



Published in final edited form as:

*J Res Pers.* 2021 February ; 90: . doi:10.1016/j.jrp.2020.104059.

## On the importance of the assessment and conceptualization of Agreeableness: A commentary on “Agreeableness and the common core of dark traits are functionally different constructs”

Colin E. Vize, Donald R. Lynam

Purdue University

Moshagen and colleagues (2020) sought to examine the distinctiveness of Agreeableness and the shared variance among so-called “dark” personality traits (referred to as D or the D factor; Moshagen, Hilbig, & Zettler, 2018). To date, previous work on the D factor had only examined its overlap with more limited assessments of Agreeableness and related traits (e.g., Honesty/Humility from the HEXACO). Thus, we commend the authors for attempting to more comprehensively examine the question that naturally arises when considering the D factor—in what ways does D differ from (low) Agreeableness? The authors also deserve credit for their work in advancing research on “dark” traits given recent critiques of the literature (Miller, Vize, Crowe, & Lynam, 2019). However, we disagree with their conclusion that Agreeableness and the D factor are functionally different constructs.

We believe that Moshagen et al. (2020) too easily dismiss data that run contrary to their conclusion. For example, computing the similarities between the correlational profiles of latent D and A from Table 1 (Moshagen et al, 2020) yields a similarity coefficient of  $-0.95$  which suggests that D and A are exact opposites.<sup>1</sup> The similarity coefficient actually reported by Moshagen et al. of 0.746 is still remarkably high given that “the outcome criteria were selected to represent theoretically implied differences between Agreeableness and D” (Moshagen et al., p. 9). These results are noteworthy, as studies have highlighted how correlational profiles provide essential information regarding construct (dis)similarity (Westen & Rosenthal, 2003; Furr & Heuckeroth, 2019).

In a similar vein, we believe an even more fundamental problem lies in the way they chose to conceptualize and assess the Agreeableness domain. As they note, their results hinge on whether, “the chosen operationalizations can be seen as comprehensive indicators of the constructs they intend to represent.” (p. 8). Unfortunately, their operationalization of Agreeableness was limited. When both D and Agreeableness are comprehensively assessed,

---

*Corresponding Author:* Colin E. Vize, University of Pittsburgh, Department of Psychiatry, Pittsburgh, PA.: vizece@upmc.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

<sup>1</sup>The similarity coefficient of 0.746 reported by Moshagen et al. depends on reversing the signs of the correlations for A, which reduces the variance of the correlations being examined and, in turn, reduces the size of the squared differences in the z-scores of the two columns being compared. This reduces the correlation between the columns, as can be seen in the following formula for the correlation (Cohen & Cohen, 1983):  $r = 1 - (0.5 * (\sum(Z_x - Z_y)^2 / (n - 1)))$ . See Roger and Nicewander (1988) for a more thorough review of the correlation coefficient and various ways to compute it.

there is little to distinguish the constructs from one another (Vize, Miller, & Lynam, 2020). The comprehensiveness of the operationalization of A matters greatly.

## Conceptualization of the Agreeableness Domain

Before considering the empirical results, we note that the authors have yet to offer a compelling conceptual framework for how D differs from (low) Agreeableness. Moshagen and colleagues (2020) critique previously offered definitions of Agreeableness as too broad (Graziano & Tobin, 2017). This assessment seems fair—to our knowledge, no attempt has been made to offer a rigorous definition of the domain. It is typically described in terms of trait adjectives. However, the Antagonism domain in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2011), understood to be the low end of Agreeableness, is defined as, “behaviors that put the individual at odds with other people, including an exaggerated sense of self-importance and a concomitant expectation of special treatment, as well as a callous antipathy toward others, encompassing both an unawareness of others’ needs and feelings and a readiness to use others in the service of self-enhancement.” (p. 780). The definition of D – “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.” (Moshagen et al., 2018, p. 657) – bears a remarkable similarity to DSM-5 Antagonism.

Though newer constructs are not required to always give deference to old constructs, the conceptual similarities between D and (low) Agreeableness require a rigorous conceptual analysis to make clear the purported distinction between D and (low) Agreeableness. Conceptually, the purported differences between Agreeableness and D appear to us to be an instance of the “jangle” fallacy, where the same construct is given different names (Block, 2000).<sup>2</sup> Moshagen and colleagues (2018; 2020) have offered a clear definition for the shared variance among socially aversive traits, and highlighted ways in which definitions of Agreeableness can be further developed. Nonetheless, the definition of D does not appear to sufficiently distinguish it from (low) Agreeableness.

## Assessment of the Agreeableness Domain

Turning to the empirical evidence, Moshagen et al. (2020) modeled a latent Agreeableness domain using five Agreeableness scales from established personality inventories: the Big Five Inventory-2 (Soto & John, 2017), the Big Five Aspects Scale (DeYoung et al., 2007), the IPIP-BFI (Goldberg, 1992), the NEO-FFI (Costa & McCrae, 1992), and the HEXACO-100 (Lee & Ashton, 2018). Importantly, across these Agreeableness scales, there is little content related to modesty or straightforwardness (e.g., Miller et al., 2011). These traits have consistently been linked, both conceptually and empirically, to the Agreeableness domain. For example, in their attempt to identify the hierarchical structure of Agreeableness,

---

<sup>2</sup>Relatedly, the ongoing use of the term “dark” to describe individuals’ personalities not only obscures similarities between (low) Agreeableness and D, but also strikes us as unscientific and potentially stigmatizing given the clinical implications surrounding these traits. To be clear, the “dark” terminology was not coined by these authors. Nonetheless, the terminology persists and has even expanded (e.g., Kaufman et al., 2019; Webster & Wongsomboon, 2020). Future work would likely benefit from exploring alternative ways to frame these socially aversive traits.

Crowe, Lynam, & Miller (2018) subjected 104 items from 22 measures of Agreeableness and related traits (e.g., Honesty Humility) to a bass-ackwards factor analysis. At the final level of the hierarchy, both a Morality (vs. Immorality) and Modesty (vs. Arrogance) factor were identified suggesting that across existing measures of Agreeableness, such content is present consistently enough for these factors to emerge. This structure was replicated by Vize et al. (2020) using the 104 items from Crowe et al. and the 70 items from the D-70 scale. More recently, similar factors emerged in a bass-ackwards analysis of commonly used Antagonism items from pathological trait measures (Sleep et al., in press), again speaking to their ubiquity across models and assessments of Agreeableness/Antagonism.

In order to adequately test the distinctiveness of Agreeableness and D, it is essential that both constructs are assessed comprehensively so that they are on equal footing. Using the same 104 items identified in Crowe et al.'s (2018) explication of Agreeableness, Vize et al. (2020) aimed to ensure adequate coverage of the domain. Using a preregistered approach (<https://osf.io/gj5bp>), Vize et al. (2020) found that latent Agreeableness was more strongly related to latent D (latent  $r = -.90$ ) compared to the latent correlation ( $r = -.64$ ) in Moshagen et al. (2020) and no specific subfactors composed of solely Agreeableness or D items emerged from a joint bass-ackwards factor analysis. However, the profile (dis)similarity of D and Agreeableness was comparable in both studies ( $r_{ICC} = -.99$  in Vize et al., 2020;  $r_{ICC} = -.95$  in Moshagen et al., 2020). The similarity in the  $r_{ICC}$ s is noteworthy, since the studies utilized different measures to assess the empirical profiles of D and Agreeableness.<sup>3</sup> When D and Agreeableness are comprehensively assessed, the evidence does not support the view that the constructs are different from one another.

### Agreeableness in the Broader Context of the Five-factor Model

Moshagen and colleagues (2020) suggest that the latent Agreeableness factor identified in Vize et al. (2020) is too broad and does not resemble Agreeableness as conceptualized by the Five-factor Model (FFM), where it is understood to be orthogonal to the remaining FFM dimensions. Using the available data from Vize et al., (2020), Moshagen et al. (2020) report a high degree of overlap between Agreeableness and Neuroticism ( $r = -.35$ ), Extraversion ( $r = .28$ ), Openness ( $r = .42$ ), and Conscientiousness ( $r = .57$ ).

First, it is noteworthy that Moshagen et al. (2020) do not provide any evidence that their latent Agreeableness factor possesses this orthogonality. Second, we believe there is little reason to expect that the five factors are actually orthogonal. This is not how personality traits behave. FFM domain scales have consistently been shown to be interrelated. Meta-analyses of the FFM scales have shown that across measurement approaches and sample type, the FFM domains are not orthogonal. In their large meta-analysis on the FFM ( $N = 144,117$ ;  $k = 212$ ), van der Linden et al. (2010) found that Agreeableness correlated .14 with Openness (O), .31 with Conscientiousness (C), .18 with Extraversion (E), and  $-.26$  with Neuroticism (N).<sup>4</sup> More recently, Park et al. (2020) also found a notable degree of

<sup>3</sup>In Moshagen et al. (2020), the 7 outcomes examined were behavioral dishonesty, competitive and dangerous worldviews, empathy, guilt proneness, internet trolling, and a sexual behavior outcome. In Vize et al., (2020), the outcomes included all 30 facets of the IPIP-NEO, various Agreeableness and "dark" personality scales, measures of reactive and proactive aggression, and substance use scales.

<sup>4</sup>The respective corrected correlations (corrected for measurement error and range restriction) were .21, .43, .26, and  $-.36$ .

interrelation among the FFM domains (median  $r = .28$ ). Agreeableness was positively related to Openness ( $r = .19$ ), Conscientiousness ( $r = .29$ ), and Extraversion ( $r = .20$ ) but was more weakly related to Neuroticism (reported as Emotional Stability;  $r = .16$ ).<sup>5</sup> In fact, some researchers choose to work at different personality hierarchies (i.e., 2 factor models: alpha vs. beta; three-factor models where A and C are blended) because of the natural attraction these dimensions (N, A, and C vs. E and O) have for one another (Digman, 1997, Markon et al., 2005; DeYoung, 2006).

In their integration of Big Five (B5) and circumplex approaches to personality, Hofstee, de Raad, and Goldberg (1992) provide compelling evidence that most personality items represent blends of multiple factors. Of 456 terms that loaded at least 0.20 on one of the five factors, only 67 terms (15%) could be taken as pure factor markers (i.e., uncorrelated with any of the other domains). For B5 Agreeableness, only 8 pairs of words could be considered pure markers of A. This narrow definition omits many terms that most personality psychologists would consider core to Agreeableness including soft-hearted, agreeable, obliging, trustful, pleasant, charitable, amiable, soft, affectionate, cooperative, considerate, and genial at the high end of A, and cold, cynical, insensitive, antagonistic, scornful, disagreeable, rude, egocentric, inconsiderate, uncharitable, ruthless, and callous at the low end of A.

So the claim by Moshagen et al. (2020) that an Agreeableness factor that, “carries substantial content of other FFM dimension...cannot be readily interpreted as one of few basic and largely orthogonal dimensions of personality as conceptualized in the FFM” (p. 8) seems unsupported. By this standard, nearly all existing FFM measures and 85% of personality trait terms are invalid. More generally, Moshagen et al.’s (2020) emphasis on the orthogonality of Agreeableness envisions a construct with which we, and the field in general, are unfamiliar.

In sum, there are issues surrounding the conceptual distinction between the D factor and (low) Agreeableness that remain unaddressed. Additionally, Moshagen et al.’s dismissal (2020) of evidence (i.e., the findings from Vize et al., (2020)) that runs contrary to their claims comes from an overly narrow and idiosyncratic view of Agreeableness that differs sharply from how the field currently conceptualizes and assesses the domain. Taken together, the claim that the D factor and Agreeableness are functionally different constructs does not appear tenable. In a literature that “jingles” and “jangles”<sup>6</sup> more than it should, great care should be taken before introducing an old concept with a new name into the literature.

## References

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

<sup>5</sup>The respective corrected correlations (corrected only for measurement error) were .28, .40, .28, and .22.

<sup>6</sup>The jingle and jangle fallacies are erroneous assumptions that two different constructs are the same because they have the same name (jingle) or that two virtually identical constructs are different because they have different names (jangle).

- Block J (2000). Three tasks for personality psychology In Bergman LR, Cairns RB, Nilsson L-G, & Nystedt L (Eds.), *Developmental science and the holistic approach* (pp. 155–164). Mahwah, NJ: Erlbaum.
- Cohen J, & Cohen P (1983). *Applied Multiple Regression/correlation Analysis for the Behavioral Sciences*. Hillsdale, New Jersey: Erlbaum.
- Costa PT, & McCrae RR (1992). *Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.
- Crowe ML, Lynam DR, & Miller JD (2018). Uncovering the structure of agreeableness from self-report measures. *Journal of Personality*, 86, 771–787. 10.1111/jopy.12358 [PubMed: 29072788]
- DeYoung CG (2006). Higher-order factors of the big five in a multi-informant sample. *Journal of Personality and Social Psychology*, 91, 1138–1151. [PubMed: 17144770]
- DeYoung CG, Quilty LC, & Peterson JB (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93, 880–896. [PubMed: 17983306]
- Digman JM (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 73, 1246–1256. [PubMed: 9418278]
- Furr RM, & Heuckeroth S (2019). The “Quantifying Construct Validity” procedure: Its role, value, interpretations, and computation. *Assessment*, 26, 555–566. [PubMed: 30614243]
- Goldberg LR (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26–42.
- Graziano WG, & Tobin M (2017). Agreeableness and the Five-Factor Model In Widiger TA (Ed.), *The Oxford Handbook of the Five Factor Model* (pp. 105–132). Oxford University Press.
- Hofstee WKB, & De Raad B. (1992). Integration of the Big Five and Circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63, 146–163. 10.1037/0022-3514.63.1.146 [PubMed: 1494982]
- John OP, & Srivastava S (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives In Pervin LA & John OP (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). New York: Guilford Press.
- Lee K, & Ashton MC (2018). Psychometric properties of the HEXACO-100. *Assessment*, 25, 543–556. [PubMed: 27411678]
- Maples JL, Guan L, Carter NT, & Miller JD (2014). A test of the international personality item pool representation of the revised NEO personality inventory and development of a 120-item IPIP-based measure of the five-factor model. *Psychological Assessment*, 26, 1070–1084. 10.1037/pas0000004 [PubMed: 24932643]
- Markon KE, Krueger RF, & Watson D (2005). Delineating the structure of normal and abnormal personality: an integrative hierarchical approach. *Journal of Personality and Social Psychology*, 88, 139–157. [PubMed: 15631580]
- Miller JD, Gaughan ET, Maples J, & Price J (2011). A comparison of agreeableness scores from the Big Five Inventory and the NEO PI-R: Consequences for the study of narcissism and psychopathy. *Assessment*, 18, 335–339. [PubMed: 21665883]
- Miller JD, Vize C, Crowe ML, & Lynam DR (2019). A Critical Appraisal of the Dark-Triad Literature and Suggestions for Moving Forward. *Current Directions in Psychological Science*, 28, 353–360. 10.1177/0963721419838233
- Moshagen M, Hilbig BE, & Zettler I (2018). The dark core of personality. *Psychological Review*, 125, 656–688. 10.1037/rev0000111 [PubMed: 29999338]
- Moshagen M, Zettler I, Horsten LK, & Hilbig BE (2020). Agreeableness and the common core of dark traits are functionally different constructs. *Journal of Research in Personality*, 87, 103986 10.1016/j.jrp.2020.103986
- Park HSH, Wiernik BM, Oh IS, Gonzalez-Mulé E, Ones DS, & Lee Y (2020). Meta-Analytic Five-Factor Model Personality Intercorrelations: Eeny, Meeny, Miney, Moe, How, Which, Why, and Where to Go. *Journal of Applied Psychology*, (Advance Online Publication). 10.1037/apl0000476
- Rodgers J, & Nicewander WA (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42, 59–66.
- Sleep CE, Crowe M, Carter NT, Lynam DR, & Miller JD (in press). Uncovering the structure of antagonism. *Personality Disorders: Theory, Research, and Treatment*.

- Soto CJ, & John OP (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113, 117–143. [PubMed: 27055049]
- van der Linden D, te Nijenhuis J, & Bakker AB (2010). The General Factor of Personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, 44(3), 315–327. 10.1016/j.jrp.2010.03.003
- Vize CE, Miller JD, & Lynam DR (2020). Examining the conceptual and empirical distinctiveness of Agreeableness and “Dark” personality items. *Journal of Personality*. Advance online publication. 10.1111/jopy.12601
- Westen D, & Rosenthal R (2003). Quantifying construct validity: two simple measures. *Journal of Personality and Social Psychology*, 84, 608–618. [PubMed: 12635920]