Psychology of Consciousness: Theory, Research, and Practice

The Dissociative Experiences Scale: An Empirical Evaluation of Long-Standing Concerns

Mayson Trujillo, Alan Brown, David Watson, Katie Croft-Caderao, and Michael Chmielewski

The Dissociative Experiences Scale: An Empirical Evaluation of Long-Standing Concerns

Mayson Trujillo[1], Alan Brown[1], David Watson[2], Katie Croft-Caderao[3], and Michael Chmielewski[1]

[1] Department of Psychology, Southern Methodist University
[2] Department of Psychology, University of Notre Dame
[3] Dallas CBT, Dallas, Texas, United States

It is imperative that psychological measures demonstrate strong psychometric properties in order to increase study replicability, develop an accurate understanding of constructs, identify potential mechanisms, and accurately determine treatment efficacy. The Dissociative Experiences Scale (DES) is the most widely used measure of dissociation. However, concerns have been raised about the DES's response format and items. In addition, the measure has demonstrated poorer dependability (e.g., short-term test–retest reliability) than other dissociation measures. The current research examines these issues across two studies. The goal of Study 1 ($N = 163$ undergraduates) was to empirically test concerns regarding the DES's response format and items. Participants' responses to the DES using the standard response format did not align with their subsequent estimates of how frequently those items occurred. Moreover, participants often did not interpret the DES items in the way intended by the measure. In Study 2 ($N = 447$ undergraduates, 2-week retest interval), we attempted to improve the dependability of the DES by changing the standard DES's response format without substantially altering its items. Changing the response format did not appear to improve the dependability of the DES, suggesting other features of the measure are responsible for its poor dependability. In conclusion, the present studies provide empirical evidence for concerns about the DES's psychometric properties and indicate the DES demonstrates low reliability that appears to result, in part, from item wording.

Keywords: Dissociative Experiences Scale, psychometrics, reliability, dependability, retest

Supplemental materials: https://doi.org/10.1037/cns0000334.supp

Cited over 7,300 times (Google Scholar) and translated into at least a dozen languages, the Dissociative Experiences Scale/Dissociative Experiences Scale-II[1] (DES; Bernstein & Putnam, 1986; DES-II, Carlson & Putnam, 1993) is by far the most widely used self-report measure of dissociation. It has been used in clinical, community, and student

[1] The DES-II uses an 11-point Likert-type response scale, whereas the original DES employed a VAS. The items, instructions, and response range (0%–100%) are identical on both forms. Given their similarities, both measures will generally be referred to as the DES throughout the article. However, distinctions will be made between the DES-I and DES-II when discussing previous psychometric studies.

samples to investigate the association between trait dissociation and numerous constructs, including personality (e.g., Groth-Marnat & Jeffs, 2002; Kwapil et al., 2002), trauma (e.g., Dalenberg et al., 2012), and psychopathology (e.g., Lyssenko et al., 2018). Dissociation is the defining process underlying dissociative disorders; however, it is also a symptom of several other diagnoses from the *Diagnostic and Statistical Manual of Mental Disorders, 5th edtion* (*DSM-5*; American Psychiatric Association, 2013), including panic disorder, posttraumatic stress disorder, and borderline personality disorder. Moreover, dissociation is associated with many forms of internalizing (e.g., major depressive disorder, generalized anxiety disorder, obsessive–compulsive disorder), externalizing (e.g., alcohol use disorder, gambling use disorder), and psychotic (e.g., schizophrenia) psychopathology (e.g., Ellickson-Larew et al., 2020; Lyssenko et al., 2018). Given the DES's widespread use and its substantial influence on the scientific knowledge base of dissociation, the psychometric properties of the DES are of critical importance. Indeed, the ability to identify mechanisms underlying dissociation (i.e., genetic, biological, environmental), evaluate treatment and intervention efficacy, determine the temporal stability of dissociation, and trust the replicability of links between dissociation and other constructs, all depend on the reliability and validity of the DES (Chmielewski et al., 2017; Chmielewski & Watson, 2009; Pargent et al., 2019; Schmidt et al., 2003).

Concerns About the DES's Psychometric Properties

Despite the DES's widespread use, there are numerous long-standing concerns regarding its psychometric properties (Chmielewski, 2022; Dunn et al., 1993; Ellickson-Larew et al., 2020; Frueh et al., 1996; Goldberg, 1999; Paolo et al., 1993; Wright & Loftus, 1999). One concern is the DES's response format, which instructs respondents to indicate the percentage of time they experience the scenarios described by each item. As Wright and Loftus (1999) note, a "10%" response to the item "someone listens to another talk and does not hear all or part of the conversation" could refer to 10% of waking hours, 10% of time spent in conversation, or 10% of some other (and unknown) metric. Therefore, it is possible that one participant may be interpreting the response

format differently from other participants. Additionally, concerns regarding the factor structure of the DES have also been expressed (Chmielewski, 2022; Ellickson-Larew et al., 2020). Bernstein and Putnam (1986) did not develop the DES to have subscales; however, they later created three subscales (Carlson & Putnam, 1993). Although the three subscales are widely used in the literature, they are not consistently replicated (Bernstein et al., 2001; Dunn et al., 1994; Zingrone & Alvarado, 2001), leading to concerns about the stability of the DES's factor structure (Chmielewski, 2022; Ellickson-Larew et al., 2020).

Concerns also have been raised regarding the interpretability of the DES's items. Indeed, even the authors of the DES-II (Carlson & Putnam, 1993) warned, there might be variation in the interpretation of item content and recommended follow-up clarification for all endorsed items in clinical settings. Although the DES-II was created to assess dissociation in clinical samples, it has frequently been used in nonclinical samples. Thus, clarifying the responses of nonclinical individuals in research settings may be particularly important. Subsequent findings from Paolo et al. (1993) indicated the DES's items are written at a 10th- to 11th-grade reading level. This is substantially higher than the average reading level in the United States (Atcherson et al., 2013; Weiss & Coyne, 1997). It is also in contrast to recommendations that items be written so that even participants with low levels of education can understand them (Clark & Watson, 1995, 2019; Weiss & Coyne, 1997). Additionally, Goldberg (1999) noted that the length and redundant phrasing of DES items (e.g., "Some people sometimes have the experience of feeling as though they are standing next to themselves or watching themselves do something as if they were looking at another person. Select a number to show what percentage of the time this happens to you") may cause fatigue and lead to item skimming. To address these concerns, Goldberg created the Curious Experiences Scale (CES; Goldberg, 1999), which is a revised version of the DES containing simplified DES items, additional new items, and a 5-point Likert scale.

As noted above, substantial concerns have been expressed regarding the DES's response format and items. However, the aforementioned concerns have not been tested empirically. If the response format and items negatively influence the measurement of dissociation, then they may

muddy the meaning of the scale's scores and reduce its reliability and validity. Below we discuss other potential psychometric limitations of the DES.

Reliability of the DES

Multiple types of reliability exist, each estimated using a different indicator. Two of the most pertinent with regard to the DES are Cronbach's α (i.e., consistency across items or internal consistency: Cortina, 1993; Cronbach, 1951) and dependability (i.e., short-term test–retest reliability: Cattell et al., 1970; Chmielewski & Watson, 2009; Gnambs, 2014; McCrae et al., 2011; Watson, 2004).

Cronbach's Alpha

By far, the most widely reported indicator of reliability is Cronbach's α, which is influenced by (a) the interrelatedness of items (i.e., the extent to which they covary internal consistency) and (b) the number of items in a scale (Clark & Watson, 1995; Cortina, 1993). Alphas exceeding .70 are typically considered adequate with greater than .80 being optimal (Clark & Watson, 1995; Cortina, 1993). Previous research has found high αs for the DES (i.e., total score: .91–.96; subscales: .77–.93; Chmielewski, 2022; Watson, 2003). However, because α is based on a single assessment session, it is unable to detect transient measurement error (i.e., fluctuations in a respondent's temporary, feelings) that differentially affects responses across assessment occasions (Becker, 2000; Chmielewski & Watson, 2009; Schmidt et al., 2003). Critically, α misspecifies transient error as true score variance, causing it to overestimate reliability estimates when based on a single occasion. In addition, because transient error can create the appearance of true change when none occurred, transient error reduces the validity of measures assessing trait-like constructs. Because α only assesses consistency across items during a single assessment, it is unable to detect the extent to which a measure is susceptible to these transient errors (Chmielewski et al., 2016; Schmidt et al., 2003; Watson, 2004).

Dependability

To determine the extent to which a measure is susceptible to transient error, short-term test–retest studies (i.e., dependability analyses) are necessary (Chmielewski et al., 2016; Chmielewski & Watson, 2009; Gnambs, 2014; McCrae et al., 2011; Watson, 2004). In contrast to typical retest studies, which often examine consistency over intervals during which both transient error and true changes in the construct are likely, dependability studies employ short retest intervals over which true trait-level change in a construct (e.g., trait dissociation) is unlikely to occur. Predominant theories of dissociation such as the trauma model and sociocognitive model predict that virtually no true changes in trait levels of dissociation should occur across a typical 2-week interval for adults. Thus, instability in the dependability of trait dissociation measures over such a time frame most likely is attributable to measurement error (Chmielewski et al., 2016; Gnambs, 2014; McCrae et al., 2011; Schmidt et al., 2003; Watson, 2004).

Dependability is essential for the construct validity of trait-like constructs (Chmielewski et al., 2017; Chmielewski & Trujillo, 2020; Chmielewski & Watson, 2009; Gnambs, 2014; McCrae et al., 2011; Suzuki et al., 2017). Measures of trait or trait-like constructs (including dissociation) with high levels of transient error, by definition, cannot validly assess the construct in the manner it is conceptualized (Chmielewski et al., 2016, 2017; McCrae et al., 2011; Watson, 2004). Moreover, there is now considerable evidence that dependability is, in part, a property of a measure, in that some scales consistently demonstrate higher levels of dependability than others (Chmielewski et al., 2016; Chmielewski & Watson, 2009; Gnambs, 2014; McCrae et al., 2011). Thus, dependability is required for reliable and valid measurement of traits and trait-like constructs (Anusic et al., 2012; Chmielewski et al., 2016; Chmielewski & Trujillo, 2020; McCrae et al., 2011; Watson, 2004).

Measures immune to transient and other forms of error would have dependability coefficients of essentially 1.0; however, given the ubiquity of measurement error, achieving perfect dependability is highly unlikely (Anusic et al., 2012; Chmielewski et al., 2016; Chmielewski & Trujillo, 2020; Watson, 2004). In order to identify acceptable dependability levels, studies of dependability should include reference "benchmark" measures in the same data collection (Chmielewski et al., 2016; Chmielewski & Trujillo, 2020; Watson, 2004). Ideally, these benchmark measures would include alternative measures of the same or similar constructs, as well as established measures of other traits that should not

demonstrate true change over the dependability interval (Chmielewski et al., 2016; Watson, 2004). The DES, as well as the majority of other dissociation measures, explicitly assesses dissociation as a stable and trait-like construct; therefore, measures of personality traits could serve as suitable benchmarks.

Studies of the DES have reported a wide range of dependability estimates. One potential reason for the mixed results is that most studies have used samples that are too small to provide precise estimates. In order to have sufficient precision, and narrow confidence intervals around dependability estimates, large samples ($N > 300$) are required (Watson, 2004). In their study on the development of the DES-I, Bernstein and Putnam (1986) obtained a dependability correlation of .84 across 4–8 weeks in a nonclinical sample of only 26 participants. In a sample of this size, the 95% confidence interval would range from only .67, indicative of a high level of error, to .93, indicating low levels of transient error. Similarly, small samples were employed by Sanders (1992), who obtained correlations of .90 ($N = 16$) across 11 weeks and .79 ($N = 46$) across 6–8 weeks for the DES-I in students. Likewise, Dubester and Braun (1995; $N = 78$; 2-week interval) and Frischholz et al. (1990; $N = 30$; 4-week interval) reported correlations of .93 for the DES-I in samples of inpatients primarily diagnosed with a dissociative disorder.

In addition to small samples, none of the four studies noted above included benchmark measures, further hindering an evaluation of the DES's dependability. In contrast, Watson (2003), using a large undergraduate sample that included benchmark measures, reported an average DES-II dependability of only .66 over 2 months, indicating a substantial level of error. Moreover, results also indicated the DES-II was significantly less dependable than two benchmark measures, the Dissociative Processes Scale (DPS; Harrison & Watson, 1992; mean dependability $r = .79$) and the Big Five Inventory (BFI; John & Srivastava, 1999; mean $r = .83$). It is important to note that the majority of the above studies employed retest intervals of 4 weeks or greater. As such, it could be argued that true change in trait levels of dissociation was perhaps possible, making their results somewhat ambiguous. Recently, Chmielewski (2022) found the DES-II produced an average retest correlation of .67 over 1 week in a sample of 340 undergraduates, which was substantially less dependable than the BFI (mean $r = .86$)

and the DPS (mean $r = .76$). Furthermore, there is mixed evidence suggesting the DES may be less dependable than the CES (Chmielewski, 2022). These results indicate the DES may demonstrate poor dependability relative to benchmark measures. Taken together, the accumulated evidence suggests the DES's response format and/or items may have a negative impact on its dependability. However, it is unclear what specific characteristics of the DES are responsible for its increased susceptibility to transient error.

The Present Studies

The present studies aim to provide information regarding the psychometric properties of the DES-II. Specifically, the current research empirically tests previous concerns regarding the ambiguity of the DES-II response format, the complexity of its items, and potential causes of its poor dependability. Study 1 examined (a) the consistency of use of the DES-II's response format across participants and (b) whether participants interpret the DES-II's items as intended (i.e., assessing dissociation). Study 2 tested whether modifying the DES-II's response format without substantially altering its items could improve its dependability. Both studies were approved by their respective institutional review boards.

Study 1

Method

Participants and Procedure

One hundred sixty-three undergraduate students (122 females, 41 males) enrolled in psychology courses at a private southern university completed the DES-II. Two weeks later, as recommended by Carlson and Putnam (1993), they completed a "clarification survey," for which they (a) estimated the frequency of each previously endorsed DES-II item and (b) provided a personal experience reflecting each item's content. Participants completed both the DES-II and clarification survey online and earned class credit for their participation.

Measures

The *DES-II* (Carlson & Putnam, 1993) is a 28-item "trait measure" (Carlson & Putnam, 1993, p. 16) of dissociation with three subscales

(Amnestic Dissociation, Depersonalization, and Absorption). Respondents estimate the percentage of time they experience each statement using an 11-point response scale (0%–100% with 10% increments).

The *clarification survey* was created for the present study. It was individualized so that each participant only received the specific DES-II items they had previously endorsed (i.e., indicated happened at least 10% of the time). For each item, the participant provided two responses. First, participants estimated the experiential frequency of each item with respect to a reference period (i.e., they chose per day, week, month, year, or lifetime). For example, a participant who frequently has out of body experiences might indicate they experience them once a *week*, whereas an individual who rarely experiences them might report that they occur once a *year* (see Supplemental Material A for an example item from the clarification survey). To standardize the time frame across items and respondents, all frequencies were converted to a 1-month basis after the data were collected. For instance, once a week would convert to "4" and once a year would be "1/12, or .08."

For the second response, participants were instructed to provide a personal example of the experience, including as much specific content as they could remember. The descriptions were used to determine each participant's interpretation of the items. Participants' personal experiences were coded by three independent raters, who were trained using a detailed manual and attended regular coding meetings. A response was classified as *dissociation* if it described involuntary changes into a different state of awareness and *nondissociation* if the participant described a nondissociative experience (indicating that they misunderstood the item), denied experiencing a previously endorsed item, or did not provide a description for that item. Interrater reliability among the three raters was .68 (Krippendorff's α) across all items for the present study, indicating an acceptable level of reliability that allows for drawing tentative conclusions (Krippendorff, 2004).

Results

Consistency of Response Scale Interpretations Across Participants

Table 1 presents the correlations between each DES-II item (standard instructions and response

Table 1

Correlations Between Responses for Each Item on the DES-II and the Number of Times That Item Was Experienced per Month From Clarification Survey

Item	N	r
1	109	.36**
2	158	.27**
3	61	.16
4	25	.26
5	62	.34**
6	88	.20
7	51	.43**
8	22	−.07
9	51	.49**
10	91	.01
11	16	.40
12	51	.38*
13	42	.05
14	123	.39**
15	131	.27**
16	85	.36**
17	129	.36**
18	95	.22*
19	122	.18
20	135	.15
21	110	.23*
22	95	.31**
23	125	.31**
24	134	.33**
25	95	.08
26	63	.18
27	29	.36
28	37	.05

Note. Ns = 16–158. Participants completed the clarification survey 2 weeks after completing the DES-II. The clarification survey only contained DES-II items from Time 1 that participants endorsed as occurring 10% of the time or greater. As a result, sample sizes varied across items. DES-II = Dissociative Experiences Scale-II.
* $p < .05$.   ** $p < .001$.

format) and the identical item from the clarification survey (i.e., number of times each item was experienced). Correlations tended to be weak, ranging from −.07 to .49, with a mean correlation of only .25.[2] Moreover, when calculated within their respective subscales, the average DES-II and clarification survey item correlations were .29 (Absorption), .28 (Depersonalization/Derealization), and .14 (Amnestic Dissociation). This weak convergence suggests substantial differences between participants in interpreting the DES-II response scale.

_____

[2] As a point of reference, the average 2-week dependability correlation of the BFI items in Study 2 was $r = .61$.

Participant Personal Experience (i.e., Interpretation) of Items

Next, we examined the personal experiences provided by participants for each item. The percentage of responses classified as dissociation and nondissociation is presented in Table 2. It is striking that participants rarely provided actual dissociative experiences, with only 22% of experiences coded as such. The percentage of dissociative responses did vary across subscales: Absorption = 31%, Depersonalization/Derealization = 14%, Amnestic Dissociation = 11%. There was also considerable variability at the item level. Participants frequently provided experiences considered dissociative for missing all or part of a conversation (Item 2, 96%), being unaware of surroundings while watching television (Item 17, 86%), and highway hypnosis (Item 1, 82%). It is worth noting, however, that these were the only items with the majority of responses coded as dissociative. Moreover, for over half of the DES-II items (15/28, or 54%), fewer than 10% of responses were coded as dissociative. Notably, participants also provided nondissociative experiences 78% of the time. Nondissociative experiences were most frequently provided for items comprising the Amnestic Dissociation subscale (89%), followed by Depersonalization/ Derealization (86%) and Absorption (69%). In fact, for the majority of items (25/28; 89%), nondissociative experiences were more frequent than dissociation experiences (see Table 2).

Study 1 Summary

The results of Study 1 provide empirical support for previous concerns regarding the interpretability of the DES-II response format and item content (Goldberg, 1999; Paolo et al., 1993; Wright & Loftus, 1999). There was only a weak association between participants' responses to DES-II items using the standard response format and the number of times they indicated experiencing each item. This outcome suggests the DES-II's response scale may not accurately reflect the actual frequency of dissociative experiences. It also supports Wright and Loftus's (1999) concern that the ambiguous reference period in the DES instructions may lead to inconsistencies across participants in how the response format is interpreted.

Study 1 also addressed concerns that it may be difficult for participants to understand the lengthy and complex wording of the DES items (Goldberg, 1999; Paolo et al., 1993). The current results support these concerns, as only 22% of personal experiences were coded as dissociative, whereas 78% were coded as nondissociative experiences. Interestingly, participants were most likely to provide examples of dissociation for items comprising the Absorption subscale (31%). Compared to the two other subscales, Absorption encompasses more normative dissociative experiences (Waller et al., 1996), which may explain why participants appear to have been more accurate in interpreting these items.

Together, inconsistent use of the DES-II's response format across participants and inaccurate interpretation of the DES-II's items may negatively affect its reliability and validity. This is consistent with previous research on the CES, which demonstrated improved psychometric properties compared to the DES-II in some samples after modifying the DES-II's response format and item wording (Chmielewski, 2022). However, because the CES altered the response format, items, and instructions, it is unclear which specific changes produced these improvements in reliability. In Study 2, the DES-II's response format and instructions were modified but the items were left unchanged (except when necessary to match the revised response format) to examine the impact of these modifications on its dependability.

Study 2

Method

Participants and Procedure

Four hundred forty-seven undergraduate students (377 females, 70 males) enrolled in psychology courses at a large public Midwestern university completed a modified version of the DES-II and two benchmark instruments: a second trait measure of dissociation and a measure of the Big Five personality traits. Two weeks later, they completed all measures a second time. Because the modified version of the DES-II and the original DES-II contain identical items, we did not have participants complete the original DES-II during the current data collection. Therefore, in addition to the benchmark measures included in the present study, we also compare results for the modified DES-II to (a) the 1-week dependability estimates

Table 2
Percentages of Dissociative and Nondissociative Responses of the Clarification Survey Categorized by Subscale

Subscale and item	N	Paraphrased item	Percent dissociation	Percent nondissociation
Depersonalization/derealization			M = 14	M = 86
28	37	Foggy reality	27	73
7	51	Out of body experiences	19	81
12	51	Environment did not feel real	18	82
13	42	Felt body was not their own	7	93
27	29	Heard voices	7	93
11	16	Didn't recognize self in mirror	6	94
Amnestic dissociation			M = 11	M = 89
3	61	Unaware of how you got to a location	37	63
4	25	Doesn't remember dressing	21	79
25	95	Forgotten they had done something	7	93
6	88	Stranger mistakes you for another	6	94
26	63	Forgotten they had drawn something	6	94
8	22	Did not recognize friends/family	5	95
10	91	Accused of lying when telling truth	4	96
5	62	Finds objects doesn't remember purchasing	3	97
Absorption			M = 31	M = 69
2	158	Missed part of a conversation	96	4
17	129	Unaware of surroundings while watching TV	86	14
15	131	Uncertain if events were real or dreamt	37	63
20	135	Zoned out unaware time had passed	34	66
23	125	Does difficult things easily in certain situations	11	89
14	123	Vivid remembrance of events	7	93
18	95	Confused fantasy with reality	6	94
22	95	Acted different depending on situation	2	98
16	85	Experienced a familiar place as foreign	2	98
No subscale			M = 30	M = 70
1	109	Highway hypnosis	82	18
19	122	Ignored pain	48	52
9	51	Amnesia for significant life events	10	90
24	134	Uncertain if completed or thought about completing something	8	92
21	110	Talked to self when alone	0	100
		Total	M = 22	M = 78

Note. Ns = 16–158. The clarification survey contained DES-II items that participants endorsed as occurring 10% of the time or greater during Time 1. Participants who did not provide a personal example were classified as providing a "nondissociation" response. Percentage of dissociative responses ranged from 0 to 96 and the nondissociative responses ranged from 4 to 100. DES-II = Dissociative Experiences Scale-II.

for the original DES-II from Chmielewski (2022) and (b) the 2-month estimates from Watson (2003). The dependability correlations from the aforementioned studies were derived from large undergraduate samples (1-week sample: $N = 340$; 2-month sample: $N = 465$).

Measures

The *Modified Dissociative Experiences Scale-II (M-DES-II)* is a modified version of the DES-II created for the present study. Instead of the traditional DES-II instructions and response format, which ask participants to indicate the percentage of time they experience each item on an 11-point response scale (0%–100% with 10% increments), the M-DES-II asks participants to indicate the extent to which each statement applies to them on a 5-point scale, 1 (A) = *not at all* to 5 (E) = *very well*, which is a response scale commonly used in personality measures (see Supplemental Material B). Additionally, these instructions and response format differ from those of the CES, which instructs participants to report how often they experience dissociative experiences on 5-point scale from 1 (*this never happens to me*) to 5 (*this is almost always happening to me*).

Items on the M-DES-II were nearly identical to the original DES-II; however, some items were adjusted to ensure compatibility with the new response format. For example, the item "Some people have the experience of feeling that other people, objects, and the world around them are not real" became "I sometimes feel that other people, objects, and the world around me are not real." Furthermore, the instructions after each DES-II item ("Select a number to show what percentage of the time this happens to you") were eliminated. No other modifications were made to the DES-II. Thus, like the DES-II, the M-DES-II is a trait measure of dissociation. Critically, the M-DES-II items were more faithful to the DES-II than the CES, which modified the original items and added three non-DES-II items. Therefore, the M-DES-II will test the effect of modifying the DES-II's instructions and response format on its dependability, while keeping the item content relatively unchanged.

The *DPS* (Harrison & Watson, 1992) is a 33-item factor analytically derived measure of trait dissociation containing three subscales: Obliviousness, Detachment, and Imagination. Participants read each statement (e.g., "Sometimes the things around me do not seem quite real") and indicate their agreement on a 5-point scale from 1 (*strongly disagree*) to 5 (*strongly agree*). The DPS has demonstrated good internal consistency with αs ranging from .93 to .94 and .85 to .89 for the total and subscale scores, respectively (Watson, 2003). The DPS total score has demonstrated dependability coefficients of $r = .79$ over 1-week (Chmielewski, 2022) and $r = .81$ over 2-month (Watson, 2003) intervals in large samples of 340 and 465 undergraduates, respectively. The DPS has been used to examine how dissociation relates to the Big Five personality

traits (Watson et al., 2015b; Watson et al., 2019) and to odd and unusual experiences (anomalous sleep experiences, Watson et al., 2015a; schizotypy, Chmielewski & Watson, 2008; schizotypy and psychological trauma, Berenbaum et al., 2008). It has also been used to examine the construct and convergent validity of measures of anomalous experiences (Cicero et al., 2017), traumatic symptoms (Gootzeit et al., 2015), and psychosis proneness (Cicero et al., 2010).

In two large independent samples, the DPS and the DES-II total scores correlated .58 and .61 and formed a single dimension in confirmatory factor analyses with another dissociation measure (Watson, 2001). However, the subscales of the two measures divide the dissociation domain somewhat differently and do not form a clear convergent/ discriminant pattern with one another. Similarly, a clear convergent/discriminant pattern was not found between the M-DES-II and DPS subscales in the present study (see Table 3). Therefore, we followed the procedure of previous studies and compared the dependability coefficients of the total scores and all subscale scores to each other rather than examining specific subscale pairings.

The *Big Five Inventory* (BFI; John & Srivastava, 1999) is a widely used 44-item factor analytically derived measure of personality traits containing five scales (Neuroticism, Extraversion, Openness, Agreeableness, Conscientiousness). The items each contain a stem statement ("I see myself as someone who . . . ," which is provided once in the instructions) followed by a description (e.g., "is talkative," "can be somewhat careless," "is inventive"). Participants indicate their agreement with each statement using a 5-point scale ranging from 1 (*disagree strongly*) to 5 (*strongly agree*). The BFI has consistently demonstrated good validity and reliability and has been used as

Table 3

Convergent Correlations Between the DPS and M-DES-II in Study 2

	DPS							
	Total		Obliviousness		Imagination		Detachment	
M-DES-II	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Total	.70	.76	.65	.70	.48	.52	.54	.68
Amnestic dissociation	.52	.58	.53	.57	.28	.33	.41	.53
Absorption	.71	.79	.63	.73	.59	.62	.52	.64
Depersonalization/derealization	.53	.55	.43	.42	.35	.37	.55	.70

Note. $N = 447$ Time 1 and Time 2 assessments were 2 weeks apart. All correlations are significant at $p < .01$, two-tailed. DPS = Dissociative Processes Scale; M-DES-II = Modified Dissociative Experiences Scale-II.

a benchmark in previous dependability studies (Chmielewski, 2022; Chmielewski et al., 2016; Chmielewski & Watson, 2009; Watson, 2003; Watson et al., 2015; Watson & Wu, 2005).

Results

Internal Consistency

Alpha coefficients for the measures in the present study, as well as those for the DES-II from comparison studies (Chmielewski, 2022; Watson, 2004) are presented in Table 4. The αs of the M-DES-II, DPS, DES-II, and BFI exceeded .75, indicating all measures demonstrated adequate internal consistency.

Dependability

The dependability coefficients for the M-DES-II, DPS, and BFI from the present study and the DES-II from previous studies are presented in the third column of Table 4. Differences between the M-DES-II, DPS, and BFI dependability coefficients were examined using Pearson–Filon tests. Individual Pearson–Filon comparisons are reported in Supplemental Tables 1–3. The M-DES-II total score was significantly less dependable than the DPS total score ($p < .01$). Moreover, the M-DES-II subscales were significantly less dependable than the DPS subscales in six of nine (67%) comparisons; the DPS subscales were never less dependable than the M-DES-II subscales. In addition, the M-DES-II total score and subscales were significantly less dependable than the benchmark BFI scales in 18 of 20 (90%) comparisons. It is also noteworthy that the DPS total score and subscales performed much better, being significantly less dependable than the BFI in only 20% of comparisons and more dependable than the BFI in 10% of comparisons.

Differences between the current M-DES-II and the DES-II data from past studies were examined using Fisher's Z-transformations. There were no differences in the dependability of the M-DES-II total and subscales scores compared to their counterparts on the DES-II across 1-week (Chmielewski, 2022) or 2-month (Watson, 2004) intervals.

Study 2 Summary

Study 2 examined the extent to which modifying the DES-II's response format while retaining

Table 4

Internal Consistency and Dependability

Scale	Coefficient α reliability		Dependability
	Time 1	Time 2	
M-DES			
Total score	.93	.95	.75
Amnesia	.84	.87	.67
Depersonalization	.84	.89	.67
Absorption	.86	.88	.73
DPS			
Total score	.93	.95	.84
Obliviousness	.86	.90	.82
Detachment	.85	.89	.72
Imagination	.83	.87	.84
BFI			
Neuroticism	.82	.84	.83
Extraversion	.85	.87	.83
Openness	.82	.84	.84
Agreeableness	.81	.83	.78
Conscientiousness	.79	.81	.81
DES-II			
Total score	.96 (.91)	.96 (.94)	.69 (.69)
Amnesia	.92 (.80)	.93 (.87)	.64 (.62)
Depersonalization	.91 (.77)	.93 (.89)	.63 (.63)
Absorption	.91 (.84)	.93 (.88)	.70 (.68)

Note. $N = 447$ for all 2-week M-DES, DPS, and BFI data. For the 1-week DES-II data, $N = 555$ and $N = 383$ for the Time 1 and Time 2 coefficient αs, respectively, and $N = 340$ for the dependability coefficients (coefficients outside of the parentheses). One-week data were adapted from Chmielewski (2022). $N = 465$ for all 2-month DES-II data (coefficients inside of the parentheses). Two-month DES-II data were adapted from Watson (2004). All correlations are significant at $p < .01$, two-tailed. M-DES = Modified Dissociative Experiences Scale; DPS = Dissociative Processes Scale; BFI = Big Five Inventory; DES-II = Dissociative Experiences Scale-II.

its original items would improve the measure's dependability. These modifications to the DES-II were unsuccessful in reducing its susceptibility to transient error, suggesting that sources of error beyond the DES-II's response format are negatively affecting its dependability. Moreover, our results clearly demonstrate the construct of dissociation can be assessed with a high level of dependability, as the DPS (an alternative measure of dissociation) demonstrated a level of dependability in line with the benchmark BFI.

Discussion

The current research provides empirical evidence supporting previous concerns regarding the scientific utility of the DES-II, with respect

to response format, item content, and dependability. Study 1 demonstrated a weak association between responses on the DES-II and participants' estimates of how often they experienced each item, supporting previous concerns that the DES-II's ambiguous response format may lead to inconsistent interpretations across participants (Wright & Loftus, 1999). Moreover, participants' examples of relevant personal experiences were rarely dissociative, confirming concerns that the wording of DES-II items may be difficult to understand (Goldberg, 1999; Paolo et al., 1993). In Study 2, we attempted to improve the poor dependability of the DES-II (see Chmielewski, 2022; Watson, 2004) by modifying its response format while keeping its items the same. This modification was unsuccessful, indicating factors beyond the DES-II's response format, such as its highly complex items, are likely responsible for its low dependability. We discuss the results from both studies below.

Participants' Understanding of the DES-II Response Format and Items

The DES-II's response format was intended to assess the proportion of time participants experience the scenario depicted in each item. However, there was only a weak association between participants' responses on the DES-II and subsequent estimates of how frequently the experience occurred (Absorption $r = .29$, Depersonalization/Derealization $r = .28$, Amnestic Dissociation $r = .14$). Wright and Loftus (1999) suggested the lack of reference period in the DES-II's instructions may lead to inconsistency across participants in how the response format is interpreted. The results of the current research support their concern and suggest substantial inaccuracy in how the DES reflects differences in the frequency of dissociative experiences across individuals.

The current research also empirically demonstrates that the DES-II's items are rarely interpreted as intended, as current participants appear to respond to the majority of items with nondissociative experiences in mind. One possible reason for these findings is the high level of cognitive effort required to comprehend the DES-II items. Goldberg (1999) raised concerns that the length and redundancy of the DES-II's items may lead participants to become fatigued and skim the items. Similarly, Paolo et al. (1993) warned

that the items may be difficult to interpret for those with less than a high school reading level, including clinical patients for whom the measure was originally designed. The current results using a highly educated college student sample indicate these concerns are not limited to individuals with low reading levels. Although the DES-II items were correctly interpreted as dissociation less than half of the time, there was variability across the subscales. Items comprising the Absorption subscale were most frequently interpreted as dissociative followed by Depersonalization/Derealization and Amnestic Dissociation. Interestingly, this pattern of variability mirrors that found for response format interpretation. These findings suggest some types of dissociative experiences may be easier to interpret and rate than others.

Together, the current results and previous literature suggest the DES-II may not measure dissociative experiences in an optimal manner. Specifically, inconsistent interpretation of the DES-II's response format and item content across participants may lead to inaccurate estimates of the frequency and nature of dissociative experiences. Moreover, such inconsistencies limit researchers' ability to compare participants and draw meaningful conclusions from DES-II scores. These limitations have considerable implications for dissociation research including inaccurate estimates of associations with other constructs, hindering the search for underlying mechanisms, and leading to false results regarding treatment effectiveness.

The Influence of Transient Error on the DES-II

Some research on the CES indicates altering both the DES's response format and items may increase measurement dependability, leading some to hypothesize the measure's response format may be one possible source of the DES-II's substantial amount of transient error (Chmielewski, 2022; Watson, 2004). Study 2 tested this hypothesis by changing the DES-II response format to a more standard one used by personality measures, while changing its item content as minimally as possible. Our modifications (M-DES-II) did not improve the dependability of the instrument. Indeed, the M-DES-II was less dependable than the DPS and BFI and did not differ from the DES-II. This suggests the complexity of the DES-II's items

(see Study 1) is likely contributing to its poor dependability by causing participants to become fatigued or confused, leading to guessing, skimming, or putting forth little effort on items (Goldberg, 1999).

As previously noted, the DES-II is conceptualized as a trait measure, and predominant theories of dissociation suggest there should be little to no true change in dissociation over the current test–retest interval. For example, the trauma model contends dissociation is experienced in response to traumatic events as a way to automatize behavior and reduce the experience of pain (Dalenberg et al., 2012; Gershuny & Thayer, 1999). The present studies were conducted across 2 typical weeks of the semester (i.e., no midterms, finals, spring breaks, weather disasters, or global pandemics), reducing the likelihood of large portions of the sample experiencing trauma over the retest interval. Furthermore, the level of childhood trauma would remain unchanged between testing in the current samples. Likewise, etiological factors suggested by the sociocognitive model, including fantasy proneness and suggestibility, are conceptualized as stable traits that should not change over the course of the retest interval (Dalenberg et al., 2012; Lilienfeld et al., 1999; Lynn et al., 2019).

True change in dissociation over this time frame is also inconsistent with hypothesized transtheoretical factors. For example, some have hypothesized that disrupted sleep patterns, which often follow traumatic experiences, may induce the integration of dreams and reality and lead to the fantasy-based thinking typical of dissociation (Lynn et al., 2019; van Heugten-van der Kloet et al., 2014). However, the possibility that the present participants experienced widespread changes in their sleep is low, given that the present studies were not conducted during unusually stressful life periods. Indeed, in the present studies, average dissociation scores did not systematically increase across the time points (see Table 5). Similarly, although research has demonstrated that dissociative experiences can be reduced with treatment (Lynch et al., 2008), means from the present studies do not suggest such reduction occurred. Critically, the DES was specifically designed to assess trait-like dissociation (Bernstein & Putnam, 1986; Carlson & Putnam, 1993), further emphasizing that true change should not have occurred over the retest interval and any observed instability is likely the result of error. Moreover, the DPS—an alternative measure of trait dissociation—was significantly

Table 5
M-DES, DPS, and DES-II Means

Scale	Means	
	Time 1	Time 2
M-DES		
Total score	24.33	24.62
Amnesia	4.94	5.03
Depersonalization	2.22	2.55
Absorption	11.56	11.23
DPS		
Total score	90.18	85.44
Obliviousness	41.51	39.31
Detachment	11.48	10.96
Imagination	20.94	19.71
DES-II		
Total score	47.07 (12.51)	42.23 (12.82)
Amnesia	8.78 (6.18)	8.05 (6.80)
Depersonalization	4.54 (4.01)	4.74 (4.48)
Absorption	22.78 (20.17)	19.82 (20.22)

Note. $N = 447$ for all 2-week M-DES and DPS. For the 1-week DES-II data, $N = 555$ and $N = 383$ for the Time 1 and Time 2 means, respectively. One-week data were adapted from Chmielewski (2022). $N = 465$ for all 2-month DES-II data (coefficients inside of the parentheses). Two-month DES-II data were adapted from Watson (2004). M-DES = Modified Dissociative Experiences Scale; DPS = Dissociative Processes Scale; DES-II = Dissociative Experiences Scale-II.

more dependable, further demonstrating that true change was minimal over this period.

The DES-II's poor dependability suggests but does not prove that participants' responses are highly influenced by transient errors, such as temporary moods, feelings, and other mental states (e.g., boredom, fatigue). Because these states frequently fluctuate, participants' responses on the DES-II—and, therefore, the DES-II's associations with other constructs—will change with them. As such, results from studies using the DES-II may be less likely to replicate. Additionally, fluctuations in responses may be misinterpreted as true change in dissociation in longitudinal or intervention studies, resulting in inaccurate conclusions about the stability of dissociation and the efficacy of treatment interventions.

Indeed, Study 2 demonstrated that the DES-II is susceptible to transient error, which affects responses to DES-II items *within* individuals over time. Additionally, transient error may contribute to the *between*-person differences found in Study 1. For example, it may be that transient states affect participants' interpretation of the DES-II's response format and items. Therefore, within an assessment occasion, differences in transient states between participants may contribute to

inconsistent interpretation of the DES-II. Further research is necessary to test this hypothesis. Ultimately, the high levels of error in the DES-II hinder researchers' ability to understand why people experience dissociation, identify treatment and intervention efficacy, and refine dissociation's nomological net.

Limitations and Future Directions

In Study 1, only the DES-II items each participant endorsed as occurring at least 10% of the time were represented in the clarification survey. Although doing so reduced participant burden, it may have limited our understanding of participants' interpretations. For example, if given the opportunity, participants may have reported experiencing previously unendorsed items (given a "0" rating response). Moreover, this design resulted in small samples for the infrequently endorsed items. As such, future studies aiming to understand participants' interpretations should include all DES items in both evaluations. Furthermore, the purpose of Study 1 was to examine participants' interpretation of the DES-II; however, participants' interpretation of items from other dissociation measures was not examined. Thus, it is unclear whether this level of inaccurate interpretation is unique to the DES-II or is demonstrated by other self-report measures of dissociation. However, the fact that the DPS scales were consistently more dependable suggests inaccurate interpretation may be less of an issue with the DPS. Future studies should examine the interpretation of other measures of pathological and nonpathological dissociation across a variety of samples to examine the generalizability of the present findings. Moreover, it is unclear whether inaccurate interpretation of dissociative experiences is unique to self-report measures or generalizable to clinician assessment of dissociative experiences. Previous research has demonstrated that self-reported psychotic experiences may demonstrate worse predictive validity than clinician-rated interviews, suggesting that respondents and clinicians may be interpreting symptoms differently (Fusar-Poli et al., 2017). Given the conceptual and empirical associations between psychosis/schizotypy and dissociation, a similar pattern of results may emerge for dissociative experiences (Chmielewski & Watson, 2008; Watson, 2001). Future studies should examine these issues.

In Study 2, to reduce the participants' burden of responding to the same items twice, dependability results for the M-DES-II were compared to DES-II dependability estimates from other studies using similar samples. However, results may have been different if the same participants completed both measures. Next, Study 2 only compared the DES-II and DPS on two indicators of reliability (viz., dependability and internal consistency). Future studies should compare the reliability and validity of the DES-II to other measures of dissociation. Furthermore, although the DES-II frequently has been used in nonclinical samples, it was initially developed to assess pathological dissociation (Carlson & Putnam, 1993). Therefore, future studies should compare the properties of the DES-II to alternative measures of pathological dissociation. Additionally, Studies 1 and 2 both relied on undergraduate samples that were predominantly females. There is some research suggesting dissociation at the subscale level may vary across gender in student and nonclinical samples, suggesting the present results may not generalize to other types of samples (Spitzer et al., 2003). However, it is important to note that although the DES-II was originally designed for use in clinical samples, it has frequently been used in undergraduate and community samples. Epidemiological research indicates dissociative experiences are common with 26%–74% of the general population experiencing depersonalization/derealization during their lifetime (Hunter et al., 2004).

It remains unclear how the dependability of the DES-II might be improved. Interestingly, previously reported dependability estimates for the DES-I, which used a Visual Analog Scale (VAS) instead of the DES-II's 11-point Likert response format, tended to be somewhat higher. However, these studies used small samples resulting in imprecise dependability estimates. Moreover, studies comparing the psychometric properties of VAS and Likert response formats suggest modest to no differences between them (Hilbert et al., 2016; Kuhlmann et al., 2017; Russell & Bobko, 1992; Simms et al., 2019). Although the present findings indicate the DES-II demonstrates poor dependability and a high susceptibility to transient error, the transient sources influencing the DES-II are unknown. It has been frequently hypothesized that fluctuations in mood, health, and fatigue may contribute transient error in self-report measures (Becker, 2000; Chmielewski & Watson, 2009; Gnambs, 2014,

2015; Green, 2003; Heggestad et al., 2006; Raykov & Penev, 2005; Reeve et al., 2005; Schmidt et al., 2003; Shaffer et al., 2016; Vispoel et al., 2018; Vispoel & Tao, 2013; Watson et al., 2015). Future studies should examine the extent to which these and other transient states impact the dependability of the DES and other trait measures.

Critically, both the current results and other studies (Chmielewski, 2022; Watson, 2004) demonstrate that trait-like dissociation can be measured with greater reliability and validity than is afforded by the DES-II. The present findings demonstrate that the DPS is more dependable than the DES-II and demonstrates dependability greater than or equal to the BFI—a benchmark measure of acceptable dependability—in the majority of scale comparisons. Similarly, results from Chmielewski (2022) indicate the CES may be more dependable than the DES-II in select samples. Therefore, we suggest researchers consider using other measures instead of the DES-II to assess dissociation.

Conclusion

The present studies provide empirical support for previous concerns regarding the DES-II, demonstrate that the measure is suboptimal and that the DES-II has substantial psychometric limitations. Indeed, the DES-II items are often interpreted inaccurately and its response format may elicit different interpretations across respondents. Moreover, the DES-II item wording apparently leads to lower dependability and higher levels of transient error, as changing the DES-II's response format failed to improve its dependability. Given these limitations, continued use of the DES-II may hinder the ability to accurately identify mechanisms eliciting dissociation, evaluate treatment and intervention efficacy, and understand the association of dissociation with other constructs.

References

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

Anusic, I., Lucas, R. E., & Donnellan, M. B. (2012). Dependability of personality, life satisfaction, and affect in short-term longitudinal data. *Journal of Personality*, *80*(1), 33–58. https://doi.org/10.1111/j.1467-6494.2011.00714.x

Atcherson, S. R., Richburg, C. M., Zraick, R. I., & George, C. M. (2013). Readability of questionnaires assessing listening difficulties associated with (central) auditory processing disorders. *Language, Speech, and Hearing Services in Schools*, *44*(1), 48–60. https://doi.org/10.1044/0161-1461(2012/11-0055)

Becker, G. (2000). How important is transient error in estimating reliability? Going beyond simulation studies. *Psychological Methods*, *5*(3), 370–379. https://doi.org/10.1037/1082-989X.5.3.370

Berenbaum, H., Thompson, R. J., Milanek, M. E., Boden, M. T., & Bredemeier, K. (2008). Psychological trauma and schizotypal personality disorder. *Journal of Abnormal Psychology*, *117*(3), 502–519. https://doi.org/10.1037/0021-843X.117.3.502

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

Bernstein, I. H., Ellason, J. W., Ross, C. A., & Vanderlinden, J. (2001). On the dimensionalities of the Dissociative Experiences Scale (DES) and the Dissociation Questionnaire (DIS-Q). *Journal of Trauma & Dissociation*, *2*(3), 103–123. https://doi.org/10.1300/J229v02n03_07

Carlson, E. B., & Putnam, F. W. (1993). An update on the Dissociative Experiences Scale. *Dissociation*, *6*(1), 16–27.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF)*. Institute for Personality and Ability Testing.

Chmielewski, M. (2022). *The construct validity of dissociation measures* [Manuscript in preparation].

Chmielewski, M., Ruggero, C. J., Kotov, R., Liu, K., & Krueger, R. F. (2017). Comparing the dependability and associations with functioning of the DSM-5 Section III trait model of personality pathology and the DSM-5 Section II personality disorder model. *Personality Disorders*, *8*(3), 228–236. https://doi.org/10.1037/per0000213

Chmielewski, M., Sala, M., Tang, R., & Baldwin, A. (2016). Examining the construct validity of affective judgments of physical activity measures. *Psychological Assessment*, *28*(9), 1128–1141. https://doi.org/10.1037/pas0000322

Chmielewski, M., & Trujillo, M. (2020). The need for a more rigorous approach to diagnostic reliability: Commentary on categorical assessment of personality disorders. In C. W. Lejuez & K. L. Gratz (Eds.), *The Cambridge handbook of personality disorders* (pp. 369–371). Cambridge University Press. https://doi.org/10.1017/9781108333931.064

Chmielewski, M., & Watson, D. (2008). The heterogeneous structure of schizotypal personality disorder: Item-level factors of the Schizotypal Personality Questionnaire and their associations with obsessive-compulsive disorder symptoms,

dissociative tendencies, and normal personality. *Journal of Abnormal Psychology*, *117*(2), 364–376. https://doi.org/10.1037/0021-843X.117.2.364

Chmielewski, M., & Watson, D. (2009). What is being assessed and why it matters: The impact of transient error on trait research. *Journal of Personality and Social Psychology*, *97*(1), 186–202. https://doi.org/10.1037/a0015618

Cicero, D. C., Kerns, J. G., & McCarthy, D. M. (2010). The Aberrant Salience Inventory: A new measure of psychosis proneness. *Psychological Assessment*, *22*(3), 688–701. https://doi.org/10.1037/a0019913

Cicero, D. C., Neis, A. M., Klaunig, M. J., & Trask, C. L. (2017). The Inventory of Psychotic-Like Anomalous Self-Experiences (IPASE): Development and validation. *Psychological Assessment*, *29*(1), 13–25. https://doi.org/10.1037/pas0000304

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, *31*(12), 1412–1427. https://doi.org/10.1037/pas0000626

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

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

Dalenberg, C. J., Brand, B. L., Gleaves, D. H., Dorahy, M. J., Loewenstein, R. J., Cardeña, E., Frewen, P. A., Carlson, E. B., & Spiegel, D. (2012). Evaluation of the evidence for the trauma and fantasy models of dissociation. *Psychological Bulletin*, *138*(3), 550–588. https://doi.org/10.1037/a0027447

Dubester, K. A., & Braun, B. G. (1995). Psychometric properties of the Dissociative Experiences Scale. *Journal of Nervous and Mental Disease*, *183*(4), 231–235. https://doi.org/10.1097/00005053-199504000-00008

Dunn, G. E., Paolo, A. M., Ryan, J. J., & Van Fleet, J. (1993). Dissociative symptoms in a substance abuse population. *The American Journal of Psychiatry*, *150*(7), 1043–1047. https://doi.org/10.1176/ajp.150.7.1043

Dunn, G. E., Ryan, J. J., & Paolo, A. M. (1994). A principal components analysis of the Dissociative Experiences Scale in a substance abuse population. *Journal of Clinical Psychology*, *50*(6), 936–940. https://doi.org/10.1002/1097-4679(199411)50:6<936::AID-JCLP2270500619>3.0.CO;2-4

Ellickson-Larew, S., Stasik-O'Brien, S. M., Stanton, K., & Watson, D. (2020). Dissociation as a multidimensional transdiagnostic symptom. *Psychology of Consciousness: Theory, Research, and Practice*, *7*(2), 126–150. https://doi.org/10.1037/cns0000218

Frischholz, E. J., Braun, B. G., Sachs, R. G., & Hopkins, L. (1990). The Dissociative Experiences Scale: Further replication and validation. *Dissociation*, *3*(3), 151–153.

Frueh, B. C., Johnson, D. E., Smith, D. W., & Williams, M. A. (1996). A potential problem with the response format of the Dissociative Experiences Scale: A significant correlation with intelligence among combat veterans with PTSD. *Journal of Traumatic Stress*, *9*(3), 651–656. https://doi.org/10.1002/jts.2490090321

Fusar-Poli, P., Raballo, A., & Parnas, J. (2017). What is an attenuated psychotic symptom? On the importance of the context. *Schizophrenia Bulletin*, *43*(4), 687–692. https://doi.org/10.1093/schbul/sbw182

Gershuny, B. S., & Thayer, J. F. (1999). Relations among psychological trauma, dissociative phenomena, and trauma-related distress: A review and integration. *Clinical Psychology Review*, *19*(5), 631–657. https://doi.org/10.1016/S0272-7358(98)00103-2

Gnambs, T. (2014). A meta-analysis of dependability coefficients (test–retest reliabilities) for measures of the big five. *Journal of Research in Personality*, *52*, 20–28. https://doi.org/10.1016/j.jrp.2014.06.003

Gnambs, T. (2015). Facets of measurement error for scores of the big five: Three reliability generalizations. *Personality and Individual Differences*, *84*, 84–89. https://doi.org/10.1016/j.paid.2014.08.019

Goldberg, L. R. (1999). The Curious Experiences Survey, a revised version of the Dissociative Experiences Scale: Factor structure, reliability, and relations to demographic and personality variables. *Psychological Assessment*, *11*(2), 134–145. https://doi.org/10.1037/1040-3590.11.2.134

Gootzeit, J., Markon, K., & Watson, D. (2015). Measuring dimensions of posttraumatic stress disorder: The Iowa traumatic response inventory. *Assessment*, *22*(2), 152–166. https://doi.org/10.1177/1073191114540747

Green, S. B. (2003). A coefficient alpha for test–retest data. *Psychological Methods*, *8*(1), 88–101. https://doi.org/10.1037/1082-989X.8.1.88

Groth-Marnat, G., & Jeffs, M. (2002). Personality factors from the five-factor model of personality that predict dissociative tendencies in a clinical population. *Personality and Individual Differences*, *32*(6), 969–976. https://doi.org/10.1016/S0191-8869(01)00101-5

Harrison, J. A., & Watson, D. (1992). *The Dissociative Processes Scale* [Unpublished manuscript]. Department of Psychology, University of Iowa, Iowa City.

Heggestad, E. D., George, E., & Reeve, C. L. (2006). Transient error in personality scores: Considering honest and faked responses. *Personality and Individual Differences*, *40*(6), 1201–1211. https://doi.org/10.1016/j.paid.2005.10.014

Hilbert, S., Küchenhoff, H., Sarubin, N., Nakagawa, T. T., & Bühner, M. (2016). The influence of the

response format in a personality questionnaire: An analysis of a dichotomous, a likert-type, and a visual analogue scale. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, *23*(1), 3–24. https://doi.org/10.4473/TPM23.1.1

Hunter, E. C. M., Sierra, M., & David, A. S. (2004). The epidemiology of depersonalisation and derealisation. A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, *39*(1), 9–18. https://doi.org/10.1007/s00127-004-0701-4

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality* (2nd ed., pp. 102–138). Guilford.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage Publications.

Kuhlmann, T., Dantlgraber, M., & Reips, U. D. (2017). Investigating measurement equivalence of visual analogue scales and Likert-type scales in Internet-based personality questionnaires. *Behavior Research Methods*, *49*(6), 2173–2181. https://doi.org/10.3758/s13428-016-0850-x

Kwapil, T. R., Wrobel, M. J., & Pope, C. A. (2002). The five-factor personality structure of dissociative experiences. *Personality and Individual Differences*, *32*(3), 431–443. https://doi.org/10.1016/S0191-8869(01)00035-6

Lilienfeld, S. O., Lynn, S. J., Kirsch, I., Chaves, J. F., Sarbin, T. R., Ganaway, G. K., & Powell, R. A. (1999). Dissociative identity disorder and the socio-cognitive model: Recalling the lessons of the past. *Psychological Bulletin*, *125*(5), 507–523. https://doi.org/10.1037/0033-2909.125.5.507

Lynch, S. M., Forman, E., Mendelsohn, M., & Herman, J. (2008). Attending to dissociation: Assessing change in dissociation and predicting treatment outcome. *Journal of Trauma & Dissociation*, *9*(3), 301–319. https://doi.org/10.1080/15299730802139063

Lynn, S. J., Maxwell, R., Merckelbach, H., Lilienfeld, S. O., Kloet, D. V. H., & Miskovic, V. (2019). Dissociation and its disorders: Competing models, future directions, and a way forward. *Clinical Psychology Review*, *73*, Article 101755. https://doi.org/10.1016/j.cpr.2019.101755

Lyssenko, L., Schmahl, C., Bockhacker, L., Vonderlin, R., Bohus, M., & Kleindienst, N. (2018). Dissociation in psychiatric disorders: A meta-analysis of studies using the Dissociative Experiences Scale. *The American Journal of Psychiatry*, *175*(1), 37–46. https://doi.org/10.1176/appi.ajp.2017.17010025

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*(1), 28–50. https://doi.org/10.1177/1088868310366253

Paolo, A. M., Ryan, J. J., Dunn, G. E., & Van Fleet, J. (1993). Reading level of the Dissociative Experiences Scale. *Journal of Clinical Psychology*, *49*(2), 209–211. https://doi.org/10.1002/1097-4679(199303)49:2<209::AID-JCLP2270490212>3.0.CO;2-E

Pargent, F., Hilbert, S., Eichhorn, K., & Bühner, M. (2019). Can't make it better nor worse: An empirical study about the effectiveness of general rules of item construction on psychometric properties. *European Journal of Psychological Assessment*, *35*(6), 891–899. https://doi.org/10.1027/1015-5759/a000471

Raykov, T., & Penev, S. (2005). Estimation of reliability for multiple-component measuring instruments in test–retest designs. *British Journal of Mathematical & Statistical Psychology*, *58*(Pt 2), 285–299. https://doi.org/10.1348/000711005X38735

Reeve, C. L., Heggestad, E. D., & George, E. (2005). Estimation of transient error in cognitive ability scales. *International Journal of Selection and Assessment*, *13*(4), 316–320. https://doi.org/10.1111/j.1468-2389.2005.00328.x

Russell, C. J., & Bobko, P. (1992). Moderated regression analysis and Likert scales: Too coarse for comfort. *Journal of Applied Psychology*, *77*(3), 336–342. https://doi.org/10.1037/0021-9010.77.3.336

Sanders, B. (1992). The Dissociative Experiences Scale [Reply]. *The American Journal of Psychiatry*, *149*(1), 144. https://doi.org/10.1176/ajp.149.1.144

Schmidt, F. L., Le, H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods*, *8*(2), 206–224. https://doi.org/10.1037/1082-989X.8.2.206

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, *19*(1), 80–110. https://doi.org/10.1177/1094428115598239

Simms, L. J., Zelazny, K., Williams, T. F., & Bernstein, L. (2019). Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychological Assessment*, *31*(4), 557–566. https://doi.org/10.1037/pas0000648

Spitzer, C., Klauer, T., Grabe, H., Lucht, M., Stieglitz, R. D., Schneider, W., & Freyberger, H. J. (2003). Gender differences in dissociation a dimensional approach. *Psychopathology*, *36*(2), 65–70. http://proxy.libraries.smu.edu/login?url=https://www.proquest.com/scholarly-journals/gender-differences-dissociation-dimensional/docview/230504075/se-2?accountid=6667

Suzuki, T., Griffin, S. A., & Samuel, D. B. (2017). Capturing the DSM-5 alternative personality disorder model traits in the five-factor model's

nomological net. *Journal of Personality*, *85*(2), 220–231. https://doi.org/10.1111/jopy.12235

van Heugten-van der Kloet, D., Huntjens, R., Giesbrecht, T., & Merckelbach, H. (2014). Self-reported sleep disturbances in patients with dissociative identity disorder and post-traumatic stress disorder and how they relate to cognitive failures and fantasy proneness. *Frontiers in Psychiatry*, *5*, Article 19. https://doi.org/10.3389/fpsyt.2014.00019

Vispoel, W. P., Morris, C. A., & Kilinc, M. (2018). Using generalizability theory to disattenuate correlation coefficients for multiple sources of measurement error. *Multivariate Behavioral Research*, *53*(4), 481–501. https://doi.org/10.1080/00273171.2018.1457938

Vispoel, W. P., & Tao, S. (2013). A generalizability analysis of score consistency for the Balanced Inventory of Desirable Responding. *Psychological Assessment*, *25*(1), 94–104. https://doi.org/10.1037/a0029061

Waller, N., Putnam, F. W., & Carlson, E. B. (1996). Types of dissociation and dissociative types: A taxometric analysis of dissociative experiences. *Psychological Methods*, *1*(3), 300–321. https://doi.org/10.1037/1082-989X.1.3.300

Watson, D. (2001). Dissociations of the night: Individual differences in sleep-related experiences and their relation to dissociation and schizotypy. *Journal of Abnormal Psychology*, *110*(4), 526–535. https://doi.org/10.1037/0021-843X.110.4.526

Watson, D. (2003). Investigating the construct validity of the dissociative taxon: Stability analyses of normal and pathological dissociation. *Journal of Abnormal Psychology*, *112*(2), 298–305. https://doi.org/10.1037/0021-843X.112.2.298

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality*, *38*(4), 319–350. https://doi.org/10.1016/j.jrp.2004.03.001

Watson, D., Ellickson-Larew, S., Stanton, K., Levin-Aspenson, H. F., Khoo, S., Stasik-O'Brien, S. M., & Clark, L. A. (2019). Aspects of extraversion and their associations with psychopathology. *Journal of Abnormal Psychology*, *128*(8), 777–794. https://doi.org/10.1037/abn0000459

Watson, D., Stasik, S. M., Chmielewski, M., & Naragon-Gainey, K. (2015). Development and validation of the Temperament and Affectivity Inventory (TAI). *Assessment*, *22*(5), 540–560. https://doi.org/10.1177/1073191114557943

Watson, D., Stasik, S. M., Ellickson-Larew, S., & Stanton, K. (2015a). Explicating the psychopathological correlates of anomalous sleep experiences. *Psychology of Consciousness: Theory, Research, and Practice*, *2*(1), 57–78. https://doi.org/10.1037/cns0000038

Watson, D., Stasik, S. M., Ellickson-Larew, S., & Stanton, K. (2015b). Extraversion and psychopathology: A facet-level analysis. *Journal of Abnormal Psychology*, *124*(2), 432–446. https://doi.org/10.1037/abn0000051

Watson, D., & Wu, K. D. (2005). Development and validation of the schedule of compulsions, obsessions, and pathological impulses (SCOPI). *Assessment*, *12*(1), 50–65. https://doi.org/10.1177/1073191104271483

Weiss, B. D., & Coyne, C. (1997). Communicating with patients who cannot read. *New England Journal of Medicine*, *337*(4), 272–274. https://doi.org/10.1056/NEJM199707243370411

Wright, D. B., & Loftus, E. F. (1999). Measuring dissociation: Comparison of alternative forms of the dissociative experiences scale. *The American Journal of Psychology*, *112*(4), 497–519. https://doi.org/10.2307/1423648

Zingrone, N. L., & Alvarado, C. S. (2001). The Dissociative Experiences Scale-II: Descriptive statistics, factor analysis, and frequency of experiences. *Imagination, Cognition and Personality*, *21*(2), 145–157. https://doi.org/10.2190/K48D-XAW3-B2KC-UBB7