



# Development of a Validity Scale for the Dissociative Experience Scale-Revised: Atypicality, Structure, and Inconsistency

Ana Abu-Rus<sup>1</sup> · Kenneth J. Thompson<sup>1</sup> · Brandi L. Naish<sup>1</sup> · Cassie Brown<sup>1</sup> · Constance Dalenberg<sup>1,2</sup>

Received: 1 September 2019 / Accepted: 18 December 2019 / Published online: 13 January 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

The Dissociative Experiences Scale (DES), the most commonly used scale for assessing dissociation across settings, lacks a validity scale. In this study, six methods of enhancing validity were utilized: vocabulary and duration screening, manipulation checks, inconsistency, atypicality, and structure (unlikely pattern of responses). Six reverse-worded DES questions were developed to assess inconsistency, six questions regarding extremely rare or unknown symptoms assessed atypicality, and the difference between taxon and absorption items assessed structure. Honest, feigning, and posttraumatic stress disorder (PTSD) groups completed the assessment ( $N = 345$ ) via Amazon Mechanical Turk (MTurk) or SurveyMonkey. All groups received a brief definition of dissociation. The honest/PTSD groups were asked to complete the survey honestly. The feigning group members were asked to pretend to be someone with dissociative symptoms. Failure of the vocabulary, duration, or manipulation check validations led to 72 exclusions. The three groups differed significantly on the inconsistent items, the atypical items, and structure items,  $F(2, 271) > 7.52, p < .001$ , with the feigning group consistently performing worse than the two honest groups. The DES with validity scale may be most useful for community survey studies in which there is a high risk of malingering or feigning.

**Keywords** Dissociation · Trauma · Malingering · Forensic · DES · PTSD

The use of the Dissociative Experiences Scale (DES) in psychological research is widespread, as is suggested by over 2200 studies utilizing the scale listed in PsycINFO. Since its inception, the DES (Bernstein & Putnam, 1986) and its variants (e.g., DES-II, Dissociative Experiences Scale-Revised (DES-R), DES-Comparison) have found increasingly wide use in the assessment of psychopathology; this usage reflects the recognition of the ubiquity and complexity of dissociation in psychopathology (Lyssenko et al., 2018) and the increasing evidence for dissociation as a trauma consequence (Carlson, Dalenberg, & McDade-Montez, 2012). In addition to the disorders in which dissociation is a defining feature (e.g., dissociative amnesia as well as both dissociative and depersonalization disorders), severe dissociation is also a criterion in other disorders as well, such as posttraumatic stress disorder (PTSD) and borderline personality disorder (American

Psychiatric Association, 2013). Indeed, recognition of the role of dissociation in trauma has led to advances and refinements in the understanding of several disorders, such as the recent identification of a dissociative subtype of PTSD. A number of studies have not only supported the existence of a dissociative subtype of PTSD in veterans but have also associated it with greater symptom severity (Armour, Karstoft, & Richardson, 2014; Haagen, van Rijn, Knisheer, van der Aa, & Kleber, 2018; Tsai, Armour, Southwick, & Pietrzak, 2015; Waelde, Silvern, & Fairbank, 2005; Wolf, Lunney, et al., 2012; Wolf, Miller, et al., 2012).

Despite general acceptance in diagnostic manuals, dissociation remains a subject of debate. Researchers have cast doubt on a causal link between trauma and dissociation (Patihis & Lynn, 2017) and the role of trauma in dissociative amnesia (Pope, Poliakoff, Parker, Boynes, & Hudson, 2007) as well as dissociative identity disorder (DID; Piper & Merskey, 2004). Merckelbach and Patihis (2018) challenge the very use of the term “trauma-related dissociation” (TRD) as a potentially prejudicial term, favoring claimants who assert a causal link between past trauma and current dissociation and implying that dissociation has only one prominent cause (Merckelbach & Patihis, 2018). However, the position of these authors has been countered by Brand et al. (2018),

✉ Constance Dalenberg  
cdalenberg@alliant.edu

<sup>1</sup> Trauma Research Institute, Alliant International University, San Diego, CA, USA

<sup>2</sup> Department of Ph.D. Clinical Psychology, Alliant International University, 10455 Pomerado Road, San Diego, CA 92121, USA

who provide a robust defense of the trauma-related dissociation concept. In the context of a forensic disagreement between experts over the validity of the dissociative symptoms; however, the DES can be criticized both for its distributional qualities (skewness) and for its lack of a validity scale.

The DES has been translated into numerous languages, including Spanish, Hebrew, Italian, Dutch, Japanese, Turkish, Russian, Portuguese, German, Czech and French. A version of the DES has also been created for use with adolescent populations and, like the DES, shows strong cross-cultural validity and reliability (Soukup, Papežová, Kuběna, & Mikolajová, 2010). However, the DES and the DES-II, the original test and first revision, have been criticized for poor distributional qualities and a confusing format. The DES asks individuals what percentage of time they experience various symptoms, a difficult question to answer on subjective items such as feelings of unreality. The DES-R was a revision of the DES in which the format was changed to a frequency scale. Respondents are asked how often they experience various symptoms, with choices ranging from “never” to “more than once a week.” The changes in the DES normalized the distribution while preserving the correlation to the original scale, with  $r$  values from .80 to .90 (Coe, Dalenberg, Aransky, & Reto, 1995; Arzoumanian et al., 2017). The DES-R has been successfully used to predict relevant clinical constructs in a number of published papers (Coe et al., 1995; Kluemper & Dalenberg, 2014).

Despite improvements to the distribution of the DES in some of the recent revisions, the lack of a validity scale within the measure remains a significant drawback, particularly for a scale that may often find its way into the forensic arena (given the relationship of dissociation to trauma). At the same time, current assessment instruments that are widely used in forensic settings have not been developed to consider complex trauma profiles and, therefore, have misclassified true dissociative patients as feigners in experimental trials (Brand, Webermann, & Frankel, 2016; Palermo & Brand, 2018). The relatively recent rise of paid online survey takers (e.g., MTurk) also has provided a ready source of participants for studies, with online samples increasingly used for tests of theoretical models. Consequently, the general research field and the forensics field would greatly benefit from a scale that could identify potential instances of malingered dissociation while remaining sensitive to individuals with true elevations in dissociative symptoms.

It appears that most experts favor administration of a feigning assessment as part of standard practice and particularly forensic practice (Melton, Petrila, Poythress, & Slobogin, 2007). It has been noted that many psychological measures of personality (e.g., the Beck Depression Inventory-II; Beck, Steer, & Brown, 1996) do not consider the issue of feigning. However, recent scale development has been more consistent in incorporating a feigning screen into new measures. For

example, the Detailed Assessment of Posttraumatic Stress (DAPS), Trauma Symptom Inventory (TSI), and Minnesota Multiphasic Personality Inventory-2 (MMPI-2) all include scales which identify possible malingering, inconsistency, or feigning, thereby adding a means to assess the validity of each measure (Resnick, West, & Wooley, 2018; Gray, Elhai, & Briere, 2010; Butcher, Graham, Ben-Porath, Tellegan, & Dahlstrom, 2001). In evaluating validity, researchers tend to employ a few specific methods which have been shown to be effective in identifying feigning within an assessment. The following methods have been used in published measures to evaluate validity:

**Inconsistency** This method involves addressing a specific thought, feeling, or construct multiple times throughout an assessment in order to determine consistency in an individual's responses. As an example, an evaluation of inconsistency has been incorporated into the newer versions of the MCMI. Both the MCMI-III and MCMI-IV include an Inconsistency Scale which compares pairs of items to identify if a person appears to be responding randomly (Millon, Millon, Davis, & Grossman, 1994; Millon, Grossman, & Millon, 2015). The Variable Response Inconsistency Scale and the True Response Inconsistency Scale scores on the MMPI-2 are also inconsistency measures.

**Atypicality** To assess atypicality, measures include several items that probe for thoughts, feelings, or behaviors that are extremely rare or unknown. Multiple endorsements of such items typically indicate feigning. There are several well-known measures that include such validity items, including the Atypical Response Scale of the TSI (Briere, Elliott, Harris, & Cotman, 1995; Gray et al., 2010). The Validity scale of the MCMI-IV also includes such improbable items (Millon et al., 1994; Millon et al., 2015).

**Unlikely Extremity** Unlikely extremity is related to atypicality; however, it differs slightly in that it aims to identify those individuals who have a pattern of extremity. Examples of this type of measure are the Under-response and Hyper-response scales of the Trauma Symptom Checklist for Children (Briere, 1996). In this measure, counts are made of complete denial of common thoughts or behaviors (under-response) and responses at the ceiling for less common thoughts or behaviors (hyper-response).

**Structure** This validity check involves *patterns* of responses that do not typically occur, such as unusual, co-occurring symptoms or endorsement of more serious aspects of a disorder without endorsing the less serious aspects. The Infrequency-Psychopathology Scale (Fp) of the MMPI-2 partially addresses this form of feigning. The scale attempts to identify those individuals who are either over-reporting

symptoms or those who seem to be responding randomly, resulting in *unlikely patterns of symptoms* (Butcher et al., 2001). That is, here the issue is not that the individual was extreme but that he or she endorsed items in a pattern that was unlikely. For the DES, the structure score consisted of subtracting the taxon score (items with a low base rate) from the absorption score (items with a high base rate), with low or negative scores indicating invalidity. Base rates for the absorption and taxon items have been repeatedly established in prior studies (Olsen, Clapp, Parra, & Beck, 2013; Waller, Putnam, & Carlson, 1996).

**Language Proficiency** This validity check involves administering an examination of grade-level language proficiency using the same level of vocabulary that is used in the remainder of the assessment or in other administered assessments. Such checks are quite rare outside of the neuropsychological realm, but lack of proficiency in language clearly increases the likelihood of confusion for the participants as to the meaning of assessment questions.

**Duration** This validity check examines whether the speed of completion for a specific assessment is too slow or too fast when compared with a mean completion time for a group. In online forums, completion in a few minutes, significantly faster than the likely reading capacity of the respondents, can mean either that the test was taken by an automated system or that the individual did not read the questions.

**Manipulation Check** Experimental studies often include several questions that ensure that the respondent is aware of the directions respective to his/her group assignment (Foschi, 2014). Failure on this check could have a number of meanings, including a lack of clarity on the part of the experimenter but clearly compromises the interpretation of results.

Another type of data integrity protection that has specific applicability to some online surveys involves checking the geographic location of the participant. As more and more research relies on Internet-based solicitation, an important aspect of validity may concern the country of origin of participants. Survey sites such as MTurk pay participants according to the number of surveys that they complete. Consequently, individuals may be incentivized to participate in surveys for which they are not necessarily qualified (e.g., ignoring such qualifiers as U.S. residency or native English speaker). While imperfect, checking the Internet Protocol (IP) addresses of participants provides at least a first tool in culling potentially inappropriate participants from a data set. Although definitional questions from Language Proficiency (LP) tests will catch many of these individuals, those who value the incentive can easily use efficient internet resources to find the definitions of words and pass the LP requirement.

In light of concerns over ensuring valid administration of the DES-R, we developed a form of the assessment that incorporated an embedded validity scale comprising atypicality items, inconsistency items, and structure. The addition includes ten new questions, six for atypicality and six for inconsistency items. With regard to structure, we subtracted taxon from absorption item scores, with a cutoff set through receiver operator curve (ROC) analysis. Additionally, all participants were required to meet specific cutoffs for English proficiency as well as duration of test taking were required to pass the manipulation check and were checked for duplicate IP addresses or IP addresses from non-English speaking countries. After exclusion based on 8th-grade vocabulary scores under 60%, unlikely duration, and suspect IP addresses, structure, inconsistency, and atypicality were hypothesized to differentiate honest responders from those asked to feign dissociative symptoms. Exclusion based on IP, vocabulary, duration, and the three validity scales was expected to increase the correlation between reported trauma and dissociation, as would be predicted from theories of trauma-related dissociation, rather than decrease, as would be expected from fantasy-based theories of dissociation (see Dalenberg et al., 2012).

## Methods

### Participants

Participants were recruited via MTurk to answer a 7–15-min survey in exchange for \$0.75 to \$1.00. MTurk is a Web-based platform designed to recruit and pay participants to perform various tasks. The quality of data collected via MTurk has been shown to meet or exceed the psychometric standards associated with published research (Buhrmester, Kwang, & Gosling, 2011). Additionally, samples collected from MTurk are more representative of the US population than in-person convenience samples (Berinsky, Huber, & Lenz, 2012; Buhrmester et al., 2011). To maximize the probability that the participants were attentive and motivated during the task, only participants who had a “Master Worker” designation from MTurk were allowed to participate for the first 25% of data collection (across honest controls and feigning groups); these individuals have demonstrated a high degree of success in performing a wide range of human intelligence tasks across a large number of requesters. The second quarter of the participants was collected without Master Worker designation, and all variables were compared. When no differences were found, the remaining participants were collected without Master Worker designation.

PTSD ( $n = 40$ ) participants and Dissociative Disorder (DD) participants ( $n = 5$ ) were consecutive patients with the diagnosis of PTSD requesting therapy at a local trauma-centered clinical practice. The PTSD diagnosis was confirmed through

administration of the Clinician-Administered PTSD Scale (CAPS-5). Sixteen had a diagnosis of the dissociative subtype of PTSD. The five DD clients were not utilized in the analyses, given unacceptable power for testing as a separate group. Results will be presented for pilot purposes.

## Measures

Respondents completed a brief demographic scale, the DSM-V (see below), a brief depression inventory, and a vocabulary screen.

**Demographic Questionnaire** Participants completed a brief demographic questionnaire regarding age, gender, and race/ethnicity. Options for prior trauma were taken from the DAPS (Briere, 2001) and could range from 0 to 11.

**Patient Health Questionnaire 9** The PHQ was designed to measure depression severity in medical populations in clinical settings. The categories within the Patient Health Questionnaire 9 (PHQ-9) are derived from the DSM-IV classification system pertaining to: (1) anhedonia, (2) depressed mood, (3) trouble sleeping, (4) feeling tired, (5) change in appetite, (6) guilt or worthlessness, (7) trouble concentrating, (8) feeling slowed down or restless, and (9) suicidal thoughts. Participant response options vary from “not at all” to “nearly every day.” The PHQ-9 has been shown several times to have good reliability and validity (Kroenke, Spitzer, & Williams, 2001; Lowe, Unutzer, Callahan, Perkins, & Kroenke, 2004; Martin, Rief, Klaiberg, & Braehler, 2006; Pinto-Meza, Serrano-Blanco, Peñarrubia, Blanco, & Haro, 2005).

**English Vocabulary Screener** A vocabulary test was administered consisting of seven 8th-grade-reading level vocabulary words, as identified by the Spache Readability Formula. Respondents who received scores under 60% on the vocabulary test were excluded.

**Dissociative Experiences Scale-Revised** The Dissociative Experiences Scale (DES; Bernstein & Putnam, 1986) is a 28-item self-report measure that assesses the frequency of dissociative experiences using three major categories: (1) absorption/imaginative involvement, (2) amnesia, and (3) depersonalization/derealization. Responses range from “this happens never” to “this happens at least once a week.” The DES has high reliability ( $r = .83, p < .0001$ ) and high internal consistency ( $\alpha = .95$ ; Frischholz, Braun, Sachs, Hopkins, et al., 1990). The test also was found to differentiate well between dissociative and non-dissociative clinical groups (Dubester & Braun, 1995).

Due to repeated findings of skewness and leptokurtosis in the DES, Dalenberg et al. (1994) revised the response format of the scale to a frequency scale. This change normalized the

distribution without changing the relationship of the DES-R to other important variables (Coe et al., 1995). The relationship between the DES and DES-R reported in Coe et al. was .90.

The structure scale of the DES-V (the label given for the DES with the embedded validity scale) was based on subtraction of the taxon items from the absorption items. Absorption items were DES items 1, 2, 14, 15, 16, 17, 18, 20, 21, 22, and 23, and taxon items were DES items 3, 5, 7, 8, 12, 13, 22, and 27 (Waller et al., 1996). The Atypical Response scale was determined by the sum of six atypical items developed by the Trauma Research Institute and verified as atypical by three highly published authors on dissociation and a pilot group of 30 patients with dissociative PTSD and 5 DD clients (none of whom were included in the main study). Items were included if they were seen as atypical by 90% or more of the patients and all experts (see Table 1 for a list of DES-R items with corresponding inconsistency items and a list of atypical items.) Chosen items had a mean of 1 or less on the 0–6 Likert-scaled atypicality items. Examples of atypical items include “sometimes people find they do not feel physical pain at all” and “sometimes people find that they collect things that remind them of their trauma but they do not remember buying those things.” Finally, six of the DES-R items were reworded in a positive direction and an inconsistency score was generated. Participants had to answer with a frequency score of five or more to an item as well as on the matching reversed-inconsistency question to be considered inconsistent. Disagreements with both items were not considered inconsistencies, as it is possible to say that one is neither extremely hypervigilant as a driver nor unaware of surroundings. A score from 0 to 6 was calculated for the number of paired disagreements.

**Other Exclusion Criteria** IP addresses that indicated the participant was from a non-English speaking country were excluded. Further, we established a cutoff for time to completion that may indicate insufficient care in answering the questionnaire. The cutoff for duration of test taking was derived by asking a group of 14 Ph.D. students to complete our survey as quickly as possible while recording their completion times. All participants in the duration pilot test were members of the Trauma Research Institute and were therefore familiar with all instruments. Thus, the duration was likely to be faster than is typical. To measure the validity of responses, participants who completed the survey two standard deviation faster than our Ph.D. group were considered to have put forth less effort than required for this assessment and excluded. Lastly, as a manipulation check, all participants were required to correctly identify their group assignment. Incorrect responses on the manipulation check, unlikely duration, and IP addresses outside of English-speaking countries resulted in the participant being dropped from the study.



**Table 1** DES-R items with corresponding inconsistency items

DES-R item	Inconsistency item
Some people have the experience of driving a car and suddenly realizing that they do not remember what has happened during all or part of the trip.	<i>I consider myself a very hyper-aware driver. I always notice changes in the scenery as I drive familiar routes.</i>
Some people have the experience of not being sure if the things that they remember happening really did happen or if they just dreamed them.	<i>My dreams are not very vivid.</i>
Some people find that they are sometimes able to ignore pain.	<i>I think I am more sensitive to pain than other people.</i>
Some people find that when they are watching television or a movie they become so absorbed in the story that they are unaware of other events happening around them.	<i>Even if I am watching television, it is very easy to get my attention.</i>
Some people find that they have no memory for some important events in their lives, for example, a wedding or graduation.	<i>My memory for important emotional events in my life is very good compared to others.</i>
Some people have the experience of finding new things among their belongings that they do not remember buying.	<i>I have an excellent memory for my consumer behavior. I cannot imagine forgetting that I had bought something and buying it again.</i>
<b>Atypical items</b>	
Some people find that when they think about the trauma they faint or have a headache.	
Some people sometimes find themselves looking in the mirror and seeing a potential attacker when no one is there.	
Some people sometimes find they can only talk about the experience if they talk in a 3rd person (i.e., he was sexually abuse).	
Sometimes people find that they collect things that remind them of their trauma but they do not remember buying those things.	
Sometimes people find they do not feel physical pain at all.	
Sometimes people find they lose control of the physical parts of their body or make gestures or do things they do not mean to do.	

## Procedures

Participants were divided into three groups. All groups were given a brief definition of dissociation. Participants could stop at any time with no penalty; therefore, completion of the survey was voluntary after receiving instructions of group assignment. Dropouts are discussed in “Results.”

**Honest Control Group** Participants were asked to answer the survey as honestly as possible. The honest group was also told that “giving dishonest answers on a medical survey is like giving contaminated blood in a blood donation. It can be extremely harmful for the scientific project.” In a separate pilot project, the addition of this statement was found to increase admission of alcohol misuse ( $t = 3.24, p < .01$ ), failure to use a condom with new sexual partners ( $t = 4.41, p < .01$ ) and cheating in undergraduate school ( $t = 5.94, p < .01$ ) in an undergraduate sample of 434 students.

**Feigning Group** Participants were asked to pretend to be someone who experiences dissociative symptoms and to try to convince the researchers using their responses to the questions. To ensure the participants were attentive and motivated during the task, the feigning group was provided an extra monetary incentive in the form of an additional \$1.00 to the top 50 most believable malingerers and an additional \$5.00 to the top 5 most believable. Dissociation was described as an experience that at times occurs after negative events, wherein a person feels detached or disconnected from reality or feels fragmented or disconnected internally. Other listed symptoms

were claims of lack of recall for various activities and difficulty in feeling normal sensations.

**Posttraumatic Stress Disorder Group** The 40 PTSD participants took the survey under the same conditions as the honest control group, as did the 5 DD participants (who, again, were not included in the analyses).

## Results

Initially, 357 individuals participated in the experiment. However, 12 individuals were participating through IP addresses originating in non-English-speaking countries and were excluded from the subsequent data analysis. Of the 345 remaining participants, 20 did not complete the DES-R, 34 failed the manipulation check (incorrectly identifying their Honest/Feigning group membership), and 35 failed the vocabulary test (including 58% ( $n = 7$ ) of those with IP addresses with non-English speaking countries). Only seven individuals fell into the category of suspiciously fast completers. As there was overlap among these failures, a total of 84 individuals were eliminated. Analysis of the excluded individuals versus retained participants yielded non-significant results for age, race, and gender. Despite instructions that asked those in the feigning group to answer honestly for these demographics, feigning group members were more likely to be excluded ( $\chi^2 = 10.491, p < .001; 33$  vs. 14%). It should be noted that, of those who were excluded because they failed the vocabulary check, 94% also failed the atypicality criteria and 90.3%

failed structure. For those who failed the manipulation check, 81% ( $n = 25$ ) also failed atypicality, 18% ( $n = 6$ ) failed inconsistency, and 84% ( $n = 27$ ) failed structure. Ninety-one percent of the respondents who failed the manipulation check also failed at least one of the validity checks. Of the remaining subjects to be evaluated on the DES-V, 40 were in the PTSD group, 98 were in the feigning, and 135 were in the honest control group, for a total  $N = 273$ .

Table 2 describes the age, gender, and race distribution of the retained sample. The honest control, feigning, and PTSD groups did not differ on gender, age, or race distribution.

### Group Differences on Atypicality, Structure, and Inconsistency

Distributions for the DES-R total score, atypicality, and structure were relatively normal within the two honest groups (PTSD, honest control). Inconsistency scores were skewed, with only 26 individuals inconsistent on one item and nine inconsistent on more than one item. The vast majority (89.8%) had no inconsistencies. The inconsistency variable was thus recoded as a dichotomy.

The three groups significantly differed on the Atypical items:  $F(2, 271) = 110.12, p < .001$ , with one missing value (see Table 3 for full results). The three groups also significantly differed on the inconsistency items ( $\chi^2 = 22.58, p < .001$ ) and the structure items,  $F(2, 272) = 27.28, p < .001$ . Effect sizes were larger for atypicality ( $\eta^2 = .50$ ), than for structure ( $\eta^2 = .17$ ) or inconsistency ( $\eta^2 = .08$ ).

Results of the Pearson correlation indicated that there was a significant positive association between inconsistency and atypicality,  $r(323) = .40, p < .01$ , between inconsistency and structure,  $r(324) = .117, p < .05$  and between atypicality and structure,  $r(324) = .484, p < .01$ . Among those who were not excluded, duration of test taking did not significantly correlate with any of the three constructs.

**Logistic Results** Using the three constructs and comparing the honest with the feigning participants on their continuous scores on atypicality and structure together with their

**Table 2** Participant characteristics

Race	N (%)	Gender	N (%)	Age
Caucasian	181 (66.5%)	Male	147 (54.4%)	Mean 37.88
African American	22 (8.1%)	Female	123 (45.6%)	SD 10.33
Hispanic	13 (4.7%)			
Asian	48 (17.6%)			
Other	8 (3.0%)			

Note: One participant failed to note race and three participants did indicate identification with either gender

**Table 3** Between group differences on malingering items

	PTSD N = 40	Feigning N = 95	Honest N = 135
Atypicality	1.38a (1.68)	14.89b (8.49)	2.53a (4.47)
Inconsistency	0.05a (.22)	0.29b (.66)	0.06a (.38)
Structure	11.55a (5.58)	3.61b (6.09)	6.96c (5.74)

Note. Means with different lowercase letters are significantly different

dichotomous score in inconsistency, a logistic regression was able to correctly classify 90.8% of the honest participants (honest controls and PTSD participants) and 75.3% of the Feigning participants. The logistic regression was statistically significant,  $\chi^2(3) = 157.421, p < .001$ , and a 60.3% (Nagelkerke  $R^2$ ) variance in responding was explained by the model.

**ROC Analyses** To establish cutoffs for feigning, we used ROCs for the structure and atypicality data. For the structure items, we followed a two-step process. As previously defined, structure represents the difference between taxon and absorption scores, with the expectation that honest respondents would endorse more high-base-rate absorption items than low-base-rate taxon items. However, as the reader will recall, feigning is associated with a low rather than a high structure score here. A high score on either absorption or taxon items can only occur if the individual is expressing dissociative symptoms; if dissociation scores as a whole were low, the issue of feigning is moot. Therefore, those participants who did not elevate on any DES-R items (scoring 3 or over) were not included in the ROC analysis.

In the second step, a ROC curve was used to establish a cutoff for honest respondents on structure. The ROC curve established the score above which an individual was likely to be feigning. This second curve established a cutoff of 4.5, with an AOC of .70,  $p < .001$ . Thus, those who scored 4 or less on structure and who elevated at least one item on the DES-R were given 1 point (for possible feigning) on structure. Using these criteria, 18.6% ( $n = 19$ ) of the honest controls, 73.3% ( $n = 66$ ) of the feigning group and 12.5% ( $n = 5$ ) of the PTSD group failed on the structure variable, as did 1 of the DD case controls.

The ROC curve for atypicality established a cutoff of 5.5, above which an individual was judged to be feigning. The AOC statistic was excellent (.906,  $p < .001$ ). One of the PTSD group failed the atypicality criteria, as did 18 of the honest controls (13.3%) and 79 of the feigning group (80.6%). One of the DD group also failed the atypicality criterion.

For inconsistency, there was insufficient variance to apply the ROC procedure. Chi-squares established that 3.7% ( $n = 5$ ) of the honest control group, 22% ( $n = 22$ ) of the feigning

group and 5% ( $n = 2$ ) of the PTSD group endorsed at least one inconsistency item. None of DD cases had a positive inconsistency score.

In order to calculate the sensitivity and specificity of our cutoffs, we assigned individuals one point each for scoring in the feigning range on atypicality, structure, and inconsistency. Scores for this scale, which we call ASI, could range from 0 to 3. The full results for the three groups are in Table 4, and related sensitivity, specificity, PPV, and NPV values are in Table 5. Note that elimination of subjects who fail at least one of the three tests retained 90% of Honest responders while eliminating 71.31% of those asked to feign dissociative symptoms.

### Comparison of Correlations Before and After Exclusion

The number of types of potential traumas was summed based on reports in the initial demographic questionnaire. Using the full sample, the correlation of trauma with the DES-R was .14,  $p < .01$ . Using the sample after exclusion of the 84 subjects based on vocabulary, IP address, incomplete reports, and unlikely duration, the correlation was .19,  $p < .01$ . Using on those who passed all three validity screens ( $n = 145$ ), the correlation was .29,  $p < .001$ . The comparable figures for the depression screen (PHQ) were .19,  $p < .01$  for the full sample, .29,  $p < .001$  for the sample after initial exclusions, and .32,  $p < .001$  after screening on the validity tests.

**Table 4** Relationship between feigning scale score and group membership

	Feigning scale (ASI) total				
	0.00	1.00	2.00	3.00	Total
<b>Group</b>					
<b>Honest</b>					
Count	100	20	14	1	135
% within group	74.1%	14.8%	10.4%	0.7%	100.0%
% within ASI	69.0%	40.0%	23.0%	5.9%	49.5%
<b>Feigning</b>					
Count	11	24	47	16	98
% within group	11.2%	24.5%	48.0%	16.3%	100.0%
% within ASI	7.6%	48.0%	77.0%	94.1%	35.9%
<b>PTSD</b>					
Count	34	6	0	0	40
% within group	85.0%	15.0%	0.0%	0.0%	100.0%
% within ASI	23.4%	12.0%	0.0%	0.0%	14.7%
<b>Total</b>					
Count	145	50	61	17	273
% within group	53.1%	18.3%	22.3%	6.2%	100.0%
% within ASI	100.0%	100.0%	100.0%	100.0%	100.0%

**Table 5** Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for values of ASI

	Sensitivity	Specificity	PPV	NPV
0	71.31	90.09	88.78	74.07
1	80.77	77.42	64.29	88.89
2	94.12	62.04	16.33	99.26

*Note.* 0 = retain as Honest responder only if sum of criteria is zero; 1 = retain as Honest responder only if sum of criteria is 1 or 0; 2 = retain as Honest responder only if sum of criteria is 2 or less

### Discussion

Although the DES is commonly used in research, forensic, and clinical settings to assess for dissociative symptomatology, it currently does not have a validity scale. This limitation of the DES can have significant effects on the status of the field in a number of ways—decreasing effect sizes in survey research, undermining diagnostic accuracy, and encouraging the small group of critics who argue that the dissociative disorders themselves (most notably dissociative amnesia and dissociative identity disorder) do not exist (McNally, 2007; Merckelbach & Patihis, 2018). Arguably, the fact that we were able to distinguish known feigners from honest responders with 84.5% accuracy (in the logistic regression) suggests that the feigners and allegedly true responders were in some way different. The use of three quite different methods, each of which individually differentiated the honest responders from those told to feign symptoms, also adds weight to the argument in favor of use of such methods. It is particularly important to note that substantial differences in findings emerge with and without taking the validity scales into account. Including all samples, the correlation between total numbers of possible traumas reported and DES-R total was .19,  $p < .01$ . Considering only those who passed validity checks, the same correlation was .29,  $p < .001$ . Substantial differences in findings with strict exclusion criteria in online samples are also noted by Thomas and Clifford (2017).

In this study, we have created a validity scale for the DES-R using several different types of checks, including structure, atypicality, inconsistency, manipulation checks, and vocabulary to assess the accuracy of responses. This inclusion of a validity scale was thought to make the DES-V a more effective tool in Web-based research, as it provides a simple means of checking the validity of data gathered online. The package of assessments addressed a variety of methods suggested for protecting the integrity of data.

The elimination of IP addresses through non-English speaking countries is an option available (typically with extra cost) with many survey-responder companies, including MTurk. In our work, only 3% ( $n = 12$ ) of the sample were eliminated for this reason. Although it is reasonable to argue

that those from non-English speaking countries may well have a vocabulary sufficient to take the test, it is certainly telling that 58% of this group also failed vocabulary and 83% failed atypicality, suggesting that they were not reliable responders. Almost all of those who failed the vocabulary test (97%,  $n = 30$ ) also failed one or more of the validity checks, strongly supporting the use of such measures in online research. Failure of the manipulation check was also a strong correlate of failure of validity checks. Here, 91% ( $n = 29$ ), failed at least one of the embedded validity measures. Both of these findings also are in keeping with Thomas and Clifford's (2017) advisory for strong exclusion criteria in online research.

The atypicality scale demonstrated the strongest ability to differentiate honest respondents from feigners compared with the other measures used in the study. Scores for participants from the feigning group were more than five times higher than those of honest responders and the participants with PTSD. While the atypical items show good initial utility in research, it is important that these items be tested with a group of individuals with confirmed dissociative disorders before being utilized in a clinical setting. It is promising, however, that if a cut point of 1 was used, rather than 0 (the recommended cut point for online research), none of those with DD and only one PTSD client was excluded in the present sample.

Few participants showed elevated levels of inconsistency in responding, which was meant to pick up carelessness in those who were less committed to honest responding. Participants in the feigning group were more easily identified through atypicality or structure than through inconsistency items. However, inconsistency was here defined by extreme differences (strongly agreeing with two similar items worded in opposite directions). This disallows inconsistency between answers that were in the mid-range of the scale, which was characteristic of the majority of the sample. In retrospect, it may have been more useful to include an attention check rather than a set of inconsistency items in order to judge careless responding. The former method involves inclusion of items with an obvious correct response, requires only one to two items, and has been shown to work well in other online research (Berinsky, Margolis, & Sances, 2014). A number of studies have shown comparable careless responding in online and in-person samples, but estimates of unacceptable carelessness tend to range from 5 to 10% (Johnson, 2005; Meade & Craig, 2012).

The results also indicate that the structure variable is useful in assessing the validity of responses. The feigning group showed a smaller difference between the taxon and absorption items than did the two honest groups among the subgroup who claimed any dissociative symptoms. Currently, assessment of taxon and absorption items is unique to the DES/DES-R, but the technique could easily be incorporated into other scales with reported base rates, such as the PHQ-9 (Rief, Nanke, Klaiberg, & Braehler, 2004), the Beck Anxiety

Inventory (Gillis, Haaga, & Ford, 1995), and the Beck Depression Inventory (Dawes et al., 2010). We emphasize that structure is recommended here as a tool at present only for the measurement and validation of dissociation as symptom, including dissociation in the context of PTSD or BPD. Within populations with DID, which should be exceedingly rare in an online random sample, structure may not be valid as an indicator, as taxon scores are often elevated. Using an archival sample of DES-R data from 32 DID individuals who had participated in other studies within our laboratory, 20% would have failed the structure criteria. This failure, however, is difficult to interpret, given that several of these individuals were questionable cases of DID (as reported by their therapists) and many did not have verification of their diagnoses through a reliable assessment tool.

The idea of using duration of test taking as a validity check seems to have face validity, and it was somewhat surprising that the cutoff identified few participants. This result may have been an artifact of study design, given that highly educated students, both familiar with the screen and trained in digesting material quickly, generated the target duration figures. This method was an effort to account for the quick reaction times in professional survey takers by substituting higher average education and training, but this arrangement may have been a poor comparison.

## Limitations and Conclusions

The DES-V results using the criteria of 1 or more elevated validity scales shows general promise. The sensitivity and specificity scores of .90 and .71 are comparable with average findings of many accepted validity screens. The Test of Memory Malingering Trial 1, for instance, has an average reported specificity of .90 and sensitivity between .59 and .70 (Martin et al., 2019). The validity screen for the Connors ADHD scale again has specificity of .86–.90 but sensitivity of .44 to .63 for random responding and .31–.46 for feigning (Walls, Wallace, Brothers, & Berry, 2017). The SIRS-2 is reported to have high specificity (.90) but moderate sensitivity (.54) (Tarescavage & Glassmire, 2016). Individually, both the structure and atypicality subscales achieved comparable or better statistics with these scales when used alone, with atypicality alone rivaling use of the full set of predictors (sensitivity = .80; specificity = .91). Further work will focus on a shift to the attention check methodology to capture carelessness, expansion of the atypicality items to ensure validity across types of samples, broadening of the structure criterion (using base rates of items on the full scale, rather than simple absorption and taxon items), and inclusion of a dissociative disorder sample. Use of IP investigation or exclusion and use of vocabulary screen are also clearly supported.



Another limitation to the study, ubiquitous in this area of research, is concern about the “honest” responders. It is quite possible that a number of the honest controls were in fact malingering, despite the attempt to use social influence to increase compliance and largely accurate manipulation check responses. If so, however, the screen is likely even more effective than is presented here, and that a few of the “honest” responders labeled as false positives (feigners) were indeed true feigners. It can be argued with certainty only that the honest groups were likely more honest on the average than those told to feign symptoms.

The most important limitation to the current work is the absence of a dissociative disorder control group, which is now in progress. The types of dissociation associated with simple PTSD (versus complex PTSD) are limited (Van der Hart, Nijenhuis, & Steele, 2005) and do not include the more severe fragmentation that is characteristic of Dissociative Identity Disorder. Although those with traumatic histories in general are often shown to elevate on validity scales (Flitter, Elhai, & Gold, 2003), concern about validity scales with dissociative disorder groups is more significant (Palermo & Brand, 2018). Given that the base rate of DID is low (Ross, 1991; Şar, Akyüz, & Doğan, 2007), we believe that the current version of the scale is valuable for further research in community populations, but clinical replications are mandatory before it would be deemed usable for forensic purposes. Such research efforts are challenging, in that many institutions do not routinely screen for dissociative disorders (Ginzburg, Somer, Tamarkin, & Kramer, 2010).

The incorporation of embedded validity checks into the administration of the DES-R provides a means for the objective evaluation of feigning, a critical requirement in forensic practice. In forensic evaluation, lack of a validity scale will leave experts relying on clinical intuition or experience in judging the reliability of a specific case, a process fraught with opportunities for miscarriages of justice (Dawes, Faust, & Meehl, 1989). In research and clinical fields, the proliferation of new assessment tools, compounded by the increasing reliance on isolated and unseen (i.e., Internet) subjects, often professional survey takers who benefit more monetarily from speed than from accuracy, necessitates more attention be paid to the validity of the data produced through such means. The DES-V may be a tool to move the field toward this goal.

**Funding** This research was supported in part by a grant from the Alliant Foundation.

## References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Arlington, VA: Author.
- Armour, C., Karstoft, K. I., & Richardson, J. D. (2014). The co-occurrence of PTSD and dissociation: Differentiating severe PTSD from dissociative-PTSD. *Social Psychiatry and Psychiatric Epidemiology*, 49, 1297–1306. <https://doi.org/10.1007/s00127-014-0819-y>.
- Arzoumanian, M., Verbeck, E., Estrellado, J., Dahlin, K., Hennrich, E., Stevens, J., & Dalenberg, C. (2017). Psychometrics of three dissociation scales: Reliability and validity data on the DES-R, DES-II, & DESC. Unpublished paper, Alliant International University, San Diego.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com’s mechanical Turk. *Political Analysis*, 20, 351–368. <https://doi.org/10.1093/pan/mpr057>.
- Berinsky, A., Margolis, M., & Sances, M. (2014). Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self Administered Surveys. *American Journal of Political Science*, 58, 739–753. <https://doi.org/10.1111/ajps.12081>
- Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease*, 174, 727–735. <https://doi.org/10.1097/00005053-198612000-00004>.
- Brand, B. L., Dalenberg, C. J., Frewen, P. A., Loewenstein, R. J., Scheilke, H., Brams, J. S., & Spiegel, D. (2018). Trauma-related dissociation is no fantasy: Addressing the errors of omission and commission in Merckelbach and Patihis. *Psychological Injury and Law*, 11, 377–393.
- Brand, B. L., Webermann, A. R., & Frankel, A. S. (2016). Assessment of complex dissociative disorder patients and simulated dissociation in forensic contexts. *International Journal of Law and Psychiatry*, 11, 513–520. <https://doi.org/10.1016/j.ijlp.2016.10.006>.
- Briere, J. (1996). *Trauma symptom checklist for children: Professional manual*. Odessa, FL: Psychological Assessment Resources, Inc..
- Briere, J. (2001). *DAPS—Detailed Assessment of Posttraumatic Stress Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Briere, J., Elliott, D. M., Harris, K., & Cotman, A. (1995). Trauma symptom inventory: Psychometrics and association with childhood and adult victimization in clinical samples. *Journal of Interpersonal Violence*, 10, 387–401.
- Buhmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon’s mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3–5.
- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegan, A., & Dahlstrom, W. G. (2001). *Manual for the administration and scoring of the MMPI-2*. Minneapolis, MN: University of Minnesota Press.
- Carlson, E. B., Dalenberg, C., & McDade-Montez, E. (2012). Dissociation in posttraumatic stress disorder part I: Definitions and review of research. *Psychological Trauma: Theory, Research, Practice, and Policy*, 4(Supplemental), 479–489. <https://doi.org/10.1037/a0027748.supp>.
- Coe, M. T., Dalenberg, C. J., Aransky, K. M., & Reto, C. S. (1995). Adult attachment style, reported childhood violence history and types of dissociative experiences. *Dissociation: Progress in the Dissociative Disorders*, 8, 142–154.
- 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, 550–588. <https://doi.org/10.1037/a0027447>.
- Dalenberg, C. J., Coe, M. T., Reto, C. S., Aransky, K. M., Duvenage, C., & Weber, R. (1994). The development of a measure of dissociation for use on general psychiatric and nonpsychiatric populations. Paper

- presented at the Eighth Annual Conference on Responding to Child Maltreatment, San Diego, California.
- Dawes, R., Faust, D., & Meehl, P. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674. <https://doi.org/10.1126/science.2648573>.
- Dawes, S. E., Suarez, P., Vaida, F., Marcotte, T. D., Atkinson, J. H., Grant, I., et al. (2010). Demographic influences and suggested cut-scores for the Beck Depression Inventory in a non-clinical Spanish speaking population from the US-Mexico border region. *International Journal of Culture and Mental Health*. <https://doi.org/10.1080/17542860903533640>.
- Dubester, K. A., & Braun, B. G. (1995). Psychometric properties of the Dissociative Experiences Scale. *Journal of Nervous and Mental Disease*, *183*, 231–235. <https://doi.org/10.1097/00005053-199504000-00008>.
- Flitter, J. M. K., Elhai, J. D., & Gold, S. N. (2003). MMPI-2 F scale elevations in adult victims of child sexual abuse. *Journal of Traumatic Stress*, *16*, 269–274.
- Foschi, M. (2014). Hypotheses, operationalizations, and manipulation checks. In *Laboratory experiments in the social sciences* (second ed.). Cambridge, MA: Academic Press. <https://doi.org/10.1016/B978-0-12-404681-8.00011-X>.
- Frischholz, E. J., Braun, B. G., Sachs, R. G., Hopkins, L., et al. (1990). The Dissociative Experiences Scale: Further replication and validation. *Dissociation: Progress in the Dissociative Disorders*, *3*, 151–153.
- Gillis, M. M., Haaga, D. A. F., & Ford, G. T. (1995). Normative values for the Beck anxiety inventory, fear questionnaire, Penn State worry questionnaire, and social phobia and anxiety inventory. *Psychological Assessment*, *7*, 450–455.
- Ginzburg, K., Somer, E., Tamarkin, G., & Kramer, L. (2010). Clandestine psychopathology: Unrecognized dissociative disorders in inpatient psychiatry. *The Journal of Nervous and Mental Disease*, *198*, 378–381.
- Gray, M. J., Elhai, J. D., & Briere, J. (2010). Evaluation of the Atypical Response Scale of the Trauma Symptom Inventory-2 in detecting simulated posttraumatic stress disorder. *Journal of Anxiety Disorders*, *24*, 447–451. <https://doi.org/10.1016/j.janxdis.2010.02.011>.
- Haagen, J. F. G., van Rijn, A., Knissee, J. W., van der Aa, N., & Kleber, R. J. (2018). The dissociative post-traumatic stress disorder (PTSD) subtype: A treatment outcome cohort study in veterans with PTSD. *British Journal of Clinical Psychology*, *57*, 201–222. <https://doi.org/10.1111/bjc.12169>.
- Johnson, J. A. (2005). Ascertaining the validity of individual protocols from Webbased personality inventories. *Journal of Research in Personality*, *39*, 103–129. <https://doi.org/10.1016/j.jrp.2004.09.009>.
- Kluemper, N. S., & Dalenberg, C. (2014). Is the dissociative adult suggestible? A test of the trauma and fantasy models of dissociation. *Journal of Trauma and Dissociation*, *15*, 457–476.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606–613.
- Lowe, B., Unutzer, J., Callahan, C. M., Perkins, A. J., & Kroenke, K. (2004). Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical Care*, *42*, 1194–1201.
- 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*, 37–46. <https://doi.org/10.1176/appi.ajp.2017.17010025>.
- Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *General Hospital Psychiatry*, *28*, 71–77. <https://doi.org/10.1016/j.genhospsych.2005.07.003>.
- Martin, P. K., Schroeder, R. W., Olsen, D. H., Maloy, H., Boettcher, A., Ernst, N., & Okut, H. (2019). A systematic review and meta-analysis of the test of memory malingering in adults: Two decades of deception detection (p. 132). *The Clinical Neuropsychologist*. <https://doi.org/10.1080/13854046.2019.1637027>.
- McNally, R. (2007). Dispelling confusion about traumatic dissociative amnesia. *Mayo Clinic Proceedings*, *82*, 1083–1087. <https://doi.org/10.4065/82.9.1083>.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*, 437–455. <https://doi.org/10.1037/a0028085>.
- Melton, G., Petrila, J., Poythress, J., & Slobogin, C. (2007). *Psychological evaluations for the courts* (3rd ed.). New York, NY: Guilford.
- Merckelbach, H., & Patihis, L. (2018). Why trauma-related dissociation is a misnomer in courts: A critical analysis of Brand et al. (2017a, b). *Psychological Injury and Law*, *11*, 370–376. <https://doi.org/10.1007/s12207-018-9328-8>.
- Millon, T., Grossman, S., & Millon, C. (2015). *Millon Clinical Multiaxial Inventory-IV manual*. Minneapolis, MN: Pearson Assessments.
- Millon, T., Millon, C., Davis, R. D., & Grossman, S. (1994). *Millon Clinical Multiaxial Inventory-III (MCMI-III): Manual*. Minneapolis, MN: Pearson/PsychCorp.
- Olsen, S. A., Clapp, J. D., Parra, G. R., & Beck, J. G. (2013). Factor structure of the dissociative experiences scale: An examination across sexual assault status. *Journal of Psychopathology and Behavioral Assessment*, *35*, 394–403. <https://doi.org/10.1007/s10862-013-9347-4>.
- Palermo, C. A., & Brand, B. L. (2018). Can the trauma symptom inventory-2 distinguish coached simulators from dissociative disorder patients? *Psychological Trauma: Theory, Research, Practice, and Policy*, *11*, 477–485. <https://doi.org/10.1037/tra0000382>.
- Patihis, L., & Lynn, S. J. (2017). Psychometric comparison of dissociative experiences scales II and C: A weak trauma-dissociation link. *Applied Cognitive Psychology*, *31*, 392–403. <https://doi.org/10.1002/acp.3337>.
- Pinto-Meza, A., Serrano-Blanco, A., Peñarrubia, M. T., Blanco, E., & Haro, J. M. (2005). Assessing depression in primary care with the PHQ-9: Can it be carried out over the telephone? *Journal of General Internal Medicine*, *20*, 738–742.
- Piper, A., & Merskey, H. (2004). The persistence of folly: A critical examination of dissociative identity disorder. Part I. The excesses of an improbable concept. *Canadian Journal of Psychiatry*, *49*, 592–600.
- Pope, H. G., Poliakoff, M. B., Parker, M. P., Boynes, M., & Hudson, J. I. (2007). Is dissociative amnesia a culture-bound syndrome? Findings from a survey of historical literature. *Psychological Medicine*, *37*, 225–233.
- Resnick, P. J., West, S. G., & Wooley, C. N. (2018). *The malingering of posttraumatic disorders*. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (p. 188–211). The Guilford Press.
- Rief, W., Nanke, A., Klaiberg, A., & Braehler, E. (2004). Base rates for panic and depression according to the Brief Patient Health Questionnaire: A population-based study. *Journal of Affective Disorders*, *82*, 271–276.
- Ross, C. A. (1991). Epidemiology of multiple personality disorder and dissociation. *Psychiatric Clinics of North America*, *14*, 503–517.
- Şar, V., Akyüz, G., & Doğan, O. (2007). Prevalence of dissociative disorders among women in the general population. *Psychiatry Research*, *149*, 169–176.
- Soukup, J., Papežová, H., Kuběna, A. A., & Mikolajová, V. (2010). Dissociation in non-clinical and clinical sample of Czech adolescents reliability and validity of the Czech version of the Adolescent Dissociative Experiences Scale. *European Psychiatry*, *25*, 390–395. <https://doi.org/10.1016/j.eurpsy.2010.03.011>.

- Tarescavage, A. M., & Glassmire, D. M. (2016). Differences between structured interview of reported symptoms (SIRS) and SIRS-2 sensitivity estimates among forensic inpatients: A criterion groups comparison. *Law and Human Behavior, 40*, 488–502. <https://doi.org/10.1037/lhb0000191>.
- Thomas, K. A., & Clifford, S. (2017). Validity and mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior, 77*, 184–197. <https://doi.org/10.1016/j.chb.2017.08.038>.
- Tsai, J., Armour, C., Southwick, S. M., & Pietrzak, R. H. (2015). Dissociative subtype of DSM-5 posttraumatic stress disorder in U.S. veterans. *Journal of Psychiatric Research, 66*, 67–74. <https://doi.org/10.1016/j.jpsychires.2015.04.017>.
- Van der Hart, O., Nijenhuis, E. R., & Steele, K. (2005). Dissociation: An insufficiently recognized major feature of complex posttraumatic stress disorder. *Journal of Traumatic Stress, 18*, 413–423.
- Waelde, L. C., Silvern, L., & Fairbank, J. A. (2005). A taxometric investigation of dissociation in Vietnam veterans. *Journal of Traumatic Stress, 18*, 359–369. [https://doi.org/10.1002/\(ISSN\)1573-6598](https://doi.org/10.1002/(ISSN)1573-6598).
- Waller, N. G., Putnam, F. W., & Carlson, E. B. (1996). Types of dissociation and dissociative types: A taxometric analysis of dissociative experiences. *Psychological Methods, 1*, 300–321.
- Walls, B. D., Wallace, E. R., Brothers, S. L., & Berry, D. T. R. (2017). Utility of the Conners' Adult ADHD Rating Scale validity scales in identifying simulated attention-deficit hyperactivity disorder and random responding. *Psychological Assessment, 29*, 1437–1446. <https://doi.org/10.1037/pas0000530>.
- Wolf, E. J., Lunney, C. A., Miller, M. W., Resick, P. A., Friedman, M. J., & Schnurr, P. P. (2012). The dissociative subtype of PTSD: A replication and extension. *Depression and Anxiety, 29*, 679–688.
- Wolf, E. J., Miller, M. W., Reardon, A. F., Ryabchenko, K. A., Castillo, D., & Freund, R. (2012). A latent class analysis of dissociation and posttraumatic stress disorder: Evidence for a dissociative subtype. *Archives of General Psychiatry, 69*, 698–705.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.