Traits and Distorted Symptom Presentation: a Scoping Review

Daniël van Helvoort[1,2] · Harald Merckelbach[1] · Chijs van Nieuwenhuizen[2,3] · Henry Otgaar[1,4]

Abstract
Are personality traits related to symptom overreporting and/or symptom underreporting? With this question in mind, we evaluated studies from 1979 to 2020 ($k = 55$), in which personality traits were linked to scores on stand-alone validity tests, including symptom validity tests (SVTs) and measures of socially desirable responding (SDR) and/or supernormality. As to symptom overreporting ($k = 14$), associations with depression, alexithymia, apathy, dissociation, and fantasy proneness varied widely from weak to strong (rs .27 to .79). For underreporting ($k = 41$), inconsistent links (rs − .43 to .63) were found with narcissism, whereas alexithymia and dissociation were often associated with lower SDR tendencies, although effect sizes were small. Taken together, the extant literature mainly consists of cross-sectional studies on single traits and contexts, mostly offering weak correlations that do not necessarily reflect causation. What this field lacks is an overarching theory relating traits to symptom reporting. Longitudinal studies involving a broad range of traits, samples, and incentives would be informative. Until such studies have been done, traits are best viewed as modest concomitants of symptom distortion.

Keywords Symptom validity · Trait · Personality · Overreporting · Underreporting

To varying degrees, people tend to report their symptoms inaccurately, i.e., report more or fewer symptoms than they actually have. When such inaccuracies become large, there is cause for concern. A heavily distorted symptom presentation may compromise diagnoses and treatment plans and may bias referral letters to clinicians or expert witness opinions. Still, self-reported symptoms are often the primary source of patient information in clinical work, or in professional reviews of legal documents or archival material (Carneiro et al., 2019; Rosen, 2006; Waite & Geddes, 2006; Wisdom et al., 2014). Therefore, it is important to consider the possibility that the patient exaggerates or minimizes symptoms. Validity tests have been constructed to aid psychologists in detecting symptom distortions and should ideally be part of every test battery (e.g., Bush et al., 2005; Chafetz et al., 2015; Heilbronner et al., 2009; Institute of Medicine, 2015; Sweet & Guidotti Breting, 2013).

Some validity tests aim to identify endorsement of bizarre symptoms or overendorsement of common symptoms (i.e., symptom overreporting) on self-report or rating measures of symptoms. These validity tests are referred to as *symptom* validity tests (SVTs). Other validity tests intend to detect underreporting, and include those that index supernormality (Cima et al., 2003, 2008) or socially desirable responding (i.e., SDR; e.g., Paulhus, 1988). Supernormality refers to denial of common, "everyday" symptoms, whereas SDR concerns denial of deviations from the social norm, which is arguably broader than just denial of symptoms. Supernormality and (blatant manifestations of) SDR are both seen as manifestations of "faking good" behavior (e.g., Bensch et al., 2019a, b; De Page & Merckelbach, 2021). Yet, although supernormality and SDR are conceptual cousins, the extant literature on supernormality is scarce, whereas the literature on SDR is abundant. For the purpose of the present review, we considered both stand-alone measures of supernormality and SDR as measures of one particular type of symptom distortion, namely underreporting.

✉ Daniël van Helvoort
   d.vanhelvoort@maastrichtuniversity.nl

[1] Clinical Psychological Science, Maastricht University, Maastricht, the Netherlands

[2] GGzE Institute for Mental Health Care, Eindhoven, the Netherlands

[3] Tranzo, Scientific Center for Care and Wellbeing, Tilburg University, Tilburg, the Netherlands

[4] Leuven Institute of Criminology, Catholic University of Leuven, Leuven, Belgium

Stand-alone validity tests are entirely or primarily dedicated to detecting non-credible presentations of symptoms or impairments, whereas embedded validity scales are part of broader measures that serve a clinical purpose, but contain a validity check (e.g., on over- or underreporting). Stand-alone tests are generally considered to be superior to embedded scales in terms of discriminant validity and are less often associated with interpretational problems (see, for a discussion: Erdodi & Lichtenstein, 2017). With these considerations in mind, we narrowed the focus of our scoping review down to stand-alone validity tests, but where relevant, we will discuss converging evidence from embedded measures, but also performance validity tests (PVTs) that tap into an exaggerated presentation of cognitive impairments.

Validity tests remain silent as to "why" patients over- or underreport symptoms (Bass & Wade, 2019). Symptom overreporting is often interpreted as a sign of malingering (i.e., intentional overreporting motivated by external incentives; Martin et al., 2015; Thompson et al., 2018). The *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; American Psychiatric Association [APA], 2013) has a limited perspective on this matter. As is true for the *DSM-IV-TR*, the *DSM-5* stresses malingering as an antecedent of symptom distortions, although its entries for histrionic personality disorder, and factitious disorder also allude to overreporting. In doing so, the *DSM-5* fails to recognize more articulated views on symptom distortions (e.g., see Berry & Nelson, 2010; Niesten et al., 2015; Otto, 2008; Rogers, 1990). Malingering is just one pathway to symptom overreporting and there might exist other pathways, such as acquiescence, careless or inattentive responding, demand characteristics, and misinformation effects (Merckelbach et al., 2019). Moreover, while the *DSM-5* briefly mentions anosognosia (i.e., lack of illness awareness), its one-sided focus on overreporting foregoes — and distracts from — the phenomenon of underreporting.

Clarifying the antecedents of symptom over- and underreporting — other than malingering or faking good, respectively — is important for the following reason: a narrow view on overreporting and underreporting would exclusively link these phenomena to incentives (e.g., compensation money; a favorable outcome in a child custody dispute) and a conceptual framing in terms of deliberate attempts to deceive others. Alternatively, some authors have argued that certain personality traits may foster a distorted symptom presentation. Ideas about the link between traits and symptom distortions can be traced back to the early papers of Eysenck, who speculated that psychoticism, antisocial features, and perhaps extraversion might contribute to overreporting (e.g., Eysenck & Gudjonsson, 1989; but see Young et al., 2016).

What is the empirical support for the idea that certain personality traits are associated with distorted symptom presentation? The current review addresses this question. If traits serve as powerful drivers of overreporting and underreporting, our conceptualization of distorted symptom presentation would have to be broadened to include, for example, the extent to which a hyperbolic style or a defensive attitude of the patient fuels such symptom distortions. To the extent that there are replicable, solid, and substantial links between personality traits and symptom distortion, such findings would provide important caveats and clues for clinicians who are faced with the task to interpret overreporting or underreporting on validity tests.

What is well established is that incentives play a prominent role in distorted symptom presentation. That is, studies consistently find higher failure rates on validity tests with increasing context-related incentives. For example, Mittenberg et al. (2002) estimated the prevalence of symptom exaggeration to hover around 10% in low-stake contexts such as regular medical cases, but up to 30% in high-stake contexts such as in forensic and civil legal cases (see also Young, 2015). One could argue that the non-trivial rate of $\pm 10\%$ of invalid responding in low-stake contexts (Dandachi-Fitzgerald et al., 2011, 2017; Mittenberg et al., 2002) reflects the contribution of traits to symptom distortions. Hence, we reviewed the literature to determine whether there is solid evidence linking certain traits to distorted symptom presentation, specifically in low-stake contexts.

In what follows, we will provide a scoping review on empirical and conceptual links between traits or personality pathology/disorders, on the one hand, and raised scores on stand-alone validity tests on the other hand. In doing so, we took the *DSM-5* section III: alternative DSM-5 model for personality disorders (AMPD) as a starting point. Importantly, elevated scores on validity measures do not necessarily imply that cut points have been crossed and that research subjects fail a validity test. Thus, our review is not restricted to studies that employed a categorical interpretation (i.e., pass versus failure) of validity test outcomes.

Before turning to the results of our review, there is a preliminary issue that deserves consideration. This issue has to do with the interpretational problems that may arise when validity test scores are linked to self-report indices of traits. Heightened scores on validity measures cast doubts on the accuracy of self-reported traits, thereby restricting the confidence that one can place in any interpretation of the data (see Merten et al., 2007). Despite this challenge to our review, we anticipate that some consistent patterns may emerge. First, the majority of articles included in our review focused on self-reported traits and more subtle distortions on validity tests (i.e., elevated scores). In such samples, typically only a minority of persons obtain scores beyond categorical

cut-offs that serve as a red flag for obvious symptom distortion (i.e., validity test failure), and that would raise concerns of relevant distortion of self-reported traits. Second, an often overlooked aspect of validity tests is that in many of these measures, the items mostly allude to bizarre symptoms (e.g., in typical SVTs), or mundane experiences (as in measures of supernormality): the type of items that some people might experience as provocative, threatening, ridiculous, or boring. It might well be that such a reactive attitude (i.e., "the negative subject"; Christensen, 1977; Miron & Brehm, 2006) is not only more pronounced in persons with certain traits (e.g., neuroticism, suspiciousness, narcissism), but also in itself fosters overreporting or underreporting of symptoms. Viewed this way, it is perfectly legitimate to the raise the question whether certain traits are related to symptom overreporting or underreporting on self-report validity measures.

Method

We conducted a scoping review (Grant & Booth, 2009) to investigate overall results, trends, limitations, and gaps in the literature on traits and distorted symptom presentation. We opted for this qualitative approach, because we had reason to assume that the extant literature would be highly diverse in terms of samples, measures, and incentives. For example, both Niesten et al. (2015) and Van Impelen et al. (2017) concluded in their reviews on antisocial features and distorted symptom presentation that the links between these constructs were highly variable and context dependent. Thus, we assumed that such diverse, perhaps even incoherent patterns might also hold true for the links between other traits and distorted symptom presentations.

On May 5, 2020, the first author (DH) ran a search in APA PsycINFO (EBSCO) and PubMed to identify empirical studies that directly tapped into measures of overreporting and/or underreporting, as well as diagnoses or measures of *DSM-5* section II or III personality disorders or traits. A Boolean title/abstract search of the format "X AND Y" was used; alternative terms were specified with the OR function. More specifically, the search comprised of typical response bias equivalents, such as *(dis)simulation*, *distortion*, *response style*, *symptom validity*, *over-/underreporting*, *malingering*, *exaggeration*, *feigning*, *faking*, *desirable responding*, *self-deception/-deceptive enhancement*, *impression management*, and/or commonly used "stand-alone validity tests" such as the *Structured Interview of Reported Symptoms*, *Structured Inventory of Malingered Symptomatology*, *Miller Forensic Assessment of Symptoms Test*, *Self-Report Symptom Inventory*, *Supernormality Scale(-Revised)*, *Marlowe-Crowne Social Desirability Scale*, *and Balanced Inventory of Desirable Responding*, and their acronyms (see for an introduction to

a special issue on self-report symptom validity measures, Giromini et al., 2022). These terms were combined with variations of "trait" and "personality disorder" terms, such as the ten *DSM-5* section II categorical personality disorders (PD, i.e., *schizoid*, *schizotypal*, *paranoid*, *antisocial*, *borderline*, *narcissistic*, *avoidant*, *dependent*, and *obsessive–compulsive PD*), the 25 *DSM-5* section III AMPD *trait facets* and their variations, and other conceptually similar or related traits and terms known to the authors: *alexithymia*, *dissociation*, *fantasy proneness*, *narcissism*, *hysteria*, *paranoia*, and *psychopathy*.

The search was limited to studies in English peer-reviewed journals between 1979, the year that the oldest measure in relevant articles — the *Narcissistic Personality Inventory* (NPI; Raskin & Hall, 1979) — was published, and May 5, 2020, the date of search. All retrieved records were screened such that those records with abstracts that covered our search terms were selected for full-text retrieval and closer inspection. Studies that were inaccessible (e.g., unpublished dissertations, conference papers) were excluded, as were non-empirical articles or chapters. Next, full-text records were coded on relevant parameters (i.e., sample and setting, trait/disorder, relevant measures, and outcomes comparisons). Studies that examined whether antisocial personality disorder and/or psychopathy are linked to heightened scores or failures on stand-alone (and/or embedded) validity tests were excluded because these studies have been extensively summarized and discussed in previous reviews (e.g., Niesten et al., 2015; Van Impelen et al., 2017).

The scoping review is divided into a section on overreporting and a section on underreporting. We focus only on those traits that have been linked to symptom distortion in the literature identified by our scoping review, or by indirect evidence in articles known to us. For each section, we cluster traits based on their conceptual similarity. For each discussed trait, our evaluation of the evidence — where available — is structured in the following way: (a) conceptual definitions; (b) direct empirical evidence provided by studies that emerged from our scoping review, linking the trait to validity test outcomes. In evaluating the strength of correlational data, we followed the benchmarks of Cohen (1988) and considered $rs \geq 0.50$ as strong, rs from 0.30 to 0.50 as moderate, and rs from 0.10 to 0.30 as weak; (c) indirect empirical evidence from literature known to the authors, involving conceptually similar traits or disorders, embedded validity scales, or PVTs as indicators of impairment exaggeration; (d) theoretical or conceptual articles that provide tentative explanations for such links based on empirical evidence; and (e) prima facie tentative plausibility of links to symptom distortion, based on core features of the respective traits (e.g., cognitions, emotions, behavior).

Results

Our scoping review identified 52 articles ($k=55$ studies) that linked self-reported traits to outcomes on stand-alone validity tests (i.e., stand-alone SVTs, or measures of SDR and/or supernormality). As to overreporting, 12 articles encompassing 15 studies were found addressing depression/depressivity ($k=3$), alexithymia ($k=2$), apathy ($k=1$), dissociation/cognitive and perceptual dysregulation ($k=7$), and/or fantasy proneness/unusual beliefs and experiences ($k=2$). As to underreporting, 40 articles encompassing 41 studies were found addressing alexithymia ($k=11$), dissociation ($k=7$), and/or narcissism ($k=23$).

Overreporting

Studies in this domain have mainly focused on correlations between self-report measures of traits, for example, the *Dissociative Experiences Scale* (DES; Bernstein & Putnam, 1986) in the case of dissociativity, and SVTs consisting of a self-report checklist of bizarre symptoms as a measure of symptom overreporting, such as the *Structured Inventory of Malingered Symptomatology* (SIMS: Smith & Burger,

1997). Some researchers also relied on a group-comparison approach, typically based on scores above and below validity test cut-offs. Table 1 presents an overview of correlational studies on traits and overreporting; Table 2 summarizes group-comparison studies. Note that these and following tables only include studies that relied on stand-alone (rather than embedded) validity tests.

Depressivity

Trait depressivity refers to pervasive and/or persistent feelings of being down, miserable, or hopeless; feelings of inferior self-worth, shame or guilt; and/or pessimism or suicidal thoughts or behavior (APA, 2013; Krueger et al., 2012). Our screening of the literature did not find relevant articles on trait depressivity per se, but did retrieve (a non-exhaustive list of) studies on depression and symptom distortions, on which we report below.

Two articles (i.e., Merckelbach & Smith, 2003; Merten et al., 2019a) directly linked self-reported depression on the *Beck Depression Inventory* (BDI; Beck et al., 1961) to overreporting on SVTs and noted strong associations: Merckelbach and Smith (2003) found in undergraduates ($N=182$) that BDI scores correlated with SIMS scores.

Table 1 Traits and overreporting on stand-alone SVTs: correlational studies

Trait	Study	Sample	N	Trait measure	SVT	$r(s)$	Link yes/no
Depression ($k=2$)	Merckelbach and Smith (2003)	Undergraduates ($N=182$)	182	BDI	SIMS	.64**	Yes
	Merten et al. (2019a)	Psychosomatic inpatients (Compensation-seekers)	537	BDI-II	SIMS;	.71***	Yes
					SRSI pseudo-symptoms	.64***	Yes
Alexithymia ($k=2$)	Brady et al. (2017)	Veterans with PTSD	75	TAS-20	M-FAST	.49**	Yes
	Merckelbach et al. (2018)	Forensic outpatients	40	TAS-20	SIMS	.56**	Yes
		Non-forensic participants	40	TAS-20	SIMS	.51**	Yes
Apathy ($k=1$)	Dandachi-Fitzgerald et al. (2020)	Neurological patients	138	AES	SIMS	.27**	Yes
Dissociation ($k=4$)	Giesbrecht and Merckelbach (2006)	Students	87	DES	SIMS	.51**	Yes
	Kunst et al. (2011)	Violence victims	125	PDEQ-10	SIMS	.41***	Yes
	Merckelbach et al. (2015)	Students	113	CDS	SIMS	.45**	Yes
		Inpatients	21	CDS DES DES-T	SIMS	.32 .25 .28	No
	Van der Heide and Merckelbach (2016)	Inpatients (Asylum seekers)	27	DES	SIMS MENT	.79*** .41*	Yes Yes
Fantasy proneness ($k=2$)	Merckelbach and Smith (2003)	Undergraduates	182	CEQ	SIMS	.33**	Yes
	Kunst et al. (2011)	Violence victims	125	CEQ	SIMS	.40***	Yes

AES, Apathy Evaluation Scale; *BDI*, Beck Depression Index; *CEQ*, Creative Experiences Questionnaire; *DES*, Dissociative Experiences Scale; *DES-T*, DES-Taxon; *MENT*, Morel Emotional Numbing Test; *M-FAST*, Miller Forensic Assessment of Symptoms Test; *PDEQ-10*, Peritraumatic Dissociative Experiences Questionnaire-10 Self-Report Version; *SIMS*, Structured Inventory of Malingered Symptomatology; *SRSI*, Self-Report Symptom Inventory; *SVT*, symptom validity test; *TAS-20*, Twenty-item Toronto Alexithymia Scale

*$p<.05$; **$p<.01$; ***$p<.001$

Table 2 Traits and overreporting on stand-alone SVTs: group-comparison studies

Trait	Study	Sample	N	Trait measure	SVT	Results	Link yes/no
Depression ($k=1$)	Merten et al. (2019a)	Psychosomatic inpatients (compensation-seekers)	537	BDI-II	SIMS SRSI pseudo-symptoms	A link between elevated BDI-II scores (i.e., > 40) and SVT failure ($X^2=89.6$, $df=3$; $phi=.41$; $p<.001$) for both SVTs	Yes
Dissociation ($k=3$)	Brand et al. (2006)	DID patients	20	SCID-D-R clinician-rated	SIRS	35% of the DID patients obtained SIRS scores suspicious of overreporting	Yes
	Rogers et al. (2009)	Traumatized patients (most of whom exhibited severe dissociative symptoms)	37	SCID TSI DES DDIS-SRV	SIRS	31% of the PTSD patients obtained SIRS scores suspicious of overreporting	Yes
	Constans et al. (2014)	Veterans	124	DES	M-FAST	Veterans who exceeded the cut point on the M-FAST (overreporting; $n=20$) attained higher DES scores than those who did not (non-overreporting; $n=44$ and 60) at $p<.001$	Yes

BDI-II, Beck Depression Inventory-II; DDIS-SRV, Dissociative Disorders Interview Schedule-Self-Report Version; DES, Dissociative Experiences Scale; DID, Dissociative Identity Disorder; M-FAST, Miller Forensic Assessment of Symptoms Test; PTSD, posttraumatic stress disorder; SCID, Structured Clinical Interview for DSM-IV; SCID-D-R, SCID-Dissociative Disorders-Revised; SIMS, Structured Interview of Malingered Symptomatology; SIRS, Structured Interview of Reported Symptoms; SRSI, Self-Report Symptom Inventory; SVT, symptom validity test; TSI, Trauma Symptom Inventory

Yet, differences in depression scores were not evident when comparisons were based on groups attaining SIMS scores below or above the cut-off. Relying on compensation-seeking psychosomatic rehabilitation inpatients ($N = 537$), Merten et al. (2019b) found scores on the SIMS and the *Self-Report Symptom Inventory* (SRSI) — a recently developed measure of overreporting (Merten et al., 2016) — to be strongly related to BDI-II scores (Beck et al., 1996). Specifically, compensation-seeking psychosomatic inpatients with elevated BDI-II scores (i.e., > 40) more often failed on the SVTs than those who attained lower BDI-II scores.

Some studies focused on embedded validity scales or PVTs as indicators of impairment exaggeration. For example, a meta-analysis found that genuine major depressive disorder was linked to validity scale failures (i.e., F-r, FBS-r, and RBS; $Ms > 80$ T; Sharf et al., 2017) on the *Minnesota Multiphasic Personality Inventory-2-Restructured Form* (MMPI-2-RF; Ben-Porath & Tellegen, 2008). However, studies that relied on PVTs in inpatients (Rees et al., 2001), community-dwelling elders (Ashendorf et al., 2004), disability-seeking outpatients (Yanez et al., 2006), and claimants in personal injury litigation (Stevens et al., 2008) did not find associations between underperformance and BDI(-II) scores.

Indirect support for a link between depressivity and symptom distortion comes from studies showing that people who endorse high levels of negative emotions tend to report more psychological and physical symptoms, regardless of the actual severity of their condition (Costa & McCrae, 1987; but see Friedman et al., 2010). Along similar lines, studies observed that some patients diagnosed with a mood disorder tend to attain higher symptom ratings on the BDI than their clinicians on the observer-rated *Hamilton Depression Rating Scale* (HDRS; Hamilton, 1980). For example, Stanley and Wilson (2006) observed this discrepancy in patients with major depressive disorder and comorbid borderline personality disorder, but not in those without this comorbidity. Duberstein and Heisel (2007) noted that endorsement of high neuroticism levels is associated with symptom overreporting, whereas other researchers contend that it is self-reported depressive symptoms rather than general neuroticism that is related to inflated reports of *physical* symptoms (Howren et al., 2009).

Suls and Howren (2012) provided a theoretical rationale for this relationship. According to these authors, depression is accompanied by a negative recall bias (e.g., a heightened accessibility of negative memories), which would lead to inflated self-reports of past symptoms, whereas anxiety would be associated with an attentional bias that escalates self-reports of momentary symptoms (see also Robinson & Clore, 2002).

Alexithymia, Restricted Affectivity, Anhedonia, and Apathy

Alexithymia refers to the inability to identify and articulate internal sensations, coupled with an external orientation (i.e., deficits in emotion processing). A common self-report measure of alexithymia is the *Toronto Alexithymia Scale-20* (TAS-20; Bagby et al., 1994). Two studies addressed TAS-20 scores and symptom overreporting on stand-alone SVTs: Brady et al. (2017) found in US veterans diagnosed with posttraumatic stress disorder (PTSD; $N = 75$) that TAS-20 scores were moderately correlated with failure on the *Miller Forensic Assessment of Symptoms Test* (M-FAST; Miller, 2001). Similarly, Merckelbach et al. (2018) noted in forensic outpatients ($n = 40$) and non-forensic participants ($n = 40$) strong associations between TAS-20 and SIMS scores.

Indirect support for a link between alexithymia and symptom distortion is provided by research that related poor interoceptive accuracy to overreporting of somatic symptoms (e.g., Byrne & Ditto, 2005; Grynberg & Pollatos, 2015; Herbert et al., 2011), although one study did not observe such a relationship (Fairclough & Goodwin, 2007). De Gucht and Heiser (2003) reviewed 16 studies on alexithymia (indexed by TAS-20 or versions thereof) and symptom reporting and concluded that endorsement of alexithymia is significantly, albeit weakly ($r = 0.23$) associated with inflated self-reports of somatic symptoms. Similar findings were reported in more recent studies relying on patients (Porcelli et al., 2013) and undergraduate students (Bogaerts et al., 2015; Wearden et al., 2005). These studies converged on the preliminary hypothesis that in alexithymia, normal arousal and distress might be mislabeled as highly intense symptoms (Grynberg et al., 2012).

The AMPD traits restricted affectivity and anhedonia are conceptual neighbors of alexithymia (e.g., see Badura, 2003; Gooding & Tallent, 2003). Restricted affectivity is defined as constricted emotional experience and responsivity (i.e., emotional numbing), whereas anhedonia is a lack of enjoyment and engagement in experiences (APA, 2013; Krueger et al., 2012). Kashdan et al. (2007) investigated whether self-reported emotional numbing and anhedonia as indexed by the MMPI-2 are linked to symptom distortion, using the Fp scale as an embedded indicator of overreporting (Butcher et al., 2001). They observed that PTSD veterans who overreported symptoms (Fp > 8) on the MMPI-2 ($n = 30$) scored higher on MMPI-2 indices of emotional numbing and anhedonia than those who did not engage in overreporting ($n = 197$), with effect sizes being small to moderate. Dandachi-Fitzgerald et al. (2020) noted in a sample of neurological patients ($N = 138$) that apathy as measured by the self-report *Apathy Evaluation Scale* (AES; Marin et al., 1991) was associated with heightened scores on the SIMS and lowered performance on a PVT (i.e., the

Test of Malingered Memory; TOMM; Tombaugh, 1997; $r = -0.31$), albeit it that effect sizes were small to moderate.

Dissociativity

Dissociative symptoms constitute a heterogeneous class of experiences that may range from minor cognitive lapses such as daydreaming, to disabling symptoms such as derealization and depersonalization (Condon & Lynn, 2014; Giesbrecht et al., 2008). In the *DSM-5* AMPD, the conceptual twin of dissociation is trait cognitive and perceptual dysregulation, defined as odd or unusual thought processes including dissociative experiences, mixed sleep–wake states, and/or thought-control experiences (APA, 2013; Krueger et al., 2012). Typically, dissociativity is assessed with self-report measures, mostly the DES, although other measures exist (Merckelbach et al., 2017). A recent meta-analysis shows a link between self-reported dissociativity and alexithymia (e.g., $r = 0.56$ in clinical populations; Reyno et al., 2020). DES scores are particularly related to the TAS-20 subscale "difficulty in identifying feelings" (e.g., rs = 0.46–0.52; Elzinga et al., 2002; Evren et al., 2008).

Merckelbach et al. (2017) reviewed studies in which dissociative symptoms or diagnoses (e.g., dissociative identity disorder [DID]) were associated with overreporting on embedded validity scales and/or stand-alone SVTs. These authors identified six relevant correlational studies (four included stand-alone SVTs) and ten group-comparison studies (i.e., based on test cut-offs; four included stand-alone SVTs), pertaining to various student and/or patient samples (e.g., PTSD; DID). The correlational studies that involved stand-alone SVTs (i.e., Giesbrecht & Merckelbach, 2006; Kunst et al., 2011; Merckelbach et al., 2015; Van der Heide & Merckelbach, 2016) overall found moderate to strong links between endorsement of dissociativity and overreporting. A similar picture emerged from correlational studies that involved embedded validity tests or PVTs). Likewise, a group comparison study in a veteran sample ($N = 124$; Constans et al., 2014) found that those who failed stand-alone SVTs all were diagnosed with PTSD and scored significantly higher on a self-report measure of dissociation. Other group comparison studies including SVTs or validity scales in small patient samples (PTSD and/or DID, Ns = 19–37, i.e., Brand et al., 2006; Rogers et al., 2009, 2011) found that sizeable proportions (i.e., 25–37%) of patients — especially those with DID — failed on such validity tests.

From a theoretical stance, a potential pathway from depressivity through recall bias and then symptom distortion possesses a certain prima facie plausibility. Much the same is true for alexithymia, poor interoceptive monitoring, and escalating symptom reports. We did not come across any theoretical rationale for how the broad and diverse category of dissociativity might contribute to distorted symptom presentation, except the notion that this overlap is the product of two prominent concomitants of dissociation, namely alexithymia (cf. supra) and fantasy proneness (Merckelbach et al., 2017), to which we will turn now.

Fantasy Proneness and Cognitive Dysregulation

Fantasy proneness refers to a strong preference for vivid imagery and make-believe experiences and activities (Elzinga et al., 2002; Merckelbach et al., 1999; Rauschenberger & Lynn, 1995). A common self-report measure of fantasy proneness is the *Creative Experiences Questionnaire* (CEQ; Merckelbach et al., 2001). A recent meta-analysis showed that self-reported dissociative symptoms and CEQ scores are strongly correlated ($r = 0.52$; Merckelbach et al., 2021). Fantasy proneness comes close to the *DSM-5* AMPD trait unusual beliefs and experiences, which is described as beliefs of unusual abilities (e.g., mind-reading, telekinesis), thought-action fusion, and unusual (e.g., hallucination-like) experiences of reality (APA, 2013; Krueger et al., 2012).

Only two studies have directly examined the link between CEQ and symptom overreporting on a dedicated, stand-alone SVT. One study found in a sample of compensation seeking victims of interpersonal violence ($N = 125$) that CEQ scores were moderately correlated with symptom overreporting on the SIMS and that CEQ scores negated the link between dissociation and PTSD symptoms in this sample (Kunst et al., 2011). Similarly, another study observed in undergraduates ($N = 182$) that CEQ scores were moderately associated with overreporting tendencies on the SIMS (Merckelbach & Smith, 2003).

Other evidence for an association between self-reported fantasy proneness and symptom distortion comes from an analogue simulation study in which undergraduate students ($N = 648$) were instructed to feign traumatic stress symptoms. Specifically, those who endorse high levels of fantasy proneness more often atypically responded on the *Trauma Symptom Inventory* (TSI; Briere, 1995) and overall had inflated symptom scores as compared with those low on self-reported fantasy proneness (Peace & Masliuk, 2011). Relatedly, Merckelbach and van de Ven (2001) found that endorsement of fantasy proneness predicts hallucinatory reports during white noise exposure in students.

From a theoretical perspective, an association between fantasy proneness and distorted symptom presentation makes sense. It might reflect fantasy prone people's preference for exploring the limits of reality and convention (i.e., counterfactual thinking; see Bacon et al., 2013), which engenders low thresholds when completing self-reports that include eccentric experiences and symptoms. Alternatively, the driving force might be the overlap between fantasy proneness and schizotypal features and the cognitive dysregulation implicated in that, which promotes inattentive or

careless responding when filling out self-report instruments that list symptoms (Merckelbach et al., 2017).

Studies on patients with schizophrenia spectrum disorders indirectly support a relationship between cognitive dysregulation and symptom overreporting. That is, such diagnoses in patients are associated with failure on SVTs (e.g., Van Impelen et al., 2014), and on embedded validity indicators or PVTs (e.g., Peters et al., 2013; Gorissen et al., 2005; Van der Heide & Merckelbach, 2016, Van der Heide et al., 2017, 2019; but see Schroeder & Marshall, 2011). Several authors have suggested that deficits in reality monitoring, illness insight, and cognitive functions may drive such failures on validity tests (Radaelli et al., 2013; Shad et al., 2006; Schaefer et al., 2013; but see Stevens et al., 2014).

Underreporting

Empirical studies in the domain of traits and underreporting have largely focused on correlations between self-report measures of traits, for example, the NPI and/or the *Pathological Narcissism Inventory* (PNI; Pincus et al., 2009) in the case of narcissism, and self-report measures of SDR, such as the *Marlowe-Crowne Social Desirability Scale* (MCSDS; Crowne & Marlowe, 1960) and the *Balanced Inventory of Desirable Responding* (BIDR; Paulhus, 1988). Both the MCDS and the BIDR do contain items that allude to symptoms (e.g., It's hard for me to shut off a disturbing thought) but most of their items gauge overly optimistic self-presentation in a number of domains (e.g., prosocial behavior, ambition). Table 3 presents an overview of such studies on traits and dedicated measures of underreporting, as identified by our scoping review.

Alexithymia

Eleven studies addressed potential links between self-reports of trait alexithymia and social desirability. Most were conducted in student or community samples and involved the TAS-20 and self-report measures of social desirability. A majority observed small to moderate correlations in the inverse direction — i.e., heightened alexithymia was linked to *lower* self-reported SDR — whereas some studies found no significant links. Of note is a study of Linden et al. (1996) in undergraduates ($N = 80$) that, unlike other studies pertaining to alexithymia, used a group-comparison (i.e., based on low, moderate, and high TAS-20 scores) rather than a correlational approach to the issue. The authors found no significant differences across these groups on the BIDR. The only study that examined the relationship between self-reported alexithymia and SDR in a patient group (i.e., in veterans with schizophrenia or schizoaffective disorder; $N = 65$) is also the only study to find *positive* (and moderate)

correlations with social desirability (i.e., as indexed by the MCSDS; Fogley et al., 2014).

Taken together, these studies found endorsement of alexithymia to be associated with *lower* self-reported social desirability and/or found no significant associations between these constructs. As to a possible theoretical rationale, Messina et al. (2010) pointed to an inverse association between TAS-20, particularly the "difficulty identifying feelings" subscale of the TAS-20, and the "Lie Scale" of the *Eysenck Personality Questionnaire* (EPQ; Eysenck & Eysenck, 1985). Interestingly, this link disappeared when controlling for self-reported neuroticism on the EPQ. Whereas previous authors suggested that individuals with relatively poorer emotional differentiation (i.e., alexithymia) might be less capable or interested in social desirability (particularly its self-deception aspect; Fukunishi, 1994), Messina et al. (2010) suggested that such a lack of interest might be more primarily driven by mild diffuse emotional distress (i.e., neuroticism).

Dissociativity

Relying on a correlational approach, seven studies investigated the connection between self-reported dissociation and social desirability. Most of them involved university (Ns = 28–633) or high school students ($N = 93$; Callahan et al., 2003). None of the studies found that higher endorsement of dissociativity is accompanied by heightened social desirability scores (Beere et al., 1996; Evans et al., 2019; Hyman & Billings, 1998) and some even noted an inverse link, with self-reported dissociativity being related to *lower* social desirability (Callahan et al., 2003; Elzinga et al., 2002). The one exception to this pattern is a study pertaining to mothers of children aged 6–11 ($N = 93$) who were referred to Youth Protection Services (Collin-Vézina et al., 2005). Here, endorsement of dissociation was *positively* and moderately correlated to self-reported social desirability. Only few ($n = 6$) respondents achieved DES scores above the clinical cut-off. The authors suggested that overall, participants might have underreported on the DES given that they were interviewed following substantiated charges of child abuse or neglect.

Grandiosity

People high on trait grandiosity are self-centered, condescend, and tend to experience feelings of superiority and entitlement, often ascribed to self-presumed outstanding qualities and/or achievements (APA, 2013; Krueger et al., 2012). Grandiosity is one of two distinct dimensions of narcissism, the other being vulnerable narcissism, which includes hypersensitivity, introversion, shame, and inhibition

Table 3 Traits and underreporting on stand-alone validity measures

Trait	Study	Sample	N	Trait measure	SVT	r(s)	Link yes/no
Alexithymia (k = 11)	King et al. (1992)	Adults	155	TAS-20	MCSDS SDQ	−.20* −.10	Inverse No
	Fukunishi (1994)	College students	215	TAS-20	MCSDS	−.39**	Inverse
	Kroner and Forth (1995)	Males incarcerated for violence	508	TAS-20	BIDR	TAS-20 scales x BIDR scales: −.59 to −.06; **only rs < −.16	Inverse
	Linden et al. (1996)	Undergraduates	80	TAS-20	BIDR	Ns. difference between low (n=20), medium (n=30), and high (n=27) alexithymia groups in terms of BIDR scores	No
	Fischer and Good (1997)	Undergraduates	208	TAS-20	MCSDS-13	−.39**	Inverse
	Waldstein et al. (2002)	Adults (aged 53–83)	102	TAS-26	MCSDS	.03	No
	Bouchard (2008)	Adults	270	TAS-20	DS36	−.42***	Inverse
	Helmes et al. (2008)	Undergraduates (three samples)	219 99 153	TAS-20	BIDR	TAS-20×BIDR-IM: −.17 to −.19 TAS-20×BIDR-SDE: −.32* to −.38*	No Inverse
	Vellante et al. (2013)	Undergraduates	200	TAS-20	MCSDS	−.32**	Inverse
	Fogley et al. (2014)	Veterans with schizophrenia or schizoaffective disorder	65	TAS-20	MCSDS	TAS-20×MCSDS-IM: .53*** TAS-20×MCSDS-SD: .35**	Yes Yes
Dissociation (k = 7)	Lyvers et al. (2019)	Adult social media users	143	TAS-20	MCSDS	−.33***	Inverse
	Beere et al. (1996)	Undergraduates	136	DES DES-T	MCSDS	−.01 −.01	No No
	Hyman and Billings (1998)	Undergraduates	267	DES	MCSDS	−.04	No
	Elzinga et al. (2002)	Undergraduates	833	DIS-Q	"Two scales [IM; SD] that consist of 8 items"	DIS-Q×IM scale: −.31*** DIS-Q×SD scale: −.15***	Inverse Inverse
	Callahan et al. (2003)	High school students (aged 13–19)	526	TSCC-D	MCSDS 10-item version	−.27**	Inverse
	Collin-Vézina et al. (2005)	Mothers of children aged 6–11 referred to Youth Protection Services	93	DES	MC-C	.49*** (after exclusion of n = 5 who attained low DES and high MC-C scores)	Yes
	Otgaar et al. (2017)	University students	26	DES	MCSDS	"No statistically significant correlations emerged"	No
	Evans et al. (2019)	Undergraduates	628	DES	MCSDS	<.01	No

Table 3 (continued)

Trait	Study	Sample	N	Trait measure	SVT	r(s)	Link yes/no
Narcissism (k = 23)	Auerbach (1984)	Undergraduates	148	NPI RNPI MCMI-N	MCSDS	−.01 to .03	No
	Watson, Grisham, et al. (1984)	Undergraduates	160	NPI	MCSDS	NPI total: −.11 NPI-EE: −.43*** NPDS total: −.36***	No Inverse Inverse
				NPDS			
	Watson, Hood, et al. (1984)	Undergraduates	85	NPI	MCSDS	−.40***	Inverse
	Watson et al. (1986)	Undergraduates	145	NPI	MCSDS	NPI total: −.04 NPI-EE: −.16*	No Inverse
	Raskin et al. (1991)	Undergraduates (four samples)	60 84 30,057	NPI MNPS MNPDS	MCSDS, DSE, PD, ESDS, UV	NPI×IM composite: −.06 to .20	No
						Narcissism composite score×IM: −.28***	Inverse
						NPI×SDE composite: .30** to .56***	Yes
						Narcissism composite×SDE composite: .63***	Yes
	Watson and Morris (1991)	Undergraduates	221	NPI	MCSDS	NPI total: −.19* NPI-EE: −.27***	Inverse Inverse
	Hibbard (1992)	University students	701	NPI OMNI NPDS	MCSDS	−.08 −.36 −.36 *NB: ps not reported*	Inverse
	Fukunishi et al. (1995)	College students	215	NPI	MCSDS	NPI total: −.25**	Inverse
		Myocardial infarction patients	30	NPI	MCSDS	NPI total: −.38*	Inverse
	Paulhus (1998)	Students	124	NPI	BIDR	NPI×BIDR-SDE: .50 and .53	Yes
						NPI×BIDR-IM: .18 *NB: ps not reported*	Yes
	Sedikides et al. (2004)	Undergraduates	155	NPI	BIDR-IM	−.23**	Inverse
	Barelds and Dijkstra (2010)	Community and student sample	460,515	NPI	MCSDS	−.05	No
	Zeigler-Hill and Wallace (2011)	Undergraduates (online survey; three samples)	315,367,397	NPI-37 NPI NGS PNI HSNS OMNI	BIDR	Across studies and indices: −.31 to .23. NB: inverse correlates were often significant at p <.05, whereas positive correlates were largely ns	Inverse
	Braun et al. (2016)	Undergraduates	942	NPI	MCSDS	−.11**	Inverse

Table 3 (continued)

Trait	Study	Sample	N	Trait measure	SVT	r(s)	Link yes/no
	Jones et al. (2016)	Adults engaging in contact sports	272	NPI-16	SDS	−.25***	Inverse
	Barry et al. (2016)	Adolescent (aged 16–19) enrolled in a residential program	161	PNI	BIDR	PNI-GN×SDR: −.11 to .01	No
						PNI-VN×SDR: −.29***	Inverse
						PNI-VN×IM: −.19*	Inverse
						PNI-VN×SDE: −.28***	Inverse
	Fernie et al. (2016)	Community sample	170	PNI	MCSDS	PNI-GN×SDR: .22*	Yes
						PNI-VN×SDR: .46**	Yes
	Gamache et al. (2018)	Online sample (e.g., media and university email)	405	DD	BIDR-21	−.25	Inverse
				SD3		−.10	No
	Manley et al. (2018)	Online survey (via Mechanical Turk)	248	NPI	BIDR	NPI×BIDR-IM: −.18*	Inverse
						NPI×BIDR-SDE: .17*	Yes
				HSNS		HSNS×BIDR-IM: −.42**	Inverse
						HSNS×BIDR-SDE: −.47**	Inverse
	Yu (2018)	University counseling attendees without disorder	160	PNI	MCSDS	−.42***	Inverse
	Levi and Bachar (2019)	Israeli survivors of terrorism, or road/work accidents	152	NVS	BIDR-SDE	−.35**	Inverse
	Lyvers et al. (2019)	Adult social media users	143	NPI-16	MCSDS-13	−.07	No
	Wu et al. (2019)	Single Asian Americans	262	NPI-16	MCSDS-13	−.15*	Inverse
	De Page and Merckelbach (2021)	Adult patients	147	NPI	SS	.20*	Yes

BIDR, Balanced Inventory of Desirable Responding; *DD*, The Dirty Dozen; *DES*, Dissociative Experiences Scale; *DES-T*, DES-Taxon; *DIS-Q*, Dissociation Questionnaire; *DS36*, 36-item Social Desirability measure (Tournois et al., 2000); *DSE*, Defensive Self-Enhancement Scale; *ESDS*, Edwards Social Desirability Scale; *GN*, grandiose narcissism; *HSNS*, Hypersensitive Narcissism Scale; *IM*, impression management (scale); *MC-C*, MCSDS-Short Form C; *MCMI*, Millon Clinical Multiaxial Inventory-Narcissism Scale; *MCSDS*, Marlowe-Crowne Social Desirability Scale; *MCSDS-13*, MCSDS-13 item version; *MNPS*, Narcissistic Personality Scale; *MNPDS*, Narcissistic Personality Disorder Scale; *NGS*, Narcissistic Grandiosity Scale; *NPI*, Narcissistic Personality Inventory; *NPI-16*, NPI-16 item version; *NPI-37*, NPI-37 item version; *NPI-EE*, NPI-Exploitation/Entitlement Scale; *NVS*, Narcissism Vulnerability Scale; *OMNI*, O'Brien Multiphasic Narcissism Inventory; *PD*, positive defensiveness; *PNI*, Pathological Narcissism Inventory; *RNPI*, Revised (i.e., 40 items) NPI; *SD*, self-deception; *SD3*, Short Dark Triad; *SDE*, self-deceptive enhancement (scale); *SDR*, socially desirable responding; *SDS*, Social Desirability Scale; *SS*, Supernormality Scale; *SVT*, symptom validity test; *TAS-20*, Twenty-item Toronto Alexithymia Scale; *TSCC-D*, Trauma Checklist for Children-Dissociation Scale; *UV*, unlikely virtues; *VN*, vulnerable narcissism

*p < .05; **p < .01; ***p < .001

of grandiose desires (e.g., see Cain et al., 2008; Miller et al., 2011; Wink, 1991).

Twenty-three empirical studies examined links between stand-alone self-report measures of narcissism and SDR. These studies were all conducted in student and/or community samples, with the exception of studies (in part) in myocardial infarction patients ($n = 30$; Fukunishi et al., 1995), an incarcerated sample ($n = 703$; Sleep et al., 2017), terrorism or work/car accident survivors ($N = 152$; Levi & Bachar, 2019), and adult psychiatric patients ($N = 147$, De Page & Merckelbach, 2021). Below, we catalogued the findings by dividing these studies into two time periods: early articles published before 1998, which primarily examined links between the NPI and the MCSDS, and later articles that more often included several self-report measures of narcissism (e.g., grandiose and vulnerable dimensions; the PNI) and social desirability (e.g., self-deceptive enhancement [SDE] and impression management [IM]; the BIDR).

Overall, early studies relied on undergraduate samples (Ns = 85–221) and found little support for the idea that narcissistic features as self-reported on the NPI are linked to heightened social desirability as indexed on the MCSDS. Specifically, some studies tended to find no link (Auerbach, 1984; Watson et al., 1986) between narcissism and social desirability, whereas other studies found that the Entitlement/Exploitativeness subscale of the NPI (Watson & Morris, 1991; Watson et al., 1986, 1984) and/or total NPI scores (Fukunishi et al., 1995; Watson & Morris, 1991; Watson et al., 1984) were linked to *lower* social desirability. Furthermore, studies that included several measures of narcissism found mixed results. Using composite scores of narcissism measures, Raskin et al. (1991) replicated in undergraduate samples (Ns = 60–300) the inverse (i.e., negative) and moderate link between self-reported narcissism and SDR, but found a *positive* relationship between narcissism and self-deception (rs = 0.30–0.63). Alternatively, Hibbard (1992) found in university students ($N = 701$) null and inverse associations between various measures of narcissism and the MCSDS. Taken together, early studies found endorsement of narcissism either not to be associated with social desirability or to be associated with *lowered* SDR tendencies.

More recent studies also produced rather inconsistent results. In the only studies that examined an incarcerated sample (Sleep et al., 2017) or a patient sample (De Page & Merckelbach, 2021), small positive links emerged between (some subscales of) narcissism and underreporting. Specifically, the former study examined in a correctional subsample ($n = 703$) the connection between the NPI and underreporting scales of the MMPI-2 (i.e., the L-r and K-r scales). The link between the NPI and underreporting varied depending on the NPI scale: positive for the Leadership/Authority scale (rs = 0.13); null results for the Grandiose Exhibitionism scale (rs = 0.02 and 0.03); and inverse links

for the Entitlement/Exploitativeness scale (rs = −0.16). Furthermore, in two additional subsamples of undergraduates (ns = 228 and 482), across several self-report measures of narcissism and embedded validity indices of underreporting, inconsistent links (rs = −0.34 to 0.24) emerged for (grandiose) narcissism, whereas vulnerable narcissism was related to less underreporting (rs = −0.52 to −0.27). De Page and Merckelbach (2021) relied on adult patients and looked into the association between self-reported supernormality (i.e., denying common symptoms) and grandiose narcissism as measured by the NPI. These authors observed a small and positive correlation between these two constructs. Similarly, other more recent work (Fernie et al., 2016; Manley et al., 2018; Paulhus, 1998) found positive links between (specific subscales of) narcissism and SDR.

However, other recent studies replicated earlier findings suggesting that endorsement of narcissism is related to *lowered* self-reported SDR (Braun et al., 2016; Jones et al., 2016; Levi & Bachar, 2019; Sedikides et al., 2004; Wu et al., 2019; Yu, 2018). Still others observed no links between endorsement of narcissism and SDR (Barelds & Dijkstra, 2010; Lyvers et al., 2019). Tellingly, studies that employed multiple self-report measures of narcissism often yielded mixed findings (Barry et al., 2016; Gamache et al., 2018; Manley et al., 2018; Zeigler-Hill & Wallace, 2011) across various indices of narcissism and desirable responding.

Taken together, studies on self-reported narcissism and social desirability produced many null results and, when they did find correlations, they were typically small, non-replicable, and inconsistent. From a theoretical point of view, both a positive and a negative connection a priori make sense: that grandiose narcissists self-enhance by presenting themselves favorably on agentic and egoistic values (e.g., intelligence, dominance, and assertiveness; Campbell et al., 2002), regardless of context, stakes, and incentives (Hart et al., 2019; Maaß & Ziegler, 2017). Alternatively, (grandiose) narcissists might not value social acceptance. As things stand, the empirical literature does favor neither the first, nor the second hypothesis.

Attention Seeking

In the *DSM-5* AMPD, attention seeking refers to behavior focused on attracting notice, attention, or admiration from others (APA, 2013; Krueger et al., 2012). Historically, the concept has been associated with hysteria. One prominent idea about hysteria is that it typically involves overreporting of physical symptoms, but underreporting of psychological symptoms. One-third of hysteria patients are thought to display the phenomenon of *la belle indifference*: a careless attitude and lack of emotional distress about the somatic symptoms they present, which otherwise typically cause

grave concern in patients with the actual medical condition (Heubrock & Petermann, 1998; Stephens & Kamp, 1962). We found no studies that directly tested the connection between attention seeking/hysteria and stand-alone measures of underreporting, but there is some indirect evidence that bears on the issue. Germane to this is a study (Ornduff et al., 1988) that relied on the MMPI-2 profile related to hysteria — i.e., high elevations of the Hy scale. The authors divided the *Hy* scale into items that reflect "bodily concern" and "psychological denial" and observed that true somatic conditions were linked to moderate *Hy* scores and endorsement of "bodily concern" items, whereas somatic patients with a psychiatric history attained high *Hy* scores by prominently endorsing the "psychological denial" items. This pattern fits with the idea that neuroticism is related to overreporting of somatic symptoms (e.g., Rosmalen et al., 2007).

Suspiciousness and Withdrawal

Trait suspiciousness is related to presuming and ascribing ill-intent and harm to others, such as being mistreated, used, or persecuted. Thus, people high on suspiciousness are distrustful, alert, and sensitive in interpersonal contact. Similarly, people high on trait withdrawal have a preference to keep to themselves. This often manifests itself in a reserved, avoidant, or disinterested behavior towards social contact (APA, 2013; Krueger et al., 2012). There is no empirical literature on underreporting and suspiciousness or withdrawal, but an interesting lead can be found in the extant MMPI literature on self-deceptive underreporting (Bagby & Marshall, 2004). That is, subtle denial reflected in high scores on the K-scale (Correction scale; Meehl & Hathaway, 1946) may in part be related to self-reported cynicism, (interpersonal) mistrust, and introversion, whereas the defensiveness reflected in high scores on the S-scale (Superlative Self-Presentation; Butcher & Han, 1995) may in part be associated with self-reported hypersensitivity and suspicion on the MMPI (Nichols, 2001).

Theoretical considerations also suggest self-deceptive underreporting in paranoia. Paranoid individuals are thoroughly — and self-deceptively — convinced of external threats, such that they tend to be resistant to alternative explanations that stress internal biases and monitoring errors instead. Thus, it might well be the case that paranoid or withdrawn individuals are disinclined to share sensitive information about their symptoms, out of a fear that disclosure might reinforce external agency over them (e.g., power; negative social feedback). Along these lines, a relevant question is whether forewarning and coaching (e.g., Baer & Wetter, 1997) may make paranoid people particularly reluctant to endorse symptoms.

Discussion

Our scoping review examined whether personality traits are related to symptom distortion on stand-alone validity tests (i.e., heightened scores and/or failures) in the form of overreporting or underreporting, the latter broadly defined so as to include both denial of deviations from social norms (i.e., SDR) and denial of symptoms (i.e., supernormality). For overreporting, the literature shows strong links with depression and alexithymia, moderate links with dissociation and fantasy proneness, and weak links with apathy and (based on embedded validity scales) restricted affectivity and anhedonia. For underreporting, studies found small to moderate and difficult to replicate links with alexithymia, dissociation, and narcissism. More generally, the overwhelming majority of correlations between self-reported traits and overreporting or underreporting were in the small-to-moderate range, indicating modest effect sizes at best. This state of affairs is best viewed in light of the methodological limitations of the studies in this field, which give little reason to conclude that there exists robust evidence for traits as major causal antecedents of distorted symptom reports.

First, the modest connections between self-reported traits and validity test scores preclude any strong statements about the directionality underlying these connections. The problem here is, of course, the accuracy of self-reported traits (see also, Merten et al., 2007). This problem may be less pronounced in studies in which a majority of participants attained elevated scores on validity measures that were not sufficiently deviant to qualify as failures (i.e., surpassing an empirically established cut-off point). The distinction between scores that are slightly elevated versus those indicative of failure on a validity test is crucial, because subtle symptom distortion would not preclude clinicians from interpreting other psychometric data from the same person (e.g., self-reports of traits), whereas validity test failure would. Yet, papers often do not differentiate between subtle versus profound symptom distortion. For example, it might well be the case that traits such as depressivity or alexithymia promote mild forms of invalid symptom reporting (e.g., below cut-off), but do not serve as precursors of massive distortions (e.g., beyond the cut-off). Future studies could include separate analyses for individuals with elevated validity test scores and those who exhibit clear validity test failure. Of course, the issue touches upon a more fundamental question, namely whether symptom distortion is a dimensional or a categorical (i.e., taxon-like) phenomenon (e.g., Walters et al., 2008).

Second, the literature on traits and symptom distortion is scattered and mostly focused on single traits and

either symptom overreporting or underreporting. What is missing is an overarching and well-articulated theory that explains why and how various traits might be related to symptom distortion. What comes closest to this ideal is Suls and Howren's (2012) framework that specifies how anxiety and depression fuel attentional bias and recall bias, respectively, thereby promoting momentary and retrospective escalation of symptom reports.

Third, almost without exception, the empirical literature in this domain relied on cross-sectional set-ups and so the data are predominantly correlational in nature, precluding any strong statements with regard to causality. It is curious that so few studies have adopted a test–retest procedure: demonstrating temporal stability would be a first and essential step for research endeavors that want to clarify trait contribution to symptom overreporting or underreporting. That such an approach is, in principle, feasible is illustrated by the study of An et al. (2012), who administered PVTs to a small student sample ($N = 36$) on two occasions and found that participants who failed at one point in time also tended to fail at a later time. Of course, this is only tentative evidence for a trait-like component, because there might be other, more parsimonious explanations for this temporal stability. For example, inattentive or random responding might spuriously create temporal stability and even inflate correlations between trait measures and indices of overreporting or underreporting (see also Bowling et al., 2016).

A fourth methodological consideration concerns the small sample sizes in this research domain. In order to obtain stable estimates of medium correlations of $r = 0.3$, sample sizes of $N \geq 212$ are recommended, although strong or very large correlations are likely to be stable in relatively smaller samples (Schönbrodt & Perugini, 2013). Based on our scoping review, it is evident that many of the included studies do not meet this recommendation. Relatedly, structural equation modelling approaches in which the merits of several causal interpretations can be evaluated are conspicuously absent, precisely because such approaches require substantial sample sizes. It is obvious which rivaling interpretations would need to be tested in this field. For example, robust associations between depression and symptom overreporting might reflect depression leading to overreporting due to recall bias (Suls & Howren, 2012), it might reflect overreporting leading to inflated depression scores (Merten et al., 2019a), or it might reflect inattentive/careless responding leading to both symptom overreporting and escalated depression scores (Bowling et al., 2016).

In sum then, there is no convincing evidence that traits are such powerful antecedents of symptom distortion that they may contribute to a complete failure on a validity test. As things stand, certain traits can best be viewed as modest concomitants of symptom overreporting and underreporting, concomitants of which we do presently not know what their causal status is. Authors who failed to find clinically meaningful and statistically relevant associations between traits and symptom distortion echo our somewhat pessimistic summary of the findings in this field and symptom distortion, leading some of them to conclude that symptom distortion is primarily a situationally specific phenomenon (e.g., Young et al., 2016).

Several limitations apply to our scoping review. First, we did not perform a systematic review or meta-analysis. The restrictions in number and methodology of empirical studies in this domain make a quantitative review impossible. As said, studies typically relied on cross-sectional designs and remained silent on confounders such as trait-context interactions and carry-over effects (see also Niesten et al., 2015; Rogers, 1990; Walters, 1988). Moreover, they typically did not adhere to a "multi-method approach" of including multiple validity tests, such that a "two-failure rule" could be applied on identifying distorted symptom presentation (Victor et al., 2009). Second, we defined personality in the context of "pathological" traits. Other approaches would have been conceivable, including those focused on trait domains rather than facets (APA, 2013), models on maladaptive self-functioning and other-functioning (e.g., Livesley, 2003, 2006), and models of "normal" personality such as the Five Factor Model (FFM, e.g., McCrae & John, 1992). Indeed, symptom distortion might not be a function of personality pathology per se, rather normal personality functioning might contribute to symptom distortions in the lower range of the distribution.

Third, our review was perhaps naïve to assume that traits function independently of context (e.g., student populations, bona fide or contaminated clinical or forensic patient samples, criminal offenders). Arguably, incentives inherent to these contexts may foment symptom distortion (e.g., Mittenberg et al., 2002; Walters, 1988; Young, 2015) and their interaction with traits largely remains an open question. What we do know is that the effect of incentives is not more profound in those with antisocial traits or disorders, individuals who may be generally indifferent about impressions in the absence of incentives (for reviews, see Niesten et al., 2015; Van Impelen et al., 2017). Moreover, administering lengthy and boring tests to people may evoke in some of them careless/inattentive responding, which in itself is linked to certain traits (e.g., low conscientiousness; Bowling et al., 2016). Another clue is that people with heightened anxiety in interpersonal contact are prone to attentional bias, which may involve inflated interpretations of momentary symptoms (Suls & Howren, 2012). We suspect that their coping (e.g., approach or avoidant behavior) and symptom reporting (e.g., overreporting versus underreporting) may well depend on the extent to which clinicians are seen as a signal of safety (e.g., support, alleviation) or danger (e.g., untrustworthy) in that very moment. Future research may wish to focus on such trait-context interactions.

Conclusions

Symptom overreporting and underreporting occur on a regular basis. Thus, identification of their antecedents might inform assessment and treatment. Are there good reasons to suspect that personality traits are associated with distorted symptom presentation? The current review tried to address that question. Our scoping review of the empirical literature shows that there is currently little reason to believe that certain traits are conducive to symptom distortion. To take this research domain one step further, the focus on single traits such as the AMPD trait facets should be abandoned. Further progress depends on conceptual housecleaning, on the one hand, and broad inductive studies, on the other hand.

As to the first priority, the transdiagnostic literature on broad dimensions of psychopathology, including the domain level of traits — such as neuroticism/negative emotionality (N/NE) — might inspire more articulated theories on how traits, biases, and distorted symptom reports are related to each other. One recent idea is that the experience of symptoms involves both a sensory-perceptual component and an affective-motivational component. In individuals high on N/NE, the perceptual detail input would be poor, whereas the affective component would be strong, making these individuals vulnerable to symptom reports not related to physiological dysfunction (Van den Bergh et al., 2021). A similar theory is that N/NE is related to a negative interpretation bias in the form of negative disambiguation, i.e., interpretation of ambiguous signals in a catastrophic way. Prospective studies provide evidence for this line of reasoning (e.g., Engelhard & van den Hout, 2007).

As to the second priority, it would be informative to conduct cross-sectional studies that include broad measures of multiple traits and symptom distortions (e.g., overreporting and underreporting) in large samples with low or high incentives. Conducting factor analysis on such data would provide insight into the most important trait correlates of distorted symptom presentation and how they might interact with incentives. Until such studies are done, there is little ground for the claim that people who present with distorted symptom reports do so because they possess certain personality traits.

Declarations

Informed Consent All listed authors have approved the manuscript before submission. Ethics approval, participant consent, and/or data, material, and code repositories are not applicable.

Conflict of Interest The authors declare no competing interests.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

An, K. Y., Zakzanis, K. K., & Joordens, S. (2012). Conducting research with non-clinical healthy undergraduates: Does effort play a role in neuropsychological test performance? *Archives of Clinical Neuropsychology, 27*(8), 849–857. https://doi.org/10.1093/arclin/acs085

Ashendorf, L., Constantinou, M., & McCaffrey, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology, 19*(1), 125–130. https://doi.org/10.1093/arclin/19.1.125

Auerbach, J. S. (1984). Validation of two scales for narcissistic personality disorder. *Journal of Personality Assessment, 48*(6), 649–653. https://doi.org/10.1207/s15327752jpa4806_13

Bacon, A. M., Walsh, C. R., & Martin, L. (2013). Fantasy proneness and counterfactual thinking. *Personality and Individual Differences, 54*(4), 469–473. https://doi.org/10.1016/j.paid.2012.09.023

Badura, A. S. (2003). Theoretical and empirical exploration of the similarities between emotional numbing in posttraumatic stress disorder and alexithymia. *Journal of Anxiety Disorders, 17*(3), 349–360. https://doi.org/10.1016/s0887-6185(02)00201-3

Baer, R. A., & Wetter, M. W. (1997). Effects of information about validity scales on underreporting of symptoms on the personality assessment inventory. *Journal of Personality Assessment, 68*(2), 402–413. https://doi.org/10.1207/s15327752jpa6802_10

Bagby, R. M., & Marshall, M. B. (2004). Assessing underreporting response bias on the MMPI-2. *Assessment, 11*(2), 115–126. https://doi.org/10.1177/1073191104265918

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Barelds, D. P. H., & Dijkstra, P. (2010). Narcissistic Personality Inventory: Structure of the adapted Dutch version. *Scandinavian Journal of Psychology, 51*(2), 132–138. https://doi.org/10.1111/j.1467-9450.2009.00737.x

Barry, C. T., Lui, J. H. L., & Anderson, A. C. (2016). Adolescent narcissism, aggression, and prosocial behavior: The relevance of socially desirable responding. *Journal of Personality Assessment, 99*(1), 46–55. https://doi.org/10.1080/00223891.2016.1193812

Bass, C., & Wade, D. T. (2019). Malingering and factitious disorder. *Practical Neurology, 19*(2), 96–105. https://doi.org/10.1136/practneurol-2018-001950

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for Beck Depression Inventory-II*. Psychological Corporation.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1961). An inventory for measuring depression. *Archives*

of General Psychiatry, 4, 561–571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Beere, D. B., Pica, M., & Maurer, L. (1996). Social desirability and the Dissociative Experiences Scale. *Dissociation: Progress in the Dissociative Disorders, 9*(2), 130–133.

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2-Restructured Form): Manual for administration, scoring, and interpretation*. University of Minnesota Press.

Bensch, D., Maaß, U., Greiff, S., Horstmann, K. T., & Ziegler, M. (2019a). The nature of faking: A homogeneous and predictable construct? *Psychological Assessment, 31*(4), 532–544. https://doi.org/10.1037/pas0000619

Bensch, D., Paulhus, D. L., Stankov, L., & Ziegler, M. (2019b). Teasing apart overclaiming, overconfidence, and socially desirable responding. *Assessment, 26*(3), 351–363. https://doi.org/10.1177/1073191117700268

Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease, 174*(12), 727–735. https://doi.org/10.1097/00005053-198612000-00004

Berry, D. T. R., & Nelson, N. W. (2010). DSM-5 and malingering: A modest proposal. *Psychological Injury and Law, 3*(4), 295–303. https://doi.org/10.1007/s12207-010-9087-7

Bogaerts, K., Rayen, L., Lavrysen, A., van Diest, I., Janssens, T., Schruers, K. & van den Bergh, O. (2015). Unraveling the relationship between trait negative affectivity and habitual symptom reporting. *PloS ONE, 10*(1), Article e0115748. https://doi.org/10.1371/journal.pone.0115748

Bouchard, G. (2008). Alexithymia among students and professionals in function of disciplines. *Social Behavior and Personality: An International Journal, 36*(3), 303–314. https://doi.org/10.2224/sbp.2008.36.3.303

Bowling, N. A., Huang, J. L., Bragg, C. B., Khazon, S., Liu, M., & Blackmore, C. E. (2016). Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality. *Journal of Personality and Social Psychology, 111*(2), 218–229. https://doi.org/10.1037/pspp0000085

Brady, R. E., Bujarski, S. J., Feldner, M. T., & Pyne, J. M. (2017). Examining the effects of alexithymia on the relation between posttraumatic stress disorder and over-reporting. *Psychological Trauma: Theory, Research, Practice, and Policy, 9*(1), 80–87. https://doi.org/10.1037/tra0000155

Brand, B. L., McNary, S. W., Loewenstein, R. J., Kolos, A. C., & Barr, S. R. (2006). Assessment of genuine and simulated dissociative identity disorder on the Structured Interview of Reported Symptoms. *Journal of Trauma & Dissociation, 7*(1), 63–85. https://doi.org/10.1300/j229v07n01_06

Braun, S., Kempenaers, C., Linkowski, P., & Loas, G. (2016). French adaptation of the Narcissistic Personality Inventory in a Belgian French-speaking sample. *Frontiers in Psychology, 7*, Article 01980. https://doi.org/10.3389/fpsyg.2016.01980

Briere, J. (1995). *Trauma Symptom Inventory professional manual*. Psychological Assessment Resources.

Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., Reynolds, C. R., & Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology, 20*(4), 419–426. https://doi.org/10.1016/j.acn.2005.02.002

Butcher, J., Graham, J., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W., & Kaemmer, B. (2001). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration, scoring, and interpretation-Revised edition*. University of Minnesota Press.

Butcher, J. N., & Han, K. (1995). *Development of an MMPI-2 scale to assess the presentation of self in a superlative manner: The S Scale*. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 10, pp. 25–50). Lawrence Erlbaum Associates, Inc.

Byrne, N., & Ditto, B. (2005). Alexithymia, cardiovascular reactivity, and symptom reporting during blood donation. *Psychosomatic Medicine, 67*(3), 471–475. https://doi.org/10.1097/01.psy.0000160471.66399.12

Cain, N. M., Pincus, A. L., & Ansell, E. B. (2008). Narcissism at the crossroads: Phenotypic description of pathological narcissism across clinical theory, social/personality psychology, and psychiatric diagnosis. *Clinical Psychology Review, 28*(4), 638–656. https://doi.org/10.1016/j.cpr.2007.09.006

Callahan, M. R., Tolman, R. M., & Saunders, D. G. (2003). Adolescent dating violence victimization and psychological well-being. *Journal of Adolescent Research, 18*(6), 664–681. https://doi.org/10.1177/0743558403254784

Campbell, W. K., Rudich, E. A., & Sedikides, C. (2002). Narcissism, self-esteem, and the positivity of self-views: Two portraits of self-love. *Personality and Social Psychology Bulletin, 28*(3), 358–368. https://doi.org/10.1177/0146167202286007

Carneiro, D. R., Sousa, M., Morgadinho, A., & Januário, C. (2019). Dopamine dysregulation syndrome induced by proxy. *BMJ Case Reports, 12*(5), Article e228495. https://doi.org/10.1136/bcr-2018-228495

Chafetz, M. D., Williams, M. A., Ben-Porath, Y. S., Bianchini, K. J., Boone, K. B., Kirkwood, M. W., Larrabee, G. J., & Ord, J. S. (2015). Official position of the American Academy of Clinical Neuropsychology Social Security Administration Policy on validity testing: Guidance and recommendations for change. *The Clinical Neuropsychologist, 29*(6), 723–740. https://doi.org/10.1080/13854046.2015.1099738

Christensen, L. (1977). The negative subject: Myth, reality, or a prior experimental experience effect? *Journal of Personality and Social Psychology, 35*(6), 392–400. https://doi.org/10.1037/0022-3514.35.6.392

Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: Supernormality. *The Clinical Neuropsychologist, 17*(2), 235–243. https://doi.org/10.1076/clin.17.2.235.16507

Cima, M., van Bergen, S., & Kremer, K. (2008). Development of the Supernormality Scale-Revised and its relationship with psychopathy. *Journal of Forensic Sciences, 53*(4), 975–981. https://doi.org/10.1111/j.1556-4029.2008.00740.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.

Collin-Vézina, D., Cyr, M., Pauzé, R., & McDuff, P. (2005). The role of depression and dissociation in the link between childhood sexual abuse and later parental practices. *Journal of Trauma & Dissociation, 6*(1), 71–97. https://doi.org/10.1300/j229v06n01_05

Condon, L. P., & Lynn, S. J. (2014). State and trait dissociation: Evaluating convergent and discriminant validity. *Imagination, Cognition and Personality, 34*(1), 25–37. https://doi.org/10.2190/ic.34.1.c

Constans, J. I., Kimbrell, T. A., Nanney, J. T., Marx, B. P., Jegley, S., & Pyne, J. M. (2014). Over-reporting and the modified Stroop effect in operation enduring and Iraqi freedom veterans with and without PTSD. *Journal of Abnormal Psychology, 123*(1), 81–90. https://doi.org/10.1037/a0035100

Costa, P. T., & McCrae, R. R. (1987). Neuroticism, somatic complaints, and disease: Is the bark worse than the bite? *Journal of Personality, 55*(2), 299–316. https://doi.org/10.1111/j.1467-6494.1987.tb00438.x

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*(4), 349–354. https://doi.org/10.1037/h0047358

Dandachi-FitzGerald, B., Duits, A. A., Leentjens, A. F., Verhey, F. R., & Ponds, R. W. H. M. (2020). Performance and symptom validity assessment in patients with apathy and cognitive impairment.

Journal of the International Neuropsychological Society, 26(3), 314–321. https://doi.org/10.1017/S1355617719001139

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25*(5), 812–828. https://doi.org/10.1080/13854046.2011.583280

Dandachi-FitzGerald, B., Merckelbach, H., & Ponds, R. W. H. M. (2017). Neuropsychologists' ability to predict distorted symptom presentation. *Journal of Clinical and Experimental Neuropsychology, 39*(3), 257–264. https://doi.org/10.1080/13803395.2016.1223278

De Gucht, V., & Heiser, W. (2003). Alexithymia and somatisation: A quantitative review of the literature. *Journal of Psychosomatic Research, 54*(5), 425–434. https://doi.org/10.1016/s0022-3999(02)00467-1

De Page, L., & Merckelbach, H. (2021). Associations between supernormality ('faking good'), narcissism and depression: An exploratory study in a clinical sample. *Clinical Psychology & Psychotherapy, 28*(1), 182–188. https://doi.org/10.1002/cpp.2500

Duberstein, P. R., & Heisel, M. J. (2007). Personality traits and the reporting of affective disorder symptoms in depressed patients. *Journal of Affective Disorders, 103*(1–3), 165–171. https://doi.org/10.1016/j.jad.2007.01.025

Elzinga, B. M., Bermond, B., & van Dyck, R. (2002). The relationship between dissociative proneness and alexithymia. *Psychotherapy and Psychosomatics, 71*(2), 104–111. https://doi.org/10.1159/000049353

Engelhard, I. M., & van den Hout, M. A. (2007). Preexisting neuroticism, subjective stressor severity, and posttraumatic stress in soldiers deployed to Iraq. *The Canadian Journal of Psychiatry, 52*(8), 505–509. https://doi.org/10.1177/070674370705200080

Erdodi, L. A., & Lichtenstein, J. D. (2017). Invalid before impaired: An emerging paradox of embedded validity indicators. *The Clinical Neuropsychologist, 31*(6–7), 1029–1046.

Evans, J., Lange, R., Houran, J., & Lynn, S. J. (2019). Further psychometric exploration of the transliminality construct. *Psychology of Consciousness: Theory, Research, and Practice, 6*(4), 417–438. https://doi.org/10.1037/cns0000163

Evren, C., Sar, V., Evren, B., Semiz, U., Dalbudak, E., & Cakmak, D. (2008). Dissociation and alexithymia among men with alcoholism. *Psychiatry and Clinical Neurosciences, 62*, 40–47. https://doi.org/10.1111/j.1440-1819.2007.01775.x

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. Plenum Press.

Eysenck, H., & Gudjonsson, G. H. (1989). *The causes and cures of criminality*. Plenum Press.

Fairclough, S. H., & Goodwin, L. (2007). The effect of psychological stress and relaxation on interoceptive accuracy: Implications for symptom perception. *Journal of Psychosomatic Research, 62*(3), 289–295. https://doi.org/10.1016/j.jpsychores.2006.10.017

Fernie, B. A., Fung, A., & Nikčević, A. V. (2016). Different coping strategies amongst individuals with grandiose and vulnerable narcissistic traits. *Journal of Affective Disorders, 205*, 301–305. https://doi.org/10.1016/j.jad.2016.08.009

Fischer, A. R., & Good, G. E. (1997). Men and psychotherapy: An investigation of alexithymia, intimacy, and masculine gender roles. *Psychotherapy: Theory, Research, Practice, Training, 34*(2), 160–170. https://doi.org/10.1037/h0087646

Fogley, R., Warman, D., & Lysaker, P. H. (2014). Alexithymia in schizophrenia: Associations with neurocognition and emotional distress. *Psychiatry Research, 218*(1–2), 1–6. https://doi.org/10.1016/j.psychres.2014.04.020

Friedman, H. S., Kern, M. L., & Reynolds, C. A. (2010). Personality and health, subjective well-being, and longevity. *Journal of*

Personality, 78(1), 179–216. https://doi.org/10.1111/j.1467-6494.2009.00613.x

Fukunishi, I. (1994). Social desirability and alexithymia. *Psychological Reports, 75*(2), 835–838. https://doi.org/10.2466/pr0.1994.75.2.835

Fukunishi, I., Hattori, M., Nakamura, H., & Nakagawa, T. (1995). Hostility is related to narcissism controlling for social desirability: Studies of college students and patients with myocardial infarction. *Journal of Psychosomatic Research, 39*(2), 215–220. https://doi.org/10.1016/0022-3999(94)00094-l

Gamache, D., Savard, C., & Maheux-Caron, V. (2018). French adaptation of the Short Dark Triad: Psychometric properties and a head-to-head comparison with the Dirty Dozen. *Personality and Individual Differences, 122*, 164–170. https://doi.org/10.1016/j.paid.2017.10.027

Giesbrecht, T., Lynn, S. J., Lilienfeld, S. O., & Merckelbach, H. (2008). Cognitive processes in dissociation: An analysis of core theoretical assumptions. *Psychological Bulletin, 134*(5), 617–647. https://doi.org/10.1037/0033-2909.134.5.617

Giesbrecht, T., & Merckelbach, H. (2006). Dissociatieve symptomen en slaap [Dissociative symptoms and sleep]. *Tijdschrift Voor Psychiatrie [journal of Psychiatry], 48*(3), 207–215.

Giromini, L., Young, G., & Sellbom, M. (2022). Assessing negative response bias using self-report measures: Introducing the special issue. *Psychological Injury and Law.*

Gooding, D. C., & Tallent, K. A. (2003). Spatial, object, and affective working memory in social anhedonia: An exploratory study. *Schizophrenia Research, 63*(3), 247–260. https://doi.org/10.1016/s0920-9964(02)00326-2

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research, 78*(2–3), 199–208. https://doi.org/10.1016/j.schres.2005.02.016

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Grynberg, D., Davydov, D. M., Vermeulen, N., & Luminet, O. (2012). Alexithymia is associated with an augmenter profile, but not only: Evidence for anticipation to arousing music. *Scandinavian Journal of Psychology, 53*(5), 375–381. https://doi.org/10.1111/j.1467-9450.2012.00962.x

Grynberg, D., & Pollatos, O. (2015). Alexithymia modulates the experience of the rubber hand illusion. *Frontiers in Human Neuroscience, 9*, Article 00357. https://doi.org/10.3389/fnhum.2015.00357

Hamilton, M. (1980). Rating depressive patients. *The Journal of Clinical Psychiatry, 41*(12, Sec 2), 21–24.

Hart, W., Breeden, C. J., & Richardson, K. (2019). Differentiating dark personalities on impression management. *Personality and Individual Differences, 147*, 58–62. https://doi.org/10.1016/j.paid.2019.04.030

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., Conference Participants 1. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23*(7), 1093–1129. https://doi.org/10.1080/13854040903155063

Helmes, E., McNeill, P. D., Holden, R. R., & Jackson, C. (2008). The construct of alexithymia: Associations with defense mechanisms. *Journal of Clinical Psychology, 64*(3), 318–331. https://doi.org/10.1002/jclp.20461

Herbert, B. M., Herbert, C., & Pollatos, O. (2011). On the relationship between interoceptive awareness and alexithymia: Is interoceptive awareness related to emotional awareness? *Journal of Personality, 79*(5), 1149–1175. https://doi.org/10.1111/j.1467-6494.2011.00717.x

Heubrock, D., & Petermann, F. (1998). Neuropsychological assessment of suspected malingering: Research results, evaluation techniques, and further directions of research and application. *European Journal of Psychological Assessment, 14*(3), 211–225. https://doi.org/10.1027/1015-5759.14.3.211

Hibbard, S. (1992). Narcissism, shame, masochism, and object relations: An exploratory correlational study. *Psychoanalytic Psychology, 9*(4), 489–508. https://doi.org/10.1037/h0079392

Howren, M. B., Suls, J., & Martin, R. (2009). Depressive symptomatology, rather than neuroticism, predicts inflated physical symptom reports in community-residing women. *Psychosomatic Medicine, 71*(9), 951–957. https://doi.org/10.1097/PSY.0b013e3181b9b2d7

Hyman, I. E., & Billings, F. J. (1998). Individual differences and the creation of false childhood memories. *Memory, 6*(1), 1–20. https://doi.org/10.1080/741941598

Institute of Medicine. (2015). *Psychological testing in the service of disability determination*. The National Academies Press. https://doi.org/10.17226/21704

Jones, B. D., Woodman, T., Barlow, M., & Roberts, R. (2016). The darker side of personality: Narcissism predicts moral disengagement and antisocial behavior in sport. *The Sport Psychologist, 31*(2), 109–116. https://doi.org/10.1123/tsp.2016-0007

Kashdan, T. B., Elhai, J. D., & Frueh, B. C. (2007). Anhedonia, emotional numbing, and symptom overreporting in male veterans with PTSD. *Personality and Individual Differences, 43*(4), 725–735. https://doi.org/10.1016/j.paid.2007.01.013

King, L. A., Emmons, R. A., & Woodley, S. (1992). The structure of inhibition. *Journal of Personality Research, 26*(1), 85–102. https://doi.org/10.1016/0092-6566(92)90061-8

Kroner, D. G., & Forth, A. E. (1995). The Toronto Alexithymia Scale with incarcerated offenders. *Personality and Individual Differences, 19*(5), 625–634. https://doi.org/10.1016/0191-8869(95)00116-N

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*(9), 1879–1890. https://doi.org/10.1017/S0033291711002674

Kunst, M., Winkel, F. W., & Bogaerts, S. (2011). Recalled peritraumatic reactions, self-reported PTSD, and the impact of malingering and fantasy proneness in victims of interpersonal violence who have applied for state compensation. *Journal of Interpersonal Violence, 26*(11), 2186–2210. https://doi.org/10.1177/0886260510383032

Levi, E., & Bachar, E. (2019). The moderating role of narcissism on the relationship between posttraumatic growth and PTSD symptoms. *Personality and Individual Differences, 138*, 292–297. https://doi.org/10.1016/j.paid.2018.10.022

Linden, W., Lenz, J. W., & Stossel, C. (1996). Alexithymia, defensiveness and cardiovascular reactivity to stress. *Journal of Psychosomatic Research, 41*(6), 575–583. https://doi.org/10.1016/s0022-3999(96)00229-2

Livesley, W. J. (2003). *Practical management of personality disorders.* Guilford Press.

Livesley, W. J. (2006). *General assessment of personality disorder (GAPD).* Unpublished manuscript. Department of Psychiatry, University of British Colombia.

Lyvers, M., Narayanan, S. S., & Thorberg, F. A. (2019). Disordered social media use and risky drinking in young adults: Differential associations with addiction-linked traits. *Australian Journal of Psychology, 71*(3), 223–231. https://doi.org/10.1111/ajpy.12236

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

Manley, H., Roberts, R., Beattie, S., & Woodman, T. (2018). I'll get there because I'm great, or am I? Narcissistic vulnerability moderates the narcissistic grandiosity – goal persistence relationship. *Personality and Individual Differences, 120*, 65–74. https://doi.org/10.1016/j.paid.2017.08.024

Marin, R. S., Biedrzycki, R. C., & Firinciogullari, S. (1991). Reliability and validity of the Apathy Evaluation Scale. *Psychiatry Research, 38*(2), 143–162.

Martin, P. K., Schroeder, R. W., & Odland, A. P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. *The Clinical Neuropsychologist, 29*(6), 741–776. https://doi.org/10.1080/13854046.2015.1087597

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology, 30*(5), 525–564. https://doi.org/10.1037/h0053634

Merckelbach, H., Boskovic, I., Pesy, D., Dalsklev, M., & Lynn, S. J. (2017). Symptom overreporting and dissociative experiences: A qualitative review. *Consciousness and Cognition, 49*, 132–144. https://doi.org/10.1016/j.concog.2017.01.007

Merckelbach, H., Dandachi-FitzGerald, B., van Helvoort, D., Jelicic, M., & Otgaar, H. (2019). When patients overreport symptoms: More than just malingering. *Current Directions in Psychological Science, 28*(3), 321–326. https://doi.org/10.1177/0963721419837681

Merckelbach, H., Giesbrecht, T., & van Heugten-van der Kloet, D., De Jong, J., Meyer, T., & Rietman, K. (2015). The overlap between dissociative symptoms and symptom over-reporting. *European Journal of Psychiatry, 29*(3), 165–172. https://doi.org/10.4321/S0213-61632015000300001

Merckelbach, H., Horselenberg, R., & Muris, P. (2001). The Creative Experiences Questionnaire (CEQ): A brief self-report measure of fantasy proneness. *Personality and Individual Differences, 31*(6), 987–995. https://doi.org/10.1016/S0191-8869(00)00201-4

Merckelbach, H., Muris, P., & Rassin, E. (1999). Fantasy proneness and cognitive failures as correlates of dissociative experiences. *Personality and Individual Differences, 26*(5), 961–967. https://doi.org/10.1016/S0191-8869(98)00193-7

Merckelbach, H., Otgaar, H., & Lynn, S. J. (2021). Empirical research on fantasy proneness and its correlates 2000–2018: A meta-analysis. *Psychology of Consciousness: Theory, Research, and Practice.* Advance online publication. https://doi.org/10.1037/cns0000272

Merckelbach, H., Prins, C., Boskovic, I., Niesten, I., & Campo, À. (2018). Alexithymia as a potential source of symptom overreporting: An exploratory study in forensic patients and non-forensic participants. *Scandinavian Journal of Psychology, 59*(2), 192–197. https://doi.org/10.1111/sjop.12427

Merckelbach, H., & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology, 18*(2), 145–152. https://doi.org/10.1093/arclin/18.2.145

Merckelbach, H., & van de Ven, V. (2001). Another White Christmas: Fantasy proneness and reports of 'hallucinatory experiences' in undergraduate students. *Journal of Behavior Therapy and Experimental Psychiatry, 32*(3), 137–144. https://doi.org/10.1016/s0005-7916(01)00029-5

Merten, T., Friedel, E., Mehren, G., & Stevens, A. (2007). Uber die Validität von Persönlichkeitsprofilen in der nervenärztlichen Begutachtung [Negative response bias and the validity of personality profiles in neuropsychiatric assessment]. *Der Nervenarzt, 78*(5), 511–520. https://doi.org/10.1007/s00115-006-2101-3

Merten, T., Giger, P., Merckelbach, H., & Stevens, A. (2019a). *Self-Report Symptom Inventory (SRSI) – Deutsche version. Manual [Manual*

of the German version of the Self-Report Symptom Inventory]. Hogrefe.

Merten, T., Kaminski, A., & Pfeiffer, W. (2019b). Prevalence of over-reporting on symptom validity tests in a large sample of psychosomatic rehabilitation inpatients. *The Clinical Neuropsychologist, 34*(5), 1004–1024. https://doi.org/10.1080/13854046.2019.1694073

Merten, T., Merckelbach, H., Giger, P., & Stevens, A. (2016). The Self-Report Symptom Inventory (SRSI): A new instrument for the assessment of symptom overreporting. *Psychological Injury and Law, 9*(2), 102–111. https://doi.org/10.1007/s12207-016-9257-3

Messina, A., Fogliani, A. M., & Paradiso, S. (2010). Association between alexithymia, neuroticism, and social desirability scores among Italian graduate students. *Psychological Reports, 107*(1), 185–192. https://doi.org/10.2466/02.07.13.15.PR0.107.4.185-192

Miller, H. A. (2001). *M-FAST: Miller Forensic Assessment of Symptoms Test and professional manual*. Psychological Assessment Resources.

Miller, J., Hoffman, B., Gaughan, E., Gentile, B., Maples, J., & Campbell, W. (2011). Grandiose and vulnerable narcissism: A nomological network analysis. *Journal of Personality, 79*(5), 1012–1042. https://doi.org/10.1111/j.1467-6494.2010.00711.x

Miron, A. M., & Brehm, J. W. (2006). Reaktanz theorie - 40 Jahre später [Reactance Theory - 40 Years Later]. *Zeitschrift Für Sozialpsychologie, 37*(1), 9–18. https://doi.org/10.1024/0044-3514.37.1.9

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24*(8), 1094–1102. https://doi.org/10.1076/jcen.24.8.1094.8379

Nichols, D. S. (2001). *Essentials of psychological assessment series. Essentials of MMPI-2 assessment.* John Wiley & Sons Inc.

Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015). Antisocial features and "faking bad": A critical note. *International Journal of Law and Psychiatry, 41*, 34–42. https://doi.org/10.1016/j.ijlp.2015.03.005

Ornduff, S. R., Frances Brennan, A., & Barrett, C. L. (1988). The Minnesota Multiphasic Personality Inventory (MMPI) Hysteria (Hy) scale: Scoring Bodily Concern and Psychological Denial subscales in chronic back pain patients. *Journal of Behavioral Medicine, 11*(2), 131–146. https://doi.org/10.1007/bf00848261

Otgaar, H., Moldoveanu, G., Wang, J., & Howe, M. L. (2017). Exploring the consequences of nonbelieved memories in the DRM paradigm. *Memory, 25*(7), 922–933. https://doi.org/10.1080/09658211.2016.1272701

Otto, R. K. (2008). Challenges and advances in assessment of response styles in forensic examination contexts. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 365–375). Guilford.

Paulhus, D. (1988). *Manual for the balanced inventory of desirable responding: Version 6.* Unpublished manuscript. University of British Columbia.

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 74*(5), 1197–1208. https://doi.org/10.1037/0022-3514.74.5.1197

Peace, K. A., & Masliuk, K. A. (2011). Do motivations for malingering matter? Symptoms of malingered PTSD as a function of motivation and trauma type. *Psychological Injury and Law, 4*(1), 44–55. https://doi.org/10.1007/s12207-011-9102-7

Peters, M. J. V., Jelicic, M., Moritz, S., Hausschildt, M., & Jelinek, L. (2013). Assessing the boundaries of symptom over-reporting using the Structured Inventory of Malingered Symptomatology in a clinical schizophrenia sample: Its relation to symptomatology and neurocognitive dysfunctions. *Journal of Experimental Psychopathology, 4*(1), 64–77. https://doi.org/10.5127/jep.023811

Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G., & Levy, K. N. (2009). Initial construction and validation of the Pathological Narcissism Inventory. *Psychological Assessment, 21*(3), 365–379. https://doi.org/10.1037/a0016530

Porcelli, P., Guidi, J., Sirri, L., Grandi, S., Grassi, L., Ottolini, F., Pasquini, P., Picardi, A., Rafanelli, C., Rigatelli, M., Sonino, N., & Fava, G. A. (2013). Alexithymia in the medically ill. Analysis of 1190 patients in gastroenterology, cardiology, oncology and dermatology. *General Hospital Psychiatry, 35*(5), 521–527. https://doi.org/10.1016/j.genhosppsych.2013.04.005

Radaelli, D., Benedetti, F., Cavallaro, R., Colombo, C., & Smeraldi, E. (2013). The reality monitoring deficit as a common neuropsychological correlate of schizophrenic and affective psychosis. *Behavioral Sciences, 3*(2), 244–252. https://doi.org/10.3390/bs3020244

Raskin, R. N., & Hall, C. S. (1979). *A Narcissistic Personality Inventory. Psychological Reports, 45*(2), 590. https://doi.org/10.2466/pr0.1979.45.2.590

Raskin, R., Novacek, J., & Hogan, R. (1991). Narcissism, self-esteem, and defensive self-enhancement. *Journal of Personality, 59*(1), 19–38. https://doi.org/10.1111/j.1467-6494.1991.tb00766.x

Rauschenberger, S. L., & Lynn, S. J. (1995). Fantasy proneness, DSM-III-R Axis I psychopathology, and dissociation. *Journal of Abnormal Psychology, 104*(2), 373–380. https://doi.org/10.1037/0021-843x.104.2.373

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the test of memory malingering. *Archives of Clinical Neuropsychology, 16*(5), 501–506. https://doi.org/10.1016/S0887-6177(00)00064-0

Reyno, S. M., Simmons, M., & Kinley, J. (2020). A meta-analytic study examining the relationship between alexithymia and dissociation in psychiatric and nonclinical populations. *Research in Psychotherapy: Psychopathology, Process and Outcome, 23*(1), 99–107. https://doi.org/10.4081/ripppo.2020.439

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*(6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934

Rogers, R. (1990). Development of a new classificatory model of malingering. *Bulletin of the American Academy of Psychiatry and the Law, 18*(3), 323–333.

Rogers, R., Gillard, N. D., Wooley, C. N., & Ross, C. A. (2011). The detection of feigned disabilities: The effectiveness of the Personality Assessment Inventory in a traumatized inpatient sample. *Assessment, 19*(1), 77–88. https://doi.org/10.1177/1073191111422031

Rogers, R., Payne, J. W., Correa, A. A., Gillard, N. D., & Ross, C. A. (2009). A study of the SIRS with severely traumatized patients. *Journal of Personality Assessment, 91*(5), 429–438. https://doi.org/10.1080/00223890903087745

Rosen, G. M. (2006). DSM's cautionary guideline to rule out malingering can protect the PTSD data base. *Journal of Anxiety Disorders, 20*(4), 530–535. https://doi.org/10.1016/j.janxdis.2005.03.004

Rosmalen, J. G., Neeleman, J., Gans, R. O., & de Jonge, P. (2007). The association between neuroticism and self-reported common somatic symptoms in a population cohort. *Journal of Psychosomatic Research, 62*(3), 305–311. https://doi.org/10.1016/j.jpsychores.2006.10.014

Shad, M. U., Tamminga, C. A., Cullum, M., Haas, G. L., & Keshavan, M. S. (2006). Insight and frontal cortical function in schizophrenia: A review. *Schizophrenia Research, 86*(1–3), 54–70. https://doi.org/10.1016/j.schres.2006.06.006

Schaefer, J., Giangrande, E., Weinberger, D. R., & Dickinson, D. (2013). The global cognitive impairment in schizophrenia: Consistent over decades and around the world. *Schizophrenia*

Research, 150(1), 42–50. https://doi.org/10.1016/j.schres.2013.07.009

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schroeder, R. W., & Marshall, P. S. (2011). Evaluation of the appropriateness of multiple symptom validity indices in psychotic and non-psychotic psychiatric populations. *The Clinical Neuropsychologist, 25*(3), 437–453. https://doi.org/10.1080/13854046.2011.556668

Sharf, A. J., Rogers, R., Williams, M. M., & Henry, S. A. (2017). The effectiveness of the MMPI-2-RF in detecting feigned mental disorders and cognitive deficits: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment, 39*(3), 441–455. https://doi.org/10.1007/s10862-017-9590-1

Sedikides, C., Rudich, E. A., Gregg, A. P., Kumashiro, M., & Rusbult, C. (2004). Are normal narcissists psychologically healthy?: Self-esteem matters. *Journal of Personality and Social Psychology, 87*(3), 400–416. https://doi.org/10.1037/0022-3514.87.3.400

Sleep, C. E., Sellbom, M., Campbell, W. K., & Miller, J. D. (2017). Narcissism and response validity: Do individuals with narcissistic features underreport psychopathology? *Psychological Assessment, 29*(8), 1059–1064. https://doi.org/10.1037/pas0000413

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the Academy of Psychiatry and the Law, 25*(2), 180–183.

Stanley, B., & Wilson, S. T. (2006). Heightened subjective experience of depression in borderline personality disorder. *Journal of Personality Disorders, 20*(4), 307–318. https://doi.org/10.1521/pedi.2006.20.4.307

Stephens, J. H., & Kamp, M. (1962). On some aspects of hysteria: A clinical study. *Journal of Nervous and Mental Disease, 134*(4), 305–315. https://doi.org/10.1097/00005053-196204000-00002

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research, 157*(1–3), 191–200. https://doi.org/10.1016/j.psychres.2007.01.003

Stevens, A., Schneider, K., Liske, B., Hermle, L., Huber, H., & Hetzel, G. (2014). Is subnormal cognitive performance in schizophrenia due to lack of effort or to cognitive impairment? *German Journal of Psychiatry, 17*(1), 1–9.

Suls, J., & Howren, M. B. (2012). Understanding the physical-symptom experience: The distinctive contributions of anxiety and depression. *Current Directions in Psychological Science, 21*(2), 129–134. https://doi.org/10.1177/0963721412439298

Sweet, J. J., & Guidotti Breting, L. M. (2013). Symptom validity test research: Status and clinical implications. *Journal of Experimental Psychopathology, 4*(1), 6–19. https://doi.org/10.5127/jep.022311

Thompson, C. M., Lin, H., & Parsloe, S. (2018). Misrepresenting health conditions through fabrication and exaggeration: An adaptation and replication of the false alarm effect. *Health Communication, 33*(5), 562–575. https://doi.org/10.1080/10410236.2017.1283563

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*(3), 260–268. https://doi.org/10.1037/1040-3590.9.3.260

Tournois, J., Mesnil, F., & Kop, J. -L. (2000). Autoduperie et héteroduperie: Un instrument de mesure de la désirabilité sociale [Self-deception and other-deception: A social desirability questionnaire]. *European Review of Applied Psychology [revue Européenne De Psychologie Appliquée], 50*(1), 219–233.

Van den Bergh, O., Brosschot, J., Critchley, H., Thayer, J. F., & Ottaviani, C. (2021). Better safe than sorry: A common signature of general vulnerability for psychopathology. *Perspectives on Psychological Science, 16*(2), 225–246. https://doi.org/10.1177/1745691620950690

Van der Heide, D., Boskovic, I., & Merckelbach, H. (2017). Standard symptom inventories for asylum seekers in a psychiatric hospital: Limited utility due to poor symptom validity. *Psychological Injury and Law, 10*(4), 358–367. https://doi.org/10.1007/s12207-017-9302-x

Van der Heide, D., Boskovic, I., Van Harten, P., & Merckelbach, H. (2019). Psychosis as a confounder of symptom credibility testing in a transcultural sample. *The Journal of Forensic Psychiatry & Psychology, 31*(1), 64–75. https://doi.org/10.1080/14789949.2019.1663899

Van der Heide, D., & Merckelbach, H. (2016). Validity of symptom reports of asylum seekers in a psychiatric hospital: A descriptive study. *International Journal of Law and Psychiatry, 49*(Part A), 40–46. https://doi.org/10.1016/j.ijlp.2016.05.007

Van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *The Clinical Neuropsychologist, 28*(8), 1336–1365. https://doi.org/10.1080/13854046.2014.984763

Van Impelen, A., Merckelbach, H., Niesten, I. J., Jelicic, M., Huhnt, B., & à Campo, J. (2017). Biased symptom reporting and antisocial behaviour in forensic samples: A weak link. *Psychiatry, Psychology and Law, 24*(4), 530–548. https://doi.org/10.1080/13218719.2016.1256017

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry, 18*(4), 326–354. https://doi.org/10.1080/13546805.2012.721728

Victor, T. L., Boone, K. B., Serpa, J. G., Buehler, J., & Ziegler, E. A. (2009). Interpreting the meaning of multiple symptom validity test failure. *The Clinical Neuropsychologist, 23*(2), 297–313. https://doi.org/10.1080/13854040802232682

Waite, S., & Geddes, A. (2006). Malingered psychosis leading to involuntary psychiatric hospitalization. *Australasian Psychiatry, 14*(4), 419–421. https://doi.org/10.1080/j.1440-1665.2006.02317.x

Waldstein, S. R., Kauhanen, J., Neumann, S. A., & Katzel, L. I. (2002). Alexithymia and cardiovascular risk in older adults: Psychosocial, psychophysiological, and biomedical correlates. *Psychology & Health, 17*(5), 597–610. https://doi.org/10.1080/08870440290025803

Walters, G. D. (1988). Assessing dissimulation and denial on the MMPI in a sample of maximum security, male inmates. *Journal of Personality Assessment, 52*(3), 465–474. https://doi.org/10.1207/s15327752jpa5203_8

Walters, G. D., Rogers, R., Berry, D. T. R., Miller, H. A., Duncan, S. A., McCusker, P. J., Payne, J. W., & Granacher, R. P., Jr. (2008). Malingering as a categorical or dimensional construct: The latent structure of feigned psychopathology as measured by the SIRS and MMPI-2. *Psychological Assessment, 20*(3), 238–247. https://doi.org/10.1037/1040-3590.20.3.238

Watson, P. J., Grisham, S. O., Trotter, M. V., & Biderman, M. D. (1984). Narcissism and empathy: Validity evidence for the Narcissistic Personality Inventory. *Journal of Personality Assessment, 48*, 301–305. https://doi.org/10.1207/s15327752jpa4803_12

Watson, P. J., Hood, R. W., & Morris, R. J. (1984). Religious orientation, humanistic values, and narcissism. *Review of Religious Research, 25*(3), 257–264. https://doi.org/10.2307/3511123

Watson, P. J., Hood, R. W., Morris, R. J., & Hall, J. R. (1986). The relationship between religiosity and narcissism. *Counseling and Values, 31*(2), 179–184. https://doi.org/10.1002/j.2161-007x.1986.tb00490.x

Watson, P. J., & Morris, R. J. (1991). Narcissism, empathy and social desirability. *Personality and Individual Differences, 12*(6), 575–579. https://doi.org/10.1016/0191-8869(91)90253-8

Wearden, A. J., Lamberton, N., Crook, N., & Walsh, V. (2005). Adult attachment, alexithymia, and symptom reporting: An extension to the four category model of attachment. *Journal of Psychosomatic Research, 58*(3), 279–288. https://doi.org/10.1016/j.jpsychores.2004.09.010

Wink, P. (1991). Two faces of narcissism. *Journal of Personality and Social Psychology, 61*(4), 590–597. https://doi.org/10.1037/0022-3514.61.4.590

Wisdom, N. M., Pastorek, N. J., Miller, B. I., Booth, J. E., Romesser, J. M., Linck, J. F., & Sim, A. H. (2014). PTSD and cognitive functioning: Importance of including performance validity testing. *The Clinical Neuropsychologist, 28*(1), 128–145. https://doi.org/10.1080/13854046.2013.863977

Wu, K., Chen, C., & Greenberger, E. (2019). Nice guys and gals can finish first: Personality and speed-dating success among Asian Americans. *Journal of Social and Personal Relationships, 36*(6), 2507–2527. https://doi.org/10.1177/0265407518790103

Yanez, Y., Fremouw, W., Tennant, J., Strunk, J., & Coker, K. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology, 21*(2), 161–165. https://doi.org/10.1016/j.acn.2005.07.009

Young, G. (2015). Malingering in forensic disability-related assessments: Prevalence 15 ± 15 %. *Psychological Injury and Law, 8*(3), 188–199. https://doi.org/10.1007/s12207-015-9232-4

Young, S., Jacobson, R., Einzig, S., Gray, K., & Gudjonsson, G. H. (2016). Can we recognise malingerers? The association between malingering, personality traits and clinical impression among complainants in civil compensation cases. *Personality and Individual Differences, 98*, 235–238. https://doi.org/10.1016/j.paid.2016.04.052

Yu, C. K. -C. (2018). Pathological narcissism, dream experiences, and personality dynamics. *Dreaming, 28*(4), 322–336. https://doi.org/10.1037/drm0000087.

Zeigler-Hill, V., & Wallace, M. T. (2011). Racial differences in narcissistic tendencies. *Journal of Research in Personality, 45*(5), 456–467. https://doi.org/10.1016/j.jrp.2011.06.001