality*, *27*, 521–531.

Jones, D. N., & Paulhus, D. L. (2011). The role of impulsivity in the Dark Triad of personality. *Personality and Individual Differences, 51*, 679–682.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3): A Brief Measure of Dark Personality Traits. *Assessment, 21*, 28–41.

Kaiser, F. G., Byrka, K., & Hartig, T. (2010). Reviving Campbell's paradigm for attitude research. *Personality and Social Psychology Review*, *14*, 351–367.

Kajonius, P. J., Persson, B. N., Rosenberg, P., & Garcia, D. (2016). The (mis)measurement of the Dark Triad Dirty Dozen: Exploitation at the core of the scale. *PeerJ*, 4:e1748.

Karagonlar, G., & Kuhlman, D. M. (2013). The role of Social Value Orientation in response to an unfair offer in the Ultimatum Game. *Organizational Behavior and Human Decision Processes*, *120*, 228–239.

King, L. A. (2010). Editorial. *Journal of Personality and Social Psychology*, *98*, 104–105.

Kollock, P. (1998). Social dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology, 24*, 183–214.

Krekels, G., & Pandelaere, M. (2015). Dispositional greed. *Personality and Individual Differences*, *74*, 225–230.

Lee, K., & Ashton, M. C. (2006). Further assessment of the HEXACO Personality Inventory: Two new facet scales and an observer report form. *Psychological Assessment, 18*, 182–191.

Lee, K., & Ashton, M. C. (2014). The dark triad, the big five, and the HEXACO model. *Personality and Individual Differences, 67*, 2–5.

Lee, K., Ashton, M. C., Wiltshire, J., Bourdage, J. S., Visser, B. A., & Gallucci, A. (2013). Sex, power, and money: Prediction from the Dark Triad and Honesty–Humility. *European Journal of Personality, 27*, 169–184.

Lilienfeld, S. O., & Widows, M. R. (2005). *Psychopathy Personality Inventory Revised (PPI-R) Manual*. Lutz: Psychological Assessment Resources.

Little, T., Henrich, C., Jones, S., & Hawley, P. (2003). Disentangling the "whys" from the "whats" of aggressive behaviour. *International Journal of Behavioral Development, 27*, 122–133.

Maaß, U., & Ziegler, M. (2017). Narcissistic self-promotion is not moderated by the strength of situational cues. *Personality and Individual Differences, 104*, 482–488. https://doi.org/10.1016/j.paid.2016.09.008

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.

Mansolf, M., & Reise, S. P. (2017). When and why the second-order and bifactor models are distinguishable. *Intelligence, 61*, 120–129.

Marcus, D. K., & Zeigler-Hill, V. (2015). A Big Tent of Dark Personality Traits. *Social and Personality Psychology Compass, 98*, 434–446.

Marcus, D. K., Zeigler-Hill, V., Mercer, S. H., & Norris, A. L. (2014). The psychology of spite and the measurement of spitefulness. *Psychological Assessment, 26*, 563–574.

Markey, P. M., & Markey, C. N. (2009). A Brief Assessment of the Interpersonal Circumplex The IPIP-IPC. *Assessment, 16*, 352–361.

McCrae, R. R., & Costa, P. T. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin, *Handbook of personality psychology: Theory and research* (3rd ed., pp. 159–181). New York, NY: Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10.

Miller, J. D., & Lynam, D. R. (2015). Using Self-and Informant Reports in the Assessment of Personality Pathology in Clinical Settings—An Easy and Effective 1–2 Combination. *Clinical Psychology: Science and Practice*, *22*(1), 20–24.

Miller, J. D., Price, J., Gentile, B., Lynam, D. R., & Campbell, W. K. (2012). Grandiose and vulnerable narcissism from the perspective of the interpersonal circumplex. *Personality and Individual Differences*, *53*(4), 507–512.

Moore, C., Detert, J. R., Treviño, L., Baker, V. L., & Mayer, D. M. (2012). Why Employees Do Bad Things: Moral Disengagement and Unethical Organizational Behavior. *Personnel Psychology*, *65*, 1–48.

Moshagen, M. (2012). The model size effect in structural equation modeling: Inflated goodness-of-fit statistics are due to the size of the covariance matrix. *Structural Equation Modeling*, *19*, 86–98.

Moshagen, M., & Auerswald, M. (in press). On congruence and incongruence of measures of fit in structural equation modeling. *Psychological Methods*.

Moshagen, M., & Erdfelder, E. (2016). A New Strategy for Testing Structural Equation Models. *Structural Equation Modeling*, *23*, 54–60.

Moshagen, M., & Hilbig, B. E. (2017). The statistical analysis of cheating paradigms. *Behavior Research Methods*, *49*, 724–732.

Moshagen, M., Hilbig, B. E., Erdfelder, E., & Moritz, A. (2014). A strategy for validating questioning techniques that assess sensitive issues. *Experimental Psychology*, *61*, 48–54.

Moshagen, M., Hilbig, B. E., & Musch, J. (2011). Defection in the dark? A randomized-response investigation of cooperativeness in social dilemma games. *European Journal of Social Psychology*, *41*, 638–644.

Moshagen, M., Hilbig, B. E., & Zettler, I. (2014). Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschen Version des 60-Item HEXACO Persönlichkeitsinventars [Factor structure, psychometric properties, and measurement invariance of the German-language version of the 60-item HEXACO Personality Inventory]. *Diagnostica*, *60*, 86–97.

Moshagen, M., & Musch, J. (2014). Sample size requirements of the robust weighted least squares estimator for confirmatory factor analysis with ordinal data. *Methodology*, *10*, 60–70.

Moshagen, M., Musch, J., Ostapczuk, M., & Zhao, Z. (2010). Reducing socially desirable responses in epidemiologic surveys: An extension of the randomized-response-technique. *Epidemiology*, *21*, 379–382.

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). *Perspectives on Psychological Science*, *12*, 183–204.

Murphy, R. O., & Ackermann, K. a. (2014). Social Value Orientation: Theoretical and Measurement Issues in the Study of Social Preferences. *Personality and Social Psychology Review*, *18*, 13–41.

Muthén, B., duToit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with

categorical and continous outcomes. Retrieved from http://statmodel.com/download/Article_075.pdf

Muthén, L. K., & Muthén, B. (2015). *Mplus user's guide* (7th ed.). Los Angeles: Muthén & Muthén.

O'Boyle, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, *97*(3), 557.

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). *Psychological Assessment, 23*, 523–531.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley, *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah: Erlbaum.

Paulhus, D. L. (2014). Toward a Taxonomy of Dark Personalities. *Current Directions in Psychological Science*, *23*, 421–426.

Paulhus, D. L., & Jones, D. N. (2015). Measures of Dark Personalities. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measures of Personality and Social Psychological Constructs* (pp. 562–594). San Diego: Academic Press.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556–563.

Postmes, T., & Spears, R. (1998). Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin*, *123*, 238–259. https://doi.org/10.1037/0033-2909.123.3.238

Rand, D. G. (2012). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, *299*, 172–179.

Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, *49*, 243–256.

Reise, S. P. (2012). The Rediscovery of Bifactor Measurement Models. *Multivariate Behavioral Research*, *47*, 667–696.

Roberts, C., Vakirtzis, A., Kristjánsdóttir, L., & Havlíček, J. (2013). Who punishes? Personality traits predict individual variation in punitive sentiment. *Evolutionary Psychology*, *11*, 186–200.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Statistical Indices Evaluating Bifactor Models : Calculating and Interpreting Statistical Indices. *Psychological Methods*, *21*, 137–150.

Ruiz, J. M., Smith, T. W., & Rhodewalt, F. (2001). Distinguishing Narcissism and Hostility: Similarities and Differences in Interpersonal Circumplex and Five-Factor Correlates. *Journal of Personality Assessment*, *76*, 537–555.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514.

Schütte, N., Blickle, G., Frieder, R. E., Wihler, A., Schnitzler, F., Heupel, J., & Zettler, I. (2017). The Role of Interpersonal Influence in Counterbalancing Psychopathic Personality Trait Facets at Work. *Journal of Management*. https://doi.org/10.1177/0149206315607967

Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In M. Zanna, *Advances in Experimental Social Psychology* (Vol. 25, pp. 1–65). New York: Academic Press.

Schyns, B. (2015). Dark Personality in the Workplace: Introduction to the Special Issue. *Applied Psychology*, *64*, 1–14.

Seuntjens, T. G., Zeelenberg, M., van de Ven, N., & Breugelmans, S. M. (2015). Dispositional greed. *Journal of Personality and Social Psychology*, *108*, 917–933.

Shalvi, S., Handgraaf, M. J. J., & De Dreu, C. K. W. (2011). Ethical manoeuvring: Why people avoid both major and minor lies. *British Journal of Management*, *22*, S16–S27.

Shirtcliff, E. A., Vitacco, M. J., Graf, A. R., Gostisha, A. J., Merz, J. L., & Zahn-Waxler, C. (2009). Neurobiology of empathy and callousness: implications for the development of antisocial behavior. *Behavioral Sciences & The Law*, *27*, 137–171.

Ten Berge, J. M. F., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*, 613–625.

Thielmann, I., Heck, D. W., & Hilbig, B. E. (2016). Anonymity and incentives: An investigation of techniques to reduce socially desirable responding in the Trust Game. *Judgment and Decision Making*, *11*, 527–536.

Tracey, T. J. G. (2000). Analysis of circumplex models. In H. E. A. Tinsley & S. D. Brown, *Handbook of applied multivariate statistics and mathematical modeling* (pp. 641–664). New York: Academic Press.

Veselka, L., & Vernon, P. A. (2014). Introduction to the special issue on the Dark Triad and related traits. *Personality and Individual Differences*, *67*, 1.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin and Review*, *11*, 192–196.

Weigel, R. H., Hessing, D. J., & Elffers, H. (1999). Egoism: Concept, measurement and implications for deviance. *Psychology, Crime & Law*, *5*, 349–378.

Weigold, A., Weigold, I. K., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. *Psychological Methods*, *18*, 53–70. https://doi.org/10.1037/a0031607

Wilson, E. O. (1975). *Sociobiology: The New Synthesis* (Vol. 233). Cambridge: Harvard University Press.

Yang, M., Jiang, G., & Yuan, K.-H. (in press). The Performance of Ten Modified Rescaled Statistics as the Number of Variables Increases. *Structural Equation Modeling*. https://doi.org/10.1080/10705511.2017.1389612

Yang-Wallentin, F., Jöreskog, K. G., & Luo, H. (2010). Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling*, *17*, 392–423.

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*, 165–200.

Yuan, K. H., Tian, Y., & Yanagihara, H. (2015). Empirical Correction to the Likelihood Ratio Statistic for Structural Equation Modeling with Many Variables. *Psychometrika*, *80*, 379–405.

Yuan, K. H., Yang-Wallentin, F., & Bentler, P. M. (2012). ML Versus MI for Missing Data With Violation of Distribution Conditions. *Sociological Methods & Research*, *41*, 598–629. https://doi.org/10.1177/0049124112460373

Yung, Y.-F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*, 113–128.

Zettler, I., & Hilbig, B. E. (2015). Honesty and Humility. In J. D. Wright, *International Encyclopedia of Social and Behavioral Sciences* (2nd ed., Vol. 11, pp. 169–174). Oxford: Elsevier.

Zettler, I., Hilbig, B. E., Moshagen, M., & de Vries, R. E. (2015). Dishonest responding or true virtue? A behavioral test of impression management. *Personality and Individual Differences*, *81*, 107–111.

Zhang, L., & Ortmann, A. (2016). Pro-social or anti-social, or both? A within-and between-subjects study of social preferences. *Journal of Behavioral and Experimental Economics*, *62*, 23–32.

Ziegler, M., MacCann, C., & Roberts, R. (2012). *New perspectives on faking in personality assessment*. Oxford: Oxford University Press.

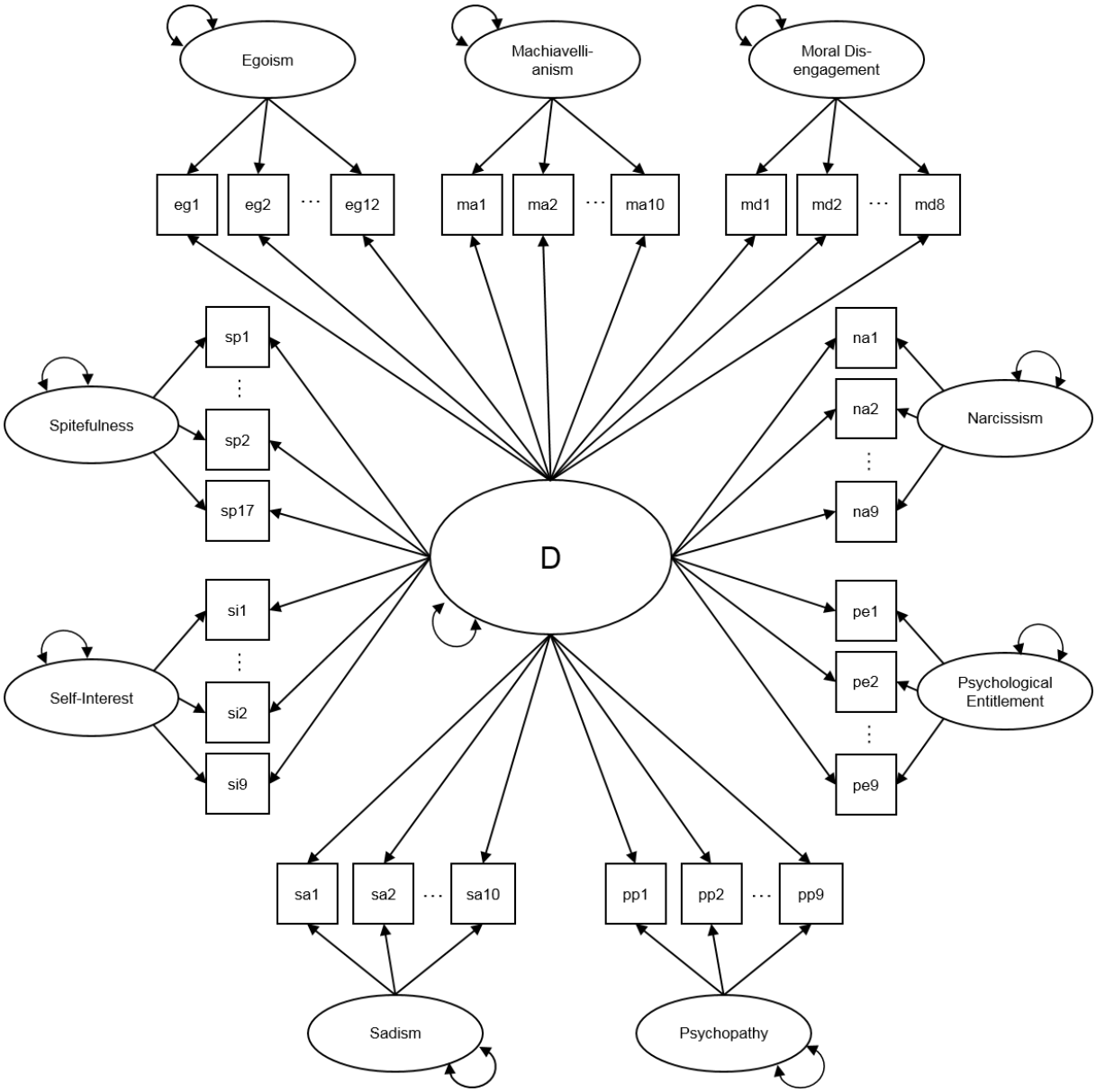

Figure 1. Bifactor model for nine dark traits and their common dark core (D). D captures the shared variance across all items, whereas the specific dark traits represent the remaining covariance among the items of the respective scale after controlling for D. To enhance clarity, only a subset of three indicators per dark trait are displayed and the residual variances of the items are omitted.

Table 1

Overview of Dark Traits Included and Corresponding Inventories

Trait	Definition	Inventory			Source
		Name	Number of Items	Sample Item	
Egoism	„the excessive concern with one's own pleasure or advantage at the expense of community well-being" (p. 349)	Egoism Scale	12 (0)	"It is hard to get ahead without cutting corners here and there."	Weigel et al. (1999)
Machiavellianism	„(a) manipulativeness, (b) callous affect, and (c) a strategic-calculating orientation" (p. 29)	Short Dark Triad	10 (1)	"I like to use clever manipulation to get my way."	Jones & Paulhus (2014)
Moral Disengagement	„a generalized cognitive orientation to the world that differentiates individuals' thinking in a way that powerfully affects unethical behavior" (p. 6)	Propensity to Morally Disengage Scale	8 (0)	"Some people have to be treated roughly because they lack feelings that can be hurt."	Moore et al. (2012)
Narcissism	„ego-reinforcement is the all-consuming motive" (p. 30)	Short Dark Triad	9 (3)	"I know that I am special because everyone keeps telling me so."	Jones & Paulhus (2014)
Psychological Entitlement	„a stable and pervasive sense that one deserves more and is entitled to more than others" (p. 31)	Psychological Entitlement Scale	9 (1)	"I honestly feel I'm just more deserving than others."	Campbell et al. (2004)
Psychopathy	„deficits in affect (i.e., callousness) and self-control (i.e., impulsivity)" (p. 29)	Short Dark Triad	9 (2)	"I'll say anything to get what I want."	Jones & Paulhus (2014)
Sadism	„a person who humiliates others, shows a longstanding pattern of cruel or demeaning behavior to others, or intentionally inflicts physical, sexual, or psychological pain or suffering on others in order to assert power and dominance or for pleasure and enjoyment" (p. 523)	Short Sadistic Impulse Scale	10 (1)	"Hurting people would be exciting."	O'Meara et al. (2011)
Self-Interest	„the pursuit of gains in socially valued domains, including material goods, social status, recognition, academic or occupational achievement, and happiness" (p. 496)	Self- and Other-Interest Inventory	9 (0)	"I try to make sure others know about my successes."	Gerbasi & Prentice (2013)
Spitefulness	„a preference that would harm another but that would also entail harm to oneself. This harm could be social, financial, physical, or an inconvenience" (p. 566)	Spitefulness Scale	17 (2)	"It is sometimes worth a little suffering on my part to see others receive the punishment they deserve."	Marcus et al. (2014)

Note. Number of negatively-keyed items in parenthesis.

Table 2

Descriptive Statistics and Correlations among the Non-Residualized Dark Traits (Study 1)

Variable	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Egoism	2.43	0.55	(.78)								
2. Machiavellianism	2.77	0.62	.78	(.81)							
3. Moral Disengagement	1.94	0.47	.75	.68	(.67)						
4. Narcissism	2.80	0.56	.17	.27	.30	(.70)					
5. Psychological Entitlement	2.53	0.64	.47	.56	.55	.59	(.84)				
6. Psychopathy	1.94	0.54	.73	.80	.68	.36	.52	(.72)			
7. Sadism	1.64	0.63	.50	.55	.47	.29	.41	.79	(.86)		
8. Self-Interest	3.42	0.58	.20	.41	.33	.53	.49	.37	.19	(.80)	
9. Spitefulness	1.77	0.50	.59	.68	.61	.25	.43	.80	.62	.24	(.86)

Note. $N = 304$. Shown are latent correlations and Cronbach's alpha estimates of internal consistency (on the diagonal). All correlations significantly differ from zero at $p < .05$.

Table 3

Mean and Range of Loadings on D and Specific Factors (Studies 1-3)

Items	Study 1 (*N* = 304)		Study 2 (*N* = 1,094)		Study 3 (*N* = 1,261)	
	D	S	D	S	D	S
Egoism	.39 (.10 - .64)	.27 (-.12 - .62)	.38 (.00 - .54)	.31 (.12 - .55)	.40 (.15 - .55)	.36 (.19 - .49)
Machiavellianism	.49 (.29 - .66)	.25 (.08 - .61)	.46 (.23 - .68)	.28 (.07 - .49)	.42 (.12 - .67)	.34 (.03 - .58)
Moral Disengagement	.36 (.23 - .48)	.29 (.09 - .63)	.38 (.20 - .58)	.30 (.16 - .43)	.44 (.29 - .60)	.35 (.19 - .61)
Narcissism	.18 (.00 - .29)	.42 (.14 - .62)	.16 (.09 - .23)	.44 (.31 - .65)	.20 (.06 - .30)	.44 (.16 - .61)
Psychological Entitlement	.37 (.19 - .54)	.48 (.34 - .66)	.38 (.29 - .53)	.49 (.33 - .63)	.39 (.19 - .53)	.51 (.31 - .66)
Psychopathy	.44 (.09 - .57)	.20 (-.14 - .42)	.47 (.14 - .62)	.22 (-.06 - .51)	.47 (.09 - .62)	.24 (.07 - .59)
Sadism	.46 (.34 - .57)	.45 (.28 - .62)	.51 (.45 - .62)	.42 (.25 - .55)	.40 (.30 - .51)	.44 (.15 - .66)
Self-Interest	.22 (.10 - .39)	.50 (.25 - .68)	.16 (-.06 - .34)	.51 (.21 - .73)	.26 (.13 - .39)	.57 (.44 - .73)
Spitefulness	.42 (-.07 - .61)	.35 (-.13 - .51)	.45 (-.10 - .61)	.27 (-.16 - .52)	.43 (-.18 - .58)	.33 (-.07 - .53)

Note. Mean and Range of loadings (min - max) of the items of the respective scale on the general factor (D) and the respective specific factor (S).

Table 4

Common Variance explained by D (Studies 1-3)

Items	Study 1 (N = 304)		Study 2 (N = 1,094)		Study 3 (N = 1,261)	
	ECV	R^2	*ECV*	R^2	*ECV*	R^2
Egoism	.57	.30	.59	.28	.57	.32
Machiavellianism	.73	.35	.69	.33	.60	.34
Moral Disengagement	.54	.25	.61	.26	.61	.26
Narcissism	.18	.24	.12	.24	.18	.43
Psychological Entitlement	.38	.39	.38	.40	.38	.35
Psychopathy	.74	.30	.75	.34	.75	.34
Sadism	.50	.42	.58	.45	.43	.39
Self-Interest	.18	.33	.13	.33	.18	.41
Spitefulness	.58	.34	.70	.33	.62	.35

Note. *ECV* denotes the percentage of common variance explained by D. R^2 gives the proportion of variance explained by the full model (D and specific factors) in the items of each respective scale.

Table 5

Descriptive Statistics and Correlations among the Non-Residualized Dark Traits (Study 2)

Variable	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Egoism	2.83	0.62	(.80)								
2. Machiavellianism	3.10	0.65	.80	(.80)							
3. Moral Disengagement	2.13	0.55	.71	.70	(.71)						
4. Narcissism	2.82	0.60	.14	.22	.31	(.71)					
5. Psychological Entitlement	2.57	0.69	.45	.55	.56	.58	(.85)				
6. Psychopathy	2.22	0.64	.70	.77	.73	.38	.52	(.76)			
7. Sadism	1.88	0.75	.60	.63	.58	.21	.40	.78	(.88)		
8. Self-Interest	3.48	0.60	.10	.32	.24	.57	.50	.21	.10	(.79)	
9. Spitefulness	2.02	0.57	.64	.70	.72	.24	.51	.84	.69	.17	(.86)
10. Dictator Game Selfishness	59.5%	18.1%	.24	.34	.34	.19	.32	.32	.27	.17	.30

Note. Dictator Game Selfishness = proportion of money kept in the dictator game. Shown are latent correlations and Cronbach's alpha estimates of internal consistency (on the diagonal). All correlations significantly differ from zero at $p < .05$.

Table 6

Latent Regression Results for Predicting the Amount of Money Kept in the Dictator Game (Studies 2 and 3) and Cheating Behavior (Study 3)

Predictor	Dictator Game Selfishness		Cheating Task (Study 3, N = 883)
	Hypothetical (Study 2, N = 1,094)	Incentivized (Study 3, N = 883)	
	β (SE)		b (SE)
D	.38 (0.04) **	.17 (0.04) **	0.19 (0.05) **
S_{Egoism}	-.07 (0.04)	-.04 (0.04)	0.04 (0.06)
$S_{Machiavellianism}$	.00 (0.05)	.04 (0.05)	-0.04 (0.06)
$S_{MoralDisengagement}$	.02 (0.05)	-.09 (0.05) *	-0.04 (0.06)
$S_{Narcissism}$	.03 (0.04)	-.10 (0.05) **	0.02 (0.05)
$S_{PsychologicalEntitlement}$	.09 (0.04) *	.09 (0.04) *	0.01 (0.06)
$S_{Psychopathy}$	-.06 (0.05)	.01 (0.05)	-0.01 (0.05)
S_{Sadism}	.00 (0.04)	-.04 (0.04)	-0.04 (0.05)
$S_{Self\text{-}Interest}$	.01 (0.04)	.04 (0.04)	0.07 (0.05)
$S_{Spitefulness}$	-.03 (0.05)	.05 (0.04)	-0.01 (0.05)
R^2	.16 **	.07 **	

Note. Latent standardized linear regression coefficients (β) with standard errors in parenthesis for the prediction of the amount of money kept (i.e., selfishness) in the dictator game by D and specific factors (S_{Trait}). Latent standardized probit regression coefficients (b) with standard errors in parenthesis for the prediction of cheating. All predictor variables are mutually uncorrelated.

* $p < .05$.

** $p < .01$.

Table 7

Overview of Self-Report Criterion Measures (Study 3)

Criterion	Selected with Regard to which Dark Trait	Reported *r*(s) with Dark Trait (in original article)	Instrument		
			Name	Number of Items	Sample Item
Self-Centeredness	Egoism	.35, .46, .47 (Weigel et al., 1999)	Self-Centeredness Scale	4 (0)	"I try to look out for myself first, even if it means making things difficult for other people."
Nurturance	Machiavellianism (Psychopathy)	-.43 (-.49) (Jones & Paulhus, 2014)	International Personality Item Pool—Interpersonal Circumplex: Nurturance Scale	32 (0)	"[I] inquire about others' well-being."
Internalized Moral Identity	Moral Disengagement	-.42, -.55 (Moore et al., 2012)	Moral Identity: Internalization Scale	5 (2)	"Being someone who has these characteristics [caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, kind] is an important part of who I am."
Dominance	Narcissism	.65 (Jones & Paulhus, 2014)	International Personality Item Pool—Interpersonal Circumplex: Dominance Scale	32 (0)	"[I] demand to be the center of interest."
Perspective Taking	Psychological Entitlement	-.43 (Campbell et al., 2004)	Interpersonal Reactivity Index: Perspective Taking Scale	7 (2)	"If I'm sure I'm right about something, I don't waste much time listening to other people's arguments." (reversed coded)
Impulsivity	Psychopathy	— ($.32 \leq r \leq .52$, Jones & Paulhus, 2011)[a]	Dysfunctional Impulsivity Scala	12 (4)	"I often say and do things without considering the consequences."
Insensitivity	Sadism	-.41 (O'Meara et al., 2011)	Empathy Quotient: Insensitivity	16 (0)	"Seeing people cry doesn't really upset me."
Power	Self-Interest	.65 (Gerbasi & Prentice, 2013)	Schwartz Value Survey: Power	4 (0)	"Please, rate the importance of the following values as a life-guiding principle for you: Social Power (control over others, dominance)"
Aggression	Spitefulness	.52, .58 (Marcus et al., 2014)	Forms and Functions of Aggression Scale	36 (0)	"To get what I want, I often put others down."

Note. Number of negatively-keyed items in parenthesis.

[a] Impulsivity was not assessed in the original article. Please see the text for further explanations.

Table 8

Descriptive Statistics and Correlations among Non-Residualized Dark Traits and Criteria (Study 3)

Variable	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Egoism	2.66	0.58	(.83)								
2. Machiavellianism	2.89	0.59	.74	(.81)							
3. Moral Disengagement	2.04	0.58	.72	.63	(.78)						
4. Narcissism	2.74	0.54	.22	.32	.25	(.73)					
5. Psychological Entitlement	2.66	0.63	.50	.58	.47	.61	(.86)				
6. Psychopathy	1.89	0.55	.63	.72	.68	.38	.49	(.76)			
7. Sadism	1.49	0.48	.45	.43	.49	.26	.33	.72	(.81)		
8. Self-Interest	3.36	0.59	.27	.45	.26	.66	.54	.31	.17	(.85)	
9. Spitefulness	1.82	0.49	.60	.59	.74	.26	.44	.74	.59	.21	(.86)
Dictator Game Selfishness	72.6%	26.4	.11	.17	.09	.04	.17	.14	.08	.11	.15
Cheating Task[a]	37.4%	48.4	.17	.13	.13	.12	.13	.17	.10	.16	.15
Self-Centeredness	2.18	0.57	**.50**	.53	.55	.30	.43	.62	.48	.30	.49
Nurturance	0.36	0.86	-.43	**-.41**	-.43	.11	-.23	**-.56**	-.49	-.01	-.45
Internalized Moral Identity	3.82	0.54	-.23	-.17	**-.24**	-.02	-.12	-.35	-.28	.10	-.25
Dominance	-1.07	0.71	.12	.23	.22	**.66**	.29	.35	.27	.34	.24
Perspective Taking	3.33	0.45	-.22	-.26	-.29	.04	**-.15**	-.35	-.25	.04	-.33
Impulsivity	2.37	0.46	.38	.24	.38	.08	.12	**.48**	.30	.11	.40
Insensitivity	2.62	0.44	.49	.45	.47	.23	.34	.60	**.47**	.21	.46
Power	2.85	0.96	.28	.47	.28	.50	.48	.36	.23	**.60**	.29
Aggression	1.63	0.53	.42	.40	.50	.20	.25	.64	.57	.15	**.52**

Note. Dictator Game Selfishness = proportion of money kept in the dictator game. Shown are latent correlations and Cronbach's alpha estimates of internal consistency (on the diagonal). Correlations between a non-residualized dark trait and its relevant target criterion (see Table 7 and the text) are printed in bold. For the correlations among the dark traits, N = 1,261. For the correlations between the dark traits and the criteria, N = 882. Correlation estimates $|r| \geq .06$ (N = 1,261) and $\geq .07$ (N = 882), respectively, differ significantly from zero at $p < .05$.

[a] Polychoric correlation estimates for the cheating task.

Table 9

Latent Regression Results for Predicting Self-Report Criteria by D and the Specific Factors (Study 3)

Predictor	Criterion Variable								
	Self-Centeredness	Nurturance	Internalized Moral Identity	Dominance	Perspective Taking	Impulsivity	Insensitivity	Power	Aggression
D	0.64 (0.03) **	-0.54 (0.03) **	-0.30 (0.04) **	0.26 (0.03) **	-0.34 (0.04) **	0.42 (0.04) **	0.57 (0.03) **	0.47 (0.04) **	0.62 (0.03) **
S_{Egoism}	**0.00 (0.04)**	0.04 (0.03)	-0.03 (0.03)	-0.10 (0.03) **	0.09 (0.04) *	0.13 (0.03) **	0.11 (0.03) **	-0.08 (0.03) *	-0.04 (0.03)
$S_{Machiavellianism}$	0.02 (0.03)	**0.11 (0.03) ****	0.10 (0.03) **	0.00 (0.03)	-0.02 (0.04)	-0.20 (0.04) **	-0.05 (0.03)	0.07 (0.03) *	-0.11 (0.03) **
$S_{MoralDisengagement}$	0.02 (0.04)	0.15 (0.03) **	**0.01 (0.04)**	-0.01 (0.03)	0.01 (0.04)	0.02 (0.04)	-0.05 (0.04)	-0.10 (0.04) *	-0.03 (0.03)
$S_{Narcissism}$	0.02 (0.03)	0.45 (0.02) **	0.00 (0.03)	**0.66 (0.02) ****	0.14 (0.04) **	-0.07 (0.03) *	0.02 (0.03)	0.13 (0.03) **	-0.02 (0.02)
$S_{PsychologicalEntitlement}$	0.04 (0.03)	-0.02 (0.02)	-0.02 (0.04)	-0.08 (0.02) **	**-0.03 (0.03)**	-0.13 (0.03) **	0.00 (0.03)	0.08 (0.03) *	-0.10 (0.02) **
$S_{Psychopathy}$	0.09 (0.04) *	**-0.06 (0.03)**	-0.11 (0.04) *	0.06 (0.03) *	-0.07 (0.04)	**0.16 (0.04) ****	0.17 (0.04) **	0.00 (0.04)	0.12 (0.03) **
S_{Sadism}	0.07 (0.03) *	-0.10 (0.03) **	-0.06 (0.04)	0.03 (0.02)	0.00 (0.03)	-0.05 (0.03)	**0.09 (0.03) ****	-0.01 (0.04)	0.16 (0.04) **
$S_{Self\text{-}Interest}$	0.03 (0.03)	0.08 (0.02) **	0.21 (0.03) **	0.02 (0.02)	0.14 (0.03) **	0.07 (0.03) *	0.01 (0.03)	**0.38 (0.03) ****	0.00 (0.02)
$S_{Spitefulness}$	-0.02 (0.04)	0.04 (0.03)	0.01 (0.04)	-0.01 (0.03)	-0.08 (0.04)	0.07 (0.04) *	-0.01 (0.03)	-0.03 (0.04)	**0.03 (0.03)**
R^2	0.43 **	0.55 **	0.16 **	0.52 **	0.17 **	0.30 **	0.38 **	0.41 **	0.45 **

Note. N = 882. Latent standardized linear regression coefficients and standard errors in parenthesis. All predictor variables are mutually uncorrelated. Regression coefficients for a specific dark trait (S_{Trait}) in the prediction of its relevant target criterion (see text for details) are printed in bold.

* $p < .05$.

** $p < .01$.

Table 10

Correlations between D, Specific Factors, and Basic Personality Traits (Study 3)

	NEO-FFI						HEXACO-60						
	OP	CO	EX	AG	NE	R^2	HH	EM	EX	AG	CO	OP	R^2
D	**-.11**	**-.27**	**-.15**	**-.69**	**.20**	.54	**-.80**	**-.09**	**-.17**	**-.45**	**-.32**	**-.05**	.70
S_{Egoism}	-.04	.03	**-.05**	-.03	**.17**	.05	**.14**	.02	**-.06**	-.02	**-.06**	**-.07**	.04
$S_{Machiavellianism}$	**-.06**	**.17**	.01	**.09**	**-.04**	.04	**.07**	-.02	.02	**.06**	**.17**	**-.10**	.05
$S_{MoralDisengagement}$	**-.17**	**.10**	**.10**	**.19**	**-.05**	.08	**.12**	.02	**.10**	**.13**	.01	**-.21**	.09
$S_{Narcissism}$	**.29**	**.15**	**.57**	**.05**	**-.45**	.43	**-.07**	**-.33**	**.71**	**.07**	**.18**	**.36**	.60
$S_{PsychologicalEntitlement}$	-.03	**.16**	**-.04**	**.04**	-.02	.03	.02	.01	.00	**.08**	**.09**	-.02	.02
$S_{Psychopathy}$	.02	.01	**-.09**	**-.18**	**.11**	.05	**.26**	.01	**-.07**	**-.34**	.01	.01	.32
S_{Sadism}	.04	**-.05**	**-.04**	**-.04**	**.05**	.01	.03	.01	**-.06**	.01	-.02	.01	.01
$S_{Self\text{-}Interest}$	**.04**	**.28**	**.11**	.02	**.11**	.19	**-.19**	**.20**	.01	**-.04**	**.24**	**.06**	.18
$S_{Spitefulness}$	-.04	.01	**.07**	**.12**	.03	.03	**.17**	**.10**	**.04**	**.06**	.02	**-.05**	.05

Note. N = 1,261. Latent correlation estimates. R^2 gives the proportion of variance explained by regressing D or a specific factor (S_{Trait}) on all traits of a given inventory (NEO-FFI or HEXACO-60). Correlations printed in bold significantly differ from zero at $p < .05$. OP = Openness to experience, CO = Conscientiousness, EX = Extraversion, AG = Agreeableness, NE = Neuroticism, HH = Honesty-Humility, EM = Emotionality.

Table 11

Correlations of Various Criteria to D and Honesty-Humility and Incremental Variance Explained by D over Honesty-Humility (Study 3).

Criterion	r_{HH}	r_D	R^2	ΔR^2
Self-Centeredness	-.56 (.03)	.64 (.03)	.31	.11**
Nurturance	.45 (.03)	-.58 (.02)	.20	. 13**
Internalized Moral Identity	.20 (.04)	-.30 (.04)	.04	.06**
Dominance	-.40 (.03)	.34 (.03)	.16	< .01
Perspective Taking	.31 (.04)	-.35 (.04)	.10	.03**
Impulsivity	-.38 (.03)	.43 (.04)	.15	.04**
Insensitivity	-.40 (.03)	.57 (.03)	.16	.18**
Power	-.53 (.03)	.47 (.04)	.28	< .01
Aggression	-.51 (.03)	.62 (.03)	.26	.13**
Dictator Game Selfishness	-.22 (.03)	.17 (.03)	.05	< .01
Cheating Task[a]	-.26 (.05)	.19 (.05)		

Note. Latent correlation estimates and standard errors in parenthesis. r_{HH} and r_D denote the bivariate correlations between the Honesty-Humility factor (HH) and D to the criterion variables, respectively. All bivariate correlations differ significantly from zero at $p < .01$. R^2 denotes the variance explained when predicting the criterion by HH. ΔR^2 represents the increase in R^2 when adding D as a predictor.

** $p < .01$.

[a] Polychoric correlation estimates for the cheating task. The log-likelihood ratio test comparing the model with HH and D as predictors vs. a model that constrained the path between cheating and D to zero was $\Delta\chi^2(1) = 0.35$, $p = .55$.

Table 12

Correlations between Complete and Reduced D Factor Scores Omitting one Dark Trait (Study 4)

Reduced D	D	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. D(−Egoism)	.99									
2. D(−Psychological Entitlement)	.99	.99								
3. D(−Machiavellianism)	.99	.98	.99							
4. D(−Moral Disengagement)	.99	.99	.99	.98						
5. D(−Narcissism)	.99	.99	.99	.99	.99					
6. D(−Psychopathy)	.97	.95	.96	.94	.96	.97				
7. D(−Sadism)	.99	.98	.98	.97	.98	.99	.98			
8. D(−Self-Interest)	.99	.99	.99	.99	.99	.99	.96	.98		
9. D(−Spitefulness)	.98	.97	.97	.95	.97	.98	.98	.99	.97	
10. D(−Dark Triad)	.95	.99	.99	.99	.99	.99	.97	.99	.99	.98

Note. Reduced D factor scores were determined by omitting the items of the trait or traits stated in the parenthesis (−Trait(s)). The mean correlation of the reduced Ds to the complete version of D was r = .99. The mean correlation among the reduced Ds was r = .98.

Table 13

Correlations between Non-Residualized Dark Traits, Target Criteria, and Reduced Ds Omitting the Respective Dark Trait (Study 4)

Dark Trait	Target Criterion	$r_{D(-trait)}$	r_{trait}	R^2	ΔR^2
Egoism	Self-Centeredness	.63 (.03)	.50 (.03)	.40	.01*
Machiavellianism	Nurturance	-.53 (.03)	-.41 (.03)	.28	< .01
Moral Disengagement	Internalized Moral Identity	-.30 (.04)	-.24 (.04)	.09	< .01
Narcissism	Dominance	.24 (.03)	.66 (.03)	.06	.41**
Psychological Entitlement	Perspective Taking	-.35 (.04)	-.15 (.04)	.12	< .01
Psychopathy	Impulsivity	.39 (.04)	.48 (.03)	.15	.06**
Sadism	Insensitivity	.55 (.03)	.47 (.03)	.30	.03**
Self-Interest	Power	.43 (.04)	.60 (.03)	.19	.21**
Spitefulness	Aggression	.59 (.03)	.52 (.03)	.35	.01**

Note. Latent correlation estimates and standard errors in parenthesis. $r_{D(-trait)}$ refers to the bivariate correlation of the target criterion to the reduced D as obtained by omitting the respective dark trait from the measurement model of D. r_{trait} gives the correlation between a non-residualized dark trait and its target criterion. All bivariate correlations differ significantly from zero at $p < .01$. R^2 denotes the proportion of variance explained when predicting the criterion by the reduced D factor (all significantly different from zero). ΔR^2 represents the increase in R^2 when adding the respective non-residualized dark trait as a predictor.

* $p < .05$.

** $p < .01$.

Table 14

Distribution of Correlations between Full and Reduced D Factor Scores Omitting 50% or 75% of Items (1,000 Random Samples each, Study 4)

% Omitted Items		5th perc.	25th perc.	Median	75th perc.	95th perc.
50%	Reduced Ds	.93	.95	.96	.96	.97
	D	.96	.97	.97	.98	.98
75%	Reduced Ds	.83	.87	.89	.90	.93
	D	.90	.92	.94	.95	.96

Note. perc. = percentile. The rows labelled "reduced Ds" show the distribution of the correlations among the reduced D factor scores as obtained from measuring D by a randomly chosen subset of items. The rows labelled D show the distributions of correlations between the reduced D factor scores and the full D factor score.