

Revisiting the Factor Structure and Construct Validity of the Cognitive Failures Questionnaire

Zachary T. Goodman, Kiara R. Timpano, Maria M. Llabre, and Sierra A. Bainter
Department of Psychology, University of Miami

The Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982) is an established and commonly used self-report measure of cognitive errors experienced in daily life, capturing perceived difficulties with forgetfulness, distractibility, and thinking blunders. Despite frequent use in clinical research and established associations with psychological and neuropsychological disorders, the psychometric properties and construct validity of the CFQ remain ambiguous. This study sought to critically assess the factor structure and external validity of the CFQ. A sample of 839 people (62% female) between 16 and 85 years of age ($M = 44.12$, $SD = 19.54$) was drawn from the Nathan Kline Institute-Rockland Sample. Previously published CFQ factor structures were compared via confirmatory factor analysis and the unique variance explained by each factor was assessed. Next, we related the CFQ latent variables to neuropsychological tasks and symptom measures of depression, anxiety domains, inattention, hyperactivity, and impulsivity. A single-factor model was best supported by the data, indicating that the CFQ represents a global measure of subjective cognitive difficulties rather than errors in specific domains. Scores on the CFQ did not predict poorer performance on objective neuropsychological tasks but were related to a range of psychological distress symptoms. Subscales derived from previously published factor structures may provide misleading impressions of the construct validity of the CFQ and are not recommended for use in future research or clinical contexts.

Public Significance Statement

The Cognitive Failures Questionnaire (CFQ) is a commonly used self-report measure to capture subjective cognitive difficulties experienced throughout the day. Our study suggests that these subjective difficulties are unrelated to underlying deficits in cognition, as traditionally measured, and instead may simply reflect elevated psychological distress. The method used to assess daily cognitive difficulties can have a considerable impact on the conclusions drawn and should be selected with intention and care.

Keywords: cognitive failures, psychometrics, neuropsychology, psychological distress

Supplemental materials: <https://doi.org/10.1037/pas0001127.supp>

Cognitive difficulties have traditionally been measured using in-person neuropsychological assessments or experimental tasks, such as the Stroop Task or Trail Making Test (TMT), among others. Task performance is derived from number of errors, completion time, or

similar variables, and is thought to be an objective indicator of cognitive function in one or more domains (Ashendorf et al., 2008; Balota et al., 2010). However, task-based cognitive assessments present limitations, including potentially lengthy batteries of tasks,

This article was published Online First April 4, 2022.

Zachary T. Goodman  <https://orcid.org/0000-0002-0126-4425>

Kiara R. Timpano  <https://orcid.org/0000-0002-0665-8722>

Maria M. Llabre  <https://orcid.org/0000-0002-7450-2498>

The authors have no conflicts of interest to report. All procedures necessary to replicate these analyses are available via request. The study was not preregistered. Data are publicly available through the Nathan Kline Institute. Zachary T. Goodman was responsible for conceptualization, analyses, methodology, and visualization. Zachary T. Goodman, Kiara R. Timpano, Maria M. Llabre, and Sierra A. Bainter were responsible for writing, draft review and editing, supervision, and administration.

Zachary T. Goodman is supported by the National Institutes of Health (T32-HL007426). Sierra A. Bainter is supported by the National Institutes of Health (K01 MH122805).

Zachary T. Goodman played lead role in conceptualization, data curation,

formal analysis, investigation, methodology, and writing of original draft and equal role in writing of review and editing. Kiara R. Timpano played lead role in writing of review and editing, and supporting role in conceptualization, formal analysis, funding acquisition, investigation, methodology, validation, and writing of original draft and equal role in project administration and supervision. Maria M. Llabre played lead role in funding acquisition and supporting role in conceptualization, formal analysis, methodology, validation, and writing of review and editing. Sierra A. Bainter played lead role in resources, software, and validation, and supporting role in conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, writing of original draft, and writing of review and editing and equal role in project administration and supervision.

Correspondence concerning this article should be addressed to Zachary T. Goodman, Department of Psychology, University of Miami, 5665 Ponce de Leon Boulevard, Office 446, Coral Gables, FL 33146-0751, United States. Email: ztg8@miami.edu

the need for trained researchers or clinicians to administer and score the assessments, and concerns about the ecological validity and clinical application (Burgess et al., 1998). As such, self-report measures of subjective cognitive functioning represent an appealing complement to task-based measurement for researchers and clinicians alike. Self-report measures are grounded in the recognition that subjective cognitive complaints are a diagnostic feature of global neurocognitive deficits and disorders (American Psychiatric Association, 2013). The assumption is that these minor mental slips can be indicative of early cognitive impairment (Clément et al., 2008) and possible precursors to more significant cognitive decline associated with neurodegenerative processes (Jessen et al., 2020; Jorm et al., 2001; St John & Montgomery, 2002; Wang et al., 2004). Furthermore, self-reported cognitive failures are linked with affective distress and psychiatric diagnoses (Hill et al., 2016), and in some instances, such as in major depressive disorder, they reflect key diagnostic criteria (American Psychiatric Association, 2013). Consequently, several self-report measures of cognitive failures have been developed, which offer unique clinical utility when considered alongside task-based measures while simultaneously assessing day-to-day difficulties.

The Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982) is the most widely used measure of subjective cognitive failures (Carrigan & Barkus, 2016). It is a self-report questionnaire with items that capture subjective failures in perception, memory, and coordination affecting daily functioning. The CFQ was developed to provide a general measure of daily cognitive failures or mental slips (Broadbent et al., 1982). Broadbent et al. (1982) noted in their original report that the CFQ consistently demonstrated the presence of a general factor, that multidimensional solutions were highly variable and sample dependent, and advised against reporting multidimensional solutions for the CFQ.

Despite the recommendation to focus on a general factor, many multidimensional models have been proposed for the CFQ ranging from 2 (Larson et al., 1997; Matthews et al., 1990) to 5 factors (Bridger et al., 2013; Pollina et al., 1992). The two most frequently used multidimensional solutions are three- and four-factor models (Rast et al., 2009; Wallace et al., 2002); items representing these models are directly compared in Table 1. Wallace et al. identified four components which they labeled *Memory*, *Distractibility*, *Blunders*, and *Names*, via principal components analysis (2002) and confirmatory factor analyses (2004). Wallace (2004) noted that the factors were highly correlated, supporting the construct of general cognitive failures as captured by the total score. Rast et al. (2009), in turn, recommended a solution with three factors labeled as *Forgetfulness*, *Distractibility*, and *False Triggering*.

Besides disagreement about the dimensionality of the CFQ, an additional concern is that CFQ items do not consistently load onto similar factors across the various published solutions (Table 1). For example, indicators of *Distractibility* in the three-factor solution load onto either the *Memory* or *Blunders* factors in the four-factor solution—of the 14 items on the three-factor solution's *Distractibility* factor, only five are represented on the four-factor solution's *Distractibility* factor. Additionally, although the *Forgetfulness* factor (derived from the three-factor solution) and *Distractibility* factor (derived from the four-factor solution) are described as measuring similar constructs, the items that measure each factor are different.

Importantly, when subscales are suggested to represent cognitive processes, empirical evidence is necessary to support such labeling

(American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). Moreover, subscales should demonstrate consistent associations with theoretically relevant constructs, including instruments assessing the same construct (AERA, APA, & NCME, 2014). Thus far, psychometric works have yet to support the validity of CFQ factors as representative of the constructs for which they are named. Considered jointly, the extant literature has not clarified how the CFQ should be appropriately implemented or interpreted in research or clinical contexts (Carrigan & Barkus, 2016). As a result, the true conceptual meaning of the CFQ, and potential subfactors, remains ambiguous.

The ambiguity in dimensionality of the CFQ also raises critical concerns about the validity and distinctiveness of subscales based on the various factors. The use of indistinct subscales is known to lead to faulty or entirely misleading results (Haberman, 2008; Reise et al., 2013). Further, if dimensions of the CFQ are not distinct, inferences about the external validity of the CFQ and other psychological constructs may be obfuscated or confounded entirely (AERA, APA, & NCME, 2014; Lynam et al., 2006).

Two separate but related literatures relying on the CFQ illustrate the ambiguity surrounding its use and subsequent conclusions. First is a body of work examining the relationship between the CFQ and objective measures of cognitive abilities (Carrigan & Barkus, 2016). Second is a literature examining the relationship between the CFQ, its subscales, and psychopathological symptoms (Carrigan & Barkus, 2016).

Findings concerning the relationship between the CFQ and objective cognitive functioning are mixed, and at least one explanation may lie in how the CFQ is inconsistently utilized across studies. One meta-analysis concluded that although the relationship between CFQ scores and sustained attention is highly variable, there is a modest significant association (Smilek et al., 2010). On the other hand, several other studies have failed to establish relationships between CFQ scores and cognitive domains (Carrigan & Barkus, 2016). CFQ scores have previously shown no relation to executive functions, working memory, or attention (Brück et al., 2019; Shakeel & Goghari, 2017), but have been associated with better visuospatial perception (Poliakoff & Smith-Spark, 2008). Of note, across this body of work, while some studies rely on the CFQ total score (e.g., Forster & Lavie, 2007; Hohman et al., 2011), others examined CFQ subscales (e.g., Smilek et al., 2010; van der Werf-Eldering et al., 2011; Wang et al., 2004). This body of literature therefore highlights that although research examining the CFQ in relation to neuropsychological functioning continues, ambiguity surrounding the psychometric properties of the CFQ, including proposed subscales, makes it challenging to draw any firm conclusions. The obscurity present in research relying on the CFQ reflects broader uncertainty in the extent to which subjective cognitive complaints reflect bona fide impairments in cognition (Burmester et al., 2016).

A second literature that is compromised by ambiguity in the factor structure of the CFQ are studies examining the relationship between the CFQ, its subscales, and psychopathological symptoms (Carrigan & Barkus, 2016). Similar to the literature investigating cognitive correlates of the CFQ, some of these investigations rely on the CFQ total score as prescribed by Broadbent (e.g., Brück et al., 2019; Hart et al., 2005), others examine subscales based on the various alternative factor structures (e.g., Mahoney et al., 1998;

Table 1
 Organization of the Factor Structure of the CFQ Identified in Previous Psychometric Analyses

#	CFQ item	Single-factor	Three-factor			Four-factor			
			Forget.	Distract.	F.T.	Distract.	Memory	Blunders	Names
1.	Do you read something and find you have not been thinking about it and must read it again?	X	X		X	X			
2.	Do you find you forget why you went from one part of the house to the other?	X	X	X	X	X			
3.	Do you fail to notice signposts on the road?	X			X	X			
4.	Do you find you confuse right and left when giving directions?	X			X	X			
5.	Do you bump into people?	X	X	X	X			X	
6.	Do you find you forget whether you have turned off a light or a fire or locked the door?	X	X		X		X		
7.	Do you fail to listen to people's names when you are meeting them?	X	X	X					X
8.	Do you say something and realize afterwards that it might be taken as insulting?	X		X				X	
9.	Do you fail to hear people speaking to you when you are doing something else?	X		X				X	
10.	Do you lose your temper and regret it?	X		X				X	
11.	Do you leave important letters unanswered for days?	X		X				X	
12.	Do you find you forget which way to turn on a road you know well but rarely use?	X			X		X		
13.	Do you fail to see what you want in a supermarket (although it is there)?	X	X		X		X		
14.	Do you find yourself suddenly wondering whether you have used a word correctly?	X		X				X	
15.	Do you have trouble making up your mind?	X		X	X	X			
16.	Do you find you forget appointments?	X	X	X			X		
17.	Do you forget where you put something like a newspaper or a book?	X	X		X		X		
18.	Do you find you accidentally throw away the thing you want and keep what you meant to throw away—as in the example of throwing away the matchbox and putting the used match in your pocket?	X		X	X		X		
19.	Do you daydream when you ought to be listening to something?	X		X		X			
20.	Do you find you forget people's names?	X	X		X				X
21.	Do you start doing one thing at home and get distracted into doing something else (unintentionally)?	X	X	X		X			
22.	Do you find you cannot quite remember something although it is on the tip of your tongue?	X	X			X			
23.	Do you find you forget what you came to the shops to buy?	X	X		X		X		
24.	Do you drop things?	X			X			X	
25.	Do you find you cannot think of anything to say?	X		X		X			

Note. Forget. = forgetfulness; Distract. = distractibility; F.T. = false triggering; single-factor (Broadbent et al., 1982); three-factor (Rast et al., 2009); four-factor (Wallace et al., 2002). Questions are reproduced from Broadbent et al. (1982) with permissions from John Wiley and Sons.

Matthews et al., 1990; Weintraub et al., 2018). Results across studies suggest that subjective cognitive failures correlate more strongly with measures of psychological distress, anxiety, depression, and posttraumatic stress symptoms compared to broad neuropsychological functioning (Balash et al., 2013; Brück et al., 2019; Wagle et al., 1999). Generally, these findings support Broadbent's original theory (1982) that individuals who react more negatively to stressful events may be more likely to notice and report daily errors. However, other studies that use the alternative factor structures draw more specific conclusions by linking psychological symptoms to specific cognitive domains, which they have interpreted based on the naming of the various subscales. For example, Weintraub et al. (2018) interpreted the association between hoarding and *Forgetfulness* and *Distractibility* factors as indicative of relationships between psychopathology and cognitive functions related to attention and memory. Overall, inconsistency in the factor solutions

implemented, in addition to ambiguous construct validity, makes drawing inferences regarding the association between psychopathology and cognitive failures difficult.

Taken as a whole, the extant literature on the CFQ highlights the continued unresolved uncertainty surrounding the appropriate factor structure, with direct and serious implications for how the CFQ is implemented and interpreted in applied research. The first aim of the present study was to investigate the internal structure of the CFQ, including the psychometric validity of distinct dimensions. We first endeavored to establish the dimensionality of the CFQ, whether unidimensional or multidimensional, by comparing the original single-factor model (Broadbent et al., 1982) to the three- (Rast et al., 2009) and four-factor (Wallace et al., 2002) models most commonly used in the literature in a large, normative lifespan sample.

Our second aim was to examine the external validity of the CFQ structure best supported in the first aim's analyses. In line with best

practices and standards (AERA, APA, & NCME, 2014), we examined support for conceptualizing the CFQ as a measure of cognitive domains by examining the associations between CFQ factors and neuropsychological tasks of executive functioning, verbal learning and recall, working memory, and attention. We also examined associations with a range of psychological distress measures, including symptoms of depression, anxiety, obsessive-compulsive disorder, attention-deficit/hyperactivity disorder, and broad impulsivity.

Our final aim was to consider the same associations between unsubstantiated factor structures and the aforementioned external criteria. For each factor, zero-order correlations and residualized associations are contrasted. Zero-order correlations represent the strength of the relationship between subscale factors and external criteria without consideration of variance explained by other subscale factors. Residualized associations are associations which partial out variance better explained by other subscale factors from the same model. In doing so, we sought to examine the propensity of inappropriate factor structures to produce spurious relationships to both neuropsychological and psychopathological outcomes. This exercise also serves to highlight inconsistent relationships depending on whether zero-order or residualized effects are considered, an outcome which instills low confidence in the construct validity of subscales (Lynam et al., 2006). Critically, we demonstrate the extent to which scores drawn from unsubstantiated multidimensional models serve to perpetuate ambiguity surrounding the CFQ and subjective cognitive complaints more broadly. In total, this study takes a comprehensive approach to clarifying the factor structure and nomological network of the CFQ for future implementation in research contexts.

Method

Participants and Procedure

The present study sampled 839 individuals (62% female) from the Nathan Kline Institute-Rockland Sample (NKI-RS; Nooner et al., 2012) who completed a battery of questionnaires, along with several neurocognitive measures. Ages ranged from 16 to 85 years old ($M = 44.12$, $SD = 19.54$). Twelve percent of the sample identified as Hispanic; 73% participants identified as White, 18% as Black or African American, 5% as Asian American/Pacific Islander, 1% as American Indian/Native American, and 3% as another racial group. Participants had an average of 15.36 years of education ($SD = 2.26$)—1% did not complete at least high school, 14% completed 2 years of college, 21% completed a bachelor's degree, and 24% completed a master's or doctoral degree. Full scale intelligence quotient (FSIQ) scores were estimated via the Wechsler Abbreviated Scale of Intelligence, 2nd edition (Wechsler, 2011). The sample demonstrated average FSIQ scores of 101.04 ($SD = 13.08$), with 5% of the sample scoring 80 or below. Previous diagnoses of intellectual disability were not known.

Staff-administered Structured Clinical Interviews for Diagnostic and Statistical Manual, Fourth Edition, Text Revision (DSM-IV-TR) Axis I Disorders—Nonpatient Edition (First et al., 2002) and Adult ADHD Clinical Diagnostic Scales (Kessler et al., 2010) were utilized to collect data on psychiatric diagnoses: 3.5% of participants met criteria for a current depressive disorder, 2.7% met criteria for a current anxiety disorder, and 1.4% met criteria for attention-deficit/

hyperactivity disorder. Medical history was gathered via a comprehensive medical history questionnaire and staff-administered medical condition form (Nooner et al., 2012). Of potential relevance, 5.6% of participants reported hypertension and 3.0% reported diabetes mellitus; no participants reported a history of neurological conditions. During data collection, some cognitive tests were added to the battery as the study progressed. As a result, some measures have more or less data available (see Measures section and Table 2)—these data are assumed to be missing completely at random. Data acquisition and use were approved by the university's institutional review board.

Measures

Cognitive Failures Questionnaire

The CFQ (Broadbent et al., 1982) is a 25-item assessment of subjective cognitive errors made in daily life. The CFQ includes questions assessing failures in the areas of memory and forgetfulness, attention and distractibility, and clumsiness (see Table 1). Items were originally assessed on a 5-point Likert-type scale with higher scores indicating more frequent errors; however, the third and

Table 2
Descriptive Statistics of Continuous Demographic Data and Study Variables

Variable	<i>n</i>	<i>M</i>	<i>SD</i>
Demographic data			
Age	839	44.12	19.54
CFQ total score	839	31.71	12.70
Years of education	794	15.36	2.26
FSIQ	834	101.04	13.08
Cognitive outcomes			
D-KEFS TMT	805	10.54	2.95
D-KEFS CST	462	10.96	3.00
D-KEFS letter fluency	793	11.11	3.56
D-KEFS category fluency	793	11.74	3.44
RAVLT trial 1	492	6.17	1.78
RAVLT trial 5	492	11.90	2.22
RAVLT immediate recall	492	9.86	3.25
RAVLT delayed recall	492	9.53	3.53
DS forward	480	8.75	1.91
DS backward	480	6.62	2.22
Psychological outcomes			
BDI-II	619	6.13	7.09
GDS	164	3.85	3.79
STAI	798	34.91	10.30
Y-BOCS obsessions	635	0.22	1.32
Y-BOCS compulsions	635	0.34	1.71
CAARS inattention/memory	800	3.65	2.71
CAARS hyperactivity/restlessness	800	4.11	2.68
CAARS impulsivity/emotional lability	800	2.13	2.17
CAARS problems w/self-concept	800	4.00	3.21
UPPS positive urgency	795	23.71	6.62
UPPS negative urgency	795	19.62	4.41
UPPS lack of premeditation	795	17.85	4.45
UPPS lack of perseverance	795	29.17	8.01
UPPS sensation seeking	795	21.68	7.42

Note. FSIQ = full scale IQ; D-KEFS = Delis-Kaplan Executive Function System; RAVLT = Rey Auditory Verbal Learning Test; DS = digit span; BDI = Beck Depression Inventory; GDS = Geriatric Depression Scale; Y-BOCS = Yale-Brown Obsessive Compulsive Scale; CAARS = Conners' Adult ADHD Rating Scales; UPPS = Urgency, Premeditation, Perseverance, Sensation Seeking Impulsive Behavior Scale; STAI = State-Trait Anxiety Inventory.

fourth response options were combined due to low endorsement of the highest option (0.1%–10.7%). Further, response options with low endorsement demonstrated poor biserial correlations and discrimination, supporting this adjustment—a similar correction has been necessary due to low endorsement rates (Rast et al., 2009). Reliabilities for each factor structure are reported in the Results section.

Delis–Kaplan Executive Function System

The Delis–Kaplan Executive Function System (D–KEFS; Delis et al., 2001) is a comprehensive battery of executive functioning tests frequently utilized in research and clinical settings. Broadly, the D–KEFS has been validated in clinical and research settings as a measure of executive functioning abilities and impairments experienced in a wide range of psychopathology, neurological injury, and cognitive decline due to aging (Homack et al., 2005; Karr et al., 2019). Performance on each D–KEFS test was transformed into age-adjusted scaled scores ($M = 10$, $SD = 3$) based on norms published with the test manual (Delis et al., 2001). Higher scores indicate better performance in each domain. Four tests were used to assess differing aspects of executive functioning.

Trail Making Test. During the TMT, participants are required to connect dots filled with numbers and/or letters in order. Although the TMT has several conditions, we relied on the primary executive-function condition, *Number–Letter Switching* (Delis et al., 2001). In this condition, participants must alternate between connecting numbers and letters, with performance based on total completion time. The *Number–Letter Switching* condition is an assessment of planning, set shifting, visual–motor scanning and sequencing, and psychomotor speed, and is sensitive to subtle deficits in cognitive flexibility (Delis et al., 2001).

Card Sorting Test. On the Card Sorting Test (CST), participants match six cards, displaying stimulus words as well as perceptual features, into two groups of three cards each. During the *Free Sorting* condition, participants are given up to 4 min to categorize cards based on either visuospatial or semantic properties. CST performance is indicative of set-shifting ability, problem-solving, sustained attention, and perseveration. The correct number of sorts completed was used as the indicator of performance on this subtest.

Color Word Interference Test. The Color Word Interference Test (CWI) requires participants to read a list of names of colors which are printed in a different color's ink. For example, the word “yellow” may be written in red ink. Depending on the condition, participants may be required to read the word itself or state the color in which the word is written. Time to completion on the *Inhibition/Switching* condition was used in this study, during which participants alternate between reading aloud the written word and stating the color of ink. The *Inhibition/Switching* condition assesses cognitive flexibility and inhibition.

Verbal Fluency. Verbal fluency is measured with a short collection of tasks assessing access to overlearned verbal information in addition to processing speed, initiation, and self-monitoring. During *Letter Fluency* participants must generate as many words for a given letter as they can within 60 s, repeated for three trials. Letter fluency measures rapid retrieval of lexical knowledge based on shared orthographic properties rather than semantic knowledge. During the *Category Fluency* task, participants generate words within two semantic categories over 60 s per category. Performance

is based on the total number of correct words generated across both trials.

Rey Auditory Verbal Learning Test

The Rey Auditory Verbal Learning Test (RAVLT; Schmidt, 1996) is a verbal learning and memory task. During learning trials, participants are presented with 15 words per trial and asked to recall as many words as they can. After five learning trials, a distractor list of 15 words is presented, after which participants are asked to recall words from the *first* list (i.e., immediate recall). After a 20-min delay, participants are again asked to recall words from the *first* list (i.e., delayed recall). Performance on learning and recall trials have been associated with normative aging (Mitrushina et al., 1991) as well as dementia (Tierney et al., 1994), and the RAVLT sensitive to structural damage of the left temporal lobe (Redoblado et al., 2003). Age-adjusted scores were used for each trial.

Digit Span

Two *Digit Span* subtests were drawn from the Wechsler Adult Intelligence Scale (WAIS-IV; Wechsler, 2008). *Digit Span Forward* (DSF) requires participants to store digits in working memory and repeat them in the same order as during presentation. *Digit Span Backward* (DSB) also requires participants to store digits but repeat them in the reverse order. Age-adjusted scores were used for both DSF and DSB performance.

Beck Depression Inventory-II

The Beck Depression Inventory-II (BDI-II; Beck et al., 1996) is a measure of depressive symptoms widely implemented in clinical and research contexts and was administered to participants aged 18–65 years ($n = 619$). Participants answer 21 questions assessing a variety of cognitive, behavioral, and emotional symptoms frequently observed in depressive disorders as well as across other psychopathologies. Responses were measured on a 4-point scale and summed ($M = 6.13$, $SD = 7.09$), with higher scores indicating greater symptom severity. The BDI-II has been well validated as a measure of depressive symptoms (Wang & Gorenstein, 2013) and demonstrated excellent reliability in this sample ($\alpha = .92$).

Geriatric Depression Scale

The Geriatric Depression Scale (GDS; Yesavage et al., 1982–1983) is a screening assessment of depressive symptom severity designed specifically for healthy as well as cognitively impaired older adults. In this sample, participants over 65 years of age completed the GDS as opposed to the BDI-II ($n = 164$). Participants responded to 30 yes-or-no questions, with higher endorsement indicating greater symptomatology ($M = 3.85$, $SD = 3.79$). The GDS has shown strong external validity with other assessments of depressive symptoms in older adults (Kørner et al., 2006) and demonstrated acceptable reliability in this sample ($\alpha = .81$).

State-Trait Anxiety Inventory

The Trait Anxiety subscale of the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1983) is a 20-item scale of a general disposition toward feeling anxious and broad psychological distress.

The STAI is assessed on a 4-point scale, with higher total scores indicating more proneness toward feeling anxious ($M = 34.91$, $SD = 10.30$). The STAI is used widely in psychological research and has been validated as a measure of broad internalization and psychological distress (Balsamo et al., 2013). The Trait subscale of STAI demonstrated strong reliability in the present study ($\alpha = .93$).

Yale-Brown Obsessive Compulsive Scale

The Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman, Price, Rasmussen, Mazure, Fleischmann, et al., 1989) is a 10-item semistructured interview checklist of obsessive and compulsive symptoms administered by trained staff. Interviewers rate the level of distress and impairment caused by each type of symptom on a 5-point scale, with higher scores indicating higher distress. Obsessive symptoms include unwanted, intrusive thoughts that are difficult to suppress or ignore ($M = 0.22$, $SD = 1.32$). Compulsive symptoms are ritualistic behaviors engaged in to reduce the distress of obsessive thoughts ($M = 0.34$, $SD = 1.71$). The psychometric properties and validity of the Y-BOCS were supported during development (Goodman, Price, Rasmussen, Mazure, Delgado, et al., 1989). Both the obsessive ($\alpha = .92$) and compulsive ($\alpha = .95$) subscales demonstrated excellent reliability in this sample.

Conners' Adult ADHD Rating Scales—Self Report: Short Form

The Conners' Adult ADHD Rating Scales (CAARS; Conners et al., 1999) is a brief, 26-item self-report assessment of inattentive and hyperactive symptoms in adulthood. Responses were coded on a 4-point scale, with higher scores indicating more frequent symptoms of inattention or hyperactivity. Four sum-scored subscales of the CAARS were utilized in this study: Inattention/Memory Problems ($M = 0.73$, $SD = 0.54$, $\alpha = .79$), Impulsivity/Emotional Lability ($M = 0.56$, $SD = 0.43$, $\alpha = .73$), Problems with Self-Concept ($M = 0.80$, $SD = 0.64$, $\alpha = .86$), and Hyperactivity/Restlessness ($M = 0.82$, $SD = 0.54$, $\alpha = .72$). The CAARS subscales have demonstrated good internal consistency in past research (Adler et al., 2008).

Urgency, Premeditation, Perseverance, Sensation Seeking Impulsive Behavior Scale

The Urgency, Premeditation, Perseverance, Sensation Seeking Impulsive Behavior Scale (UPPS) is a 59-item measure of five dimensions of impulsive behaviors associated with an array of psychopathologies (Whiteside & Lynam, 2001; Whiteside et al., 2005). The Positive Urgency ($M = 21.68$, $SD = 7.42$, $\alpha = .92$) subscale assesses one's tendency to engage in spontaneous, impulsive behaviors when in a positive affective state (e.g., happy, excited). Conversely, Negative Urgency ($M = 23.71$, $SD = 6.62$, $\alpha = .88$) refers to the tendency to be impulsive when in a negative affective state (e.g., angry, sad). Lack of Premeditation ($M = 19.62$, $SD = 4.41$, $\alpha = .80$) assesses difficulty in considering consequences before engaging in a behavior. Lack of Perseverance ($M = 17.85$, $SD = 4.45$, $\alpha = .79$) refers to the tendency to give up on difficult or tedious tasks. Lastly, the Sensation Seeking ($M = 29.17$, $SD = 8.01$, $\alpha = .87$) subscale measures the preference and openness toward exciting and potentially dangerous situations. Responses were on a

4-point scale and coded so that higher scores indicate a greater impulsivity in each domain.

Statistical Analyses

Psychometric Properties

Confirmatory factor analysis (CFA) was used to examine and compare previously identified factor structures of the CFQ. Analyses were conducted in R Version 3.6.3 with the *lavaan* package (Rosseel, 2012). Diagonally weighted least squares (DWLS) estimation was used as CFQ data are ordinal. Consequently, missing data were removed listwise. The one-factor (Broadbent et al., 1982), three-factor (Rast et al., 2009), and four-factor (Wallace et al., 2002) CFQ models previously supported were tested. On the basis of guidelines by Hu and Bentler (1999), models were considered a good fit to the data if the comparative fit index (CFI) was at or above .95, the root mean square error of approximation (RMSEA) was at or below .06, and the standardized root mean squared residual (SRMR) was at or below .08. Modification indices, which quantify the improvement in model fit should additional parameters be specified (Kline, 2015), were examined when fit was poor. Critically, modification indices are atheoretical in that suggested parameters are ordered by the extent to which included them would reduce the model χ^2 (Kline, 2015). Consequently, only theoretically justifiable respecifications were considered or implemented. Standardized solutions, including factor loadings, correlation coefficients, and regression paths, are reported throughout the results.

Following CFA analyses, the psychometric quality of previously published multidimensional solutions were evaluated via the Haberman (2008) procedure. The Haberman (2008) procedure assesses the unique contribution of each subscale above and beyond relying on the total score, or the value added by using a subscale in place of the total score. When a subscale demonstrates low uniqueness, the total score better represents that specific dimension than the subscale itself (Haberman, 2008). In other words, subscale scores for dimensions with low uniqueness provide less precise estimates than the total score. Reise et al. (2013) emphatically state that when a subscale does not meet this criterion, it is indefensible from a psychometric perspective to calculate, report, or utilize such subscales in research or decision making. This procedure is especially relevant in factor analysis, in which the focus is entirely on the internal structure of the construct with no respect for external validity and ultimately prediction (Davison et al., 2015). We calculated the Value-Added Ratio (VAR; Feinberg & Jurich, 2017), an approximation of the Haberman (2008) procedure, for each of the subscales derived from each factor solution, which provides an index of the unique variance a subscale provides, above and beyond the variance provided via the total score. According to VAR criteria (Feinberg & Jurich, 2017), subscales with VAR less than 0.90 should not be used and are misleading, whereas subscales with VAR over 1.10 provide a potentially meaningful source of variability and may be interpretable.

Nomological Network

Lastly, CFQ factor models were examined in relation to objective cognitive performance and metrics of psychological distress through a series of structural equation models (SEM). Factors from

multidimensional solutions were first examined in isolation; this approach mimics studies where only one subscale of the CFQ is examined in relationship to relevant outcomes (e.g., Kanai et al., 2011), and elucidates simple relationships between latent constructs and external criterion (Lynam et al., 2006). Standardized regression paths (β) are analogous to correlation coefficients, with absolute values closer to one indicating a strong correlation between the CFQ factor and outcomes, whereas absolute values closer to zero indicate weak correlations.

We then examined combined SEM, with all subfactors of the respective multidimensional solutions, to compare similarities and differences in the relationships of the subscales to objective measures of cognitive functioning in the presence of other latent variables (i.e., the unique variance of each subscale when controlling for other types of cognitive failures; Lynam et al., 2006). This approach mimics studies where the unique contribution of each CFQ subscale is examined in relation to a relevant outcome (e.g., Weintraub et al., 2018). This process demonstrates two critical but related issues for establishing the appropriate factor structure of the CFQ: (a) the validity of a measure is determined by consistent relationships between the construct and external criteria when the construct is examined in isolation *as well as* when partialling the variance due to related constructs (Lynam et al., 2006) and (b) factors that are highly collinear result in unstable and spurious coefficients, and can induce empirical underidentification (Rindskopf, 1984). We carry out this exercise as a demonstration of potential spurious or volatile results that may arise from relying on improper factor structures.

All procedures necessary to replicate these analyses are available via request. The study was not preregistered. Data are publicly available through the Nathan Kline Institute.

Results

Psychometric Properties

Previously published models were first examined to explore the psychometric structure of the CFQ. Fit across all three models ranged from marginally good to good (Table 3). The original single-factor model proposed by Broadbent et al. (1982) demonstrated the worst fit (marginally good), whereas the three-factor model proposed by Rast et al. (2009) fit the data the best; however, fit indices between the three- and four-factor models were only marginally different ($\Delta\text{CFI} = .01$, $\Delta\text{RMSEA} < .01$, $\Delta\text{SRMR} < .01$). The single-factor model demonstrated the strongest reliability ($\alpha = .94$), although reliabilities for each factor fell comfortably within

acceptable ranges for the three-factor ($\alpha = .89-.90$) and four-factor ($\alpha = .79-.86$) models.

Although the three-factor model fit best, there were several concerning psychometric properties. The correlations between several factors were concerningly high: *Distractibility* and *False Trigger* ($r = .92$), *Forgetfulness* and *Distractibility* ($r = .82$), and *Forgetfulness* and *False Triggering* ($r = .78$). Further inspection revealed that Item 20 demonstrated a standardized factor loading greater than one onto the *Forgetfulness* factor ($\lambda = 1.10$), indicating an improper solution and empirical underidentification (Chen et al., 2001; Rindskopf, 1984). Additionally, *Forgetfulness* standardized factor loadings on several items were low or very low, specifically Item 6 ($\lambda = .09$), Item 16 ($\lambda = .10$), Item 13 ($\lambda = .19$), and Item 1 ($\lambda = .27$). Item 5 loaded negatively ($\lambda = -.45$), which is problematic when computing reliability and subscale scores, or interpreting the factor, as this item is not intended to be reverse coded. Likewise, several items demonstrated low and/or negative standardized loadings onto the *Distractibility* factor: Item 18 ($\lambda = -.12$), Item 21 ($\lambda = .15$), Item 7 ($\lambda = -.26$), Item 5 ($\lambda = .28$), and Item 2 ($\lambda = -.94$), complicating factor interpretations. Problematic standardized loadings emerged on the *False Triggering* factor as well: Item 15 ($\lambda = -.24$) and Item 20 ($\lambda = -.45$). With respect to the unique variance contributed by each factor, the *Forgetfulness* (VAR = 0.83), *Distractibility* (VAR = 0.79), and *False Triggering* (VAR = 0.81) subscales were all under the minimum value-added threshold.

Similar observations of concerning factor structures emerged when the four-factor solution was examined. The *Names* factor is comprised of only two items, which cannot be estimated in an isolated factor model. Large correlations emerged in the four-factor solution between *Distractibility* and *Memory* ($r = .90$), *Distractibility* and *Blunders* ($r = .91$), and *Memory* and *Blunders* ($r = .88$). The model did not demonstrate concerning standardized factor loadings, such as those noted for the three-factor model; however, the unique variance contributed by each factor, the *Distractibility* (VAR = 0.79), *Memory* (VAR = 0.76), and *Blunders* (VAR = 0.85) subscales fell short of the criteria. The two-item *Names* (VAR = 1.46) subscale was acceptable.

Given the questionable factor loadings and VAR of the multifactor solutions, we next conducted a series of post hoc analyses to explore alternative solutions. First, we examined the eigenvalues from the polychoric correlation matrix for all 25 CFQ items. The eigenvalues demonstrated a large primary factor (EVs = 10.09, 1.51, 1.22, 0.99), suggesting that the CFQ is essentially unidimensional (Embretson & Reise, 2000; Reise et al., 2011). Inspection of the scree plot supports the notion that the CFQ is best represented by a single factor, and not a multidimensional factor structure.

Table 3

Confirmatory Factor Models of Past Psychometric Studies, as Well as a Single-Factor Model With Residual Correlations Between Items With High Residual Covariance

Model	χ^2	<i>df</i>	<i>p</i>	CFI	RMSEA	SRMR	Fit determination
Previously published models							
Single-factor (Broadbent et al., 1982)	1505.68	275	<.001	.92	.07	.06	Marginal
Three-factor (Rast et al., 2009)	927.33	257	<.001	.96	.06	.05	Acceptable
Four-factor (Wallace et al., 2002)	1045.08	269	<.001	.95	.06	.05	Acceptable
Final model							
Adjusted single-factor w/correlated residuals	1083.13	272	<.001	.95	.06	.05	Acceptable

Next, we examined the modification indices for the single-factor model, which suggested covariances between several items of similar content would improve fit. Two residual covariances were included between two-item pairs with similar content; a covariance ($r = .53$) between item 7 (“Do you fail to listen to people’s names when you are meeting them?”) and item 20 (“Do you find you forget people’s names?”), and a covariance ($r = .40$) between Item 5 (“Do you bump into people?”) and Item 24 (“Do you drop things?”). The addition of these covariances improved the single-factor model’s fit substantially (Table 3). The adjusted single-factor model further demonstrated no concerning factor loadings, and fit was comparable to the three- and four-factor solutions.

Nomological Network

Objective Cognitive Functions

To examine the association between the single-factor CFQ model and cognitive functioning, we fit a series of SEM in which the adjusted single-factor CFQ latent variable predicted tasks of objective cognitive functioning. The CFQ latent variable was significantly and positively related to CST ($\beta = .12$, $b = 0.36$, $SE = 0.05$, $p < .001$), CWI ($\beta = .04$, $b = 0.11$, $SE = 0.04$, $p = .008$), and letter fluency ($\beta = .05$, $b = 0.16$, $SE = 0.06$, $p = .003$) performances such that more self-reported cognitive failures were related to *better* performance on tasks of cognitive flexibility, set-shifting, response inhibition, and letter-guided verbal fluency. On the other hand, higher CFQ ratings were negatively and weakly related to category fluency ($\beta = -.04$, $b = -0.14$, $SE = 0.06$, $p = .016$), such that more self-reported cognitive failures were associated with poorer category fluency performance. Lastly, single-factor CFQ scores were unrelated to learning trials, immediate recall, or delayed recall trials of the RAVLT ($p \geq .353$) as well as unrelated to working memory based on digit span (DS) subtests ($p \geq .870$).

Zero-order correlations between a CFQ sum score and neuropsychological outcomes were of similar direction, but attenuated magnitude. CFQ total scores only demonstrated a statistically significant positive correlation with the CST ($r = .11$, $p = .019$), but were uncorrelated with TMT ($r = .01$, $p = .881$), CWI ($r = -.01$, $p = .855$), letter fluency ($r = .06$, $p = .122$), and category fluency ($r = -.02$, $p = .665$), as well as learning or recall trials of the RAVLT ($p \geq .397$).

Psychological Distress

To examine associations between the CFQ and psychological stress and vulnerability, several measures of psychological distress were predicted from the modified single-factor CFQ latent variable. CFQ total scores were moderately and positively related to depressive symptoms in both young to middle aged adults based on BDI-II scores ($\beta = .41$, $b = 2.88$, $SE = 0.27$, $p < .001$) and older adults based on GDS scores ($\beta = .35$, $b = 1.33$, $SE = 0.27$, $p < .001$). The relationship between CFQ total scores and trait anxiety ($\beta = .51$, $b = 5.23$, $SE = 0.35$, $p < .001$) was similarly strong and positive.

There were weak but significant relationships with obsessive ($\beta = .12$, $b = 0.15$, $SE = 0.03$, $p < .001$) and compulsive ($\beta = .14$, $b = 0.24$, $SE = 0.06$, $p < .001$) symptoms. We next examined the link between the CFQ and the CAARS, a measure of attention-deficit/hyperactivity disorder (ADHD) symptoms, which assesses

inattentiveness and hyperactivity. Despite the lack of significant relationships with executive functioning tasks, self-reported cognitive failures were consistently and positively associated with each of the CAARS subscales. CFQ total scores were highly related to the Inattention/Memory Problems subscale ($\beta = .62$, $b = 1.69$, $SE = 0.09$, $p < .001$). Standardized coefficients were also large for the Impulsivity/Emotional Lability ($\beta = .51$, $b = 1.10$, $SE = 0.07$, $p < .001$) and Problems with Self-Concept ($\beta = .49$, $b = 1.57$, $SE = 0.11$, $p < .001$) subscales, and moderate for the Hyperactivity/Restlessness subscale ($\beta = .41$, $b = 1.09$, $SE = 0.08$, $p < .001$).

Regarding specific facets of impulsivity, cognitive failures were positively associated with Negative Urgency ($\beta = .47$, $b = 3.11$, $SE = 0.24$, $p < .001$) and a Lack of Perseverance ($\beta = .45$, $b = 1.99$, $SE = 0.15$, $p < .001$). To a lesser extent, CFQ scores were also related to Positive Urgency ($\beta = .28$, $b = 2.10$, $SE = 0.28$, $p < .001$) and a Lack of Premeditation ($\beta = .25$, $b = 1.10$, $SE = 0.16$, $p < .001$), but not significantly related to Sensation Seeking ($\beta = .04$, $b = 0.30$, $SE = 0.29$, $p = .310$).

As with neuropsychological outcomes, correlations between CFQ total scores and indicators of psychological distress were similar in direction and attenuated in magnitude. Correlations were positive and significant between sum scores and BDI-II ($r = .40$, $p < .001$), GDS ($r = .34$, $p < .001$), trait anxiety ($r = .50$, $p < .001$), obsessive ($r = .11$, $p = .007$) and compulsive symptoms ($r = .13$, $p = .002$), Inattention/Memory Problems ($r = .61$, $p < .001$), Impulsivity/Emotional Lability ($r = .49$, $p < .001$), Problems with Self-Concept ($r = .48$, $p < .001$), and Hyperactivity/Restlessness ($r = .40$, $p < .001$). With respect to impulsivity, CFQ sum scores were again positively correlated with Negative Urgency ($r = .46$, $p < .001$), a Lack of Perseverance ($r = .43$, $p < .001$), Positive Urgency ($r = .28$, $p < .001$), and a Lack of Premeditation ($r = .25$, $p < .001$).

Zero-Order Versus Residualized Associations in Multidimensional Solutions

The CFQ models were examined in relation to objective cognitive performance and metrics of psychological distress through a series of structural equation models. The three-factor solution demonstrated a standardized factor loading greater than 1, implying a misspecified factor structure regardless of fit indices (Kolenikov & Bollen, 2012). Consequently, we do not endorse relying on this factor structure, and report correlations as a demonstration of misleading results that would be obtained should researchers rely on this solution. Zero-order and residualized associations are depicted in Table 4 for the three-factor model and Table 5 for the four-factor model; complete statistics for each of the respective models are provided in the Supplemental Materials.

Considering the misspecified three-factor model results first, the factors were generally positively, but weakly associated with performance across the various cognitive tasks (Table 4). In contrast, when controlling for the other factors, both positive and negative significant associations emerged. Zero-order associations between the CFQ factors and the psychological symptom measures considered were largely positive and moderate. Similar to results observed for the cognitive tasks, residualized associations shifted considerably. In addition to many associations flipping directionality and increasing (or decreasing) in strength, several associations also exceeded acceptable magnitudes. An almost identical pattern was observed for the four-factor model (Table 5).

Table 4

Comparisons of Standardized Structural Paths (β) From Each CFQ Factor (Forget. = Forgetfulness; Dist. = Distractibility; F.T = False Triggering) From the Three-Factor Model (Rast et al., 2009)

			Three-factor model					
			Zero-order correlations			Residualized associations		
			Forget.	Dist.	F.T.	Forget.	Dist.	F.T.
Cognitive functioning	D-KEFS	Trail making	.05	.01	.04	.40	-.50	.15
		Card sorting	.14	.12	.12	.38	-.35	.11
		Color/word int.	.08	.03	.03	.47	-.41	.01
		Category fluency	-.04	-.04	-.04	.10	-.09	-.05
		Letter fluency	.08	.03	.06	.39	-.69	.39
	RAVLT	Trial 1	-.07	.04	-.05	-.27	.93	-.72
		Trial 5	.01	.10	.01	-.23	.95	-.71
		Immediate recall	-.01	.12	-.01	-.29	1.25	-.97
		Delayed recall	.01	.11	.01	-.28	1.11	-.83
	DS	Forward	.04	.04	-.02	.20	.28	-.45
		Backward	-.01	.02	-.04	.08	.27	-.34
Psychological functioning	Distress	BDI- II	.39	.42	.39	.09	.45	-.13
		GDS	.34	.38	.32	.11	.95	-.69
		STAI	.44	.56	.48	-.25	1.39	-.62
		Obsessions	-.03	-.01	-.09	.80	.09	-.80
		Compulsions	.03	-.28	.06	-.50	.80	-.24
	CAARS	Inattention/memory	.60	.65	.58	.11	.92	-.38
		Hyperactivity/restlessness	.36	.46	.37	-.17	.88	-.30
		Impulsivity/emotional lability	.44	.56	.46	-.14	.90	-.25
		Problems w/self-concept	.44	.53	.45	-.07	1.19	-.62
	UPPS	Positive urgency	.40	.52	.43	-.27	1.02	-.28
		Negative urgency	.21	.27	.25	-.20	.20	.24
		Lack of premeditation	.42	.47	.43	-.05	.59	-.09
		Lack of perseverance	-.03	.09	.01	-.51	.75	-.23
		Sensation seeking	.22	.31	.28	-.45	.58	.14

Note. D-KEFS = Delis-Kaplan Executive Function System; RAVLT = Rey Auditory Verbal Learning Test; DS = digit span; BDI = Beck Depression Inventory; GDS = Geriatric Depression Scale; CAARS = Conners' Adult ADHD Rating Scales; UPPS = Urgency, Premeditation, Perseverance, Sensation Seeking Impulsive Behavior Scale; STAI = State-Trait Anxiety Inventory. See the online article for the color version of this table.

Discussion

This study critically examined previously published factor structures and the overarching construct validity of the CFQ as a self-report measure of cognitive and psychological functioning by leveraging a large, community-based sample of adults. Through multiple statistical techniques, the psychometric evidence of several commonly utilized factor structures was critically analyzed, and a modified single-factor model maintained the best support. Further, self-reported cognitive failures as measured by the CFQ did not reflect performance on objective assessments of cognitive functioning but are reflective of psychological processes and potentially psychopathology.¹ Moreover, subscales of the CFQ were inconsistently related to cognitive task performance, and both the significance and the effect sizes of these relationships were highly volatile across zero-order and residualized models, instilling low confidence in the validity of these solutions. Most critically, the reliance on multidimensional factor structures of the CFQ is likely to result in spurious relationships that obfuscates the construct validity of cognitive failures. These findings demonstrate that subscales derived from these factor structures should not be used in future empirical research or clinical applications.

Despite good model fit, the three-factor (Rast et al., 2009) and four-factor (Wallace et al., 2002) solutions previously published and often implemented in psychological research demonstrated several concerning properties. The correlations between factors in the CFA framework are exceedingly high, suggesting a lack of differentiation and high collinearity. Nuanced investigation of the unique variance provided by each factor of the multidimensional solutions does not support these structures as superior to the single-factor model. Moreover, the use of subscales based on factors failing the Haberman (2008) procedure, as these multidimensional factor structures did, is likely misleading and can result in inaccurate conclusions (Feinberg & Jurich, 2017; Reise et al., 2013). Although the *Names* subscale of the four-factor solution contributed unique variance above and beyond the total score, the factor is comprised of only two indicators. Such a structure is underidentified in a single-factor model, and can contribute to empirical underidentification

¹ Analyses were also conducted on a subset of the sample, aged 60 years or older. Results were consistent with those reported in the article—CFQ scores were not related to objective cognitive functioning as measured by neuropsychological tasks.

Table 5

Comparisons of Standardized Structural Paths (β) From Each CFQ Factor (Dist. = Distractibility; Mem. = Memory; Bldn = Blunders) of the Four-Factor Model (Wallace et al., 2002)

			Four-factor model							
			Zero-order correlations				Residualized associations			
			Dist.	Mem.	Bldn.	Names	Dist.	Mem.	Bldn.	Names
Cognitive functioning	D-KEFS	Trail making	.07	.02	.02	.09	.72	-.38	-.34	-.01
		Card sorting	.19	.02	.06	.14	1.39	-.86	-.37	-.12
		Color/word int.	.09	-.04	.01	.11	.84	-.70	-.13	.01
		Category fluency	.01	-.06	-.07	-.03	.73	-.30	-.38	-.14
		Letter fluency	.08	-.02	.04	.03	.77	-.61	-.05	-.12
	RAVLT	Trial 1	-.01	.05	-.08	-.13	.47	.28	-.63	-.22
		Trial 5	.08	.08	-.01	-.14	.93	-.11	-.48	-.46
		Immediate recall	.06	.11	-.01	-.13	.56	.25	-.57	-.30
		Delayed recall	.07	.10	-.01	-.15	.74	.05	-.48	-.40
	DS	Forward	.04	-.01	-.05	.10	.24	.07	-.42	.15
Backward		.02	.02	-.06	.01	.21	.21	-.46	.05	
Psychological functioning	Distress	BDI- II	.41	.38	.38	.25	.41	.01	.05	-.06
		GDS	.38	.29	.29	.29	1.51	.04	-.97	-.29
		STAI	.52	.51	.47	.24	.78	.25	-.33	-.25
		Obsessions	-.09	.09	-.09	.04	-1.08	1.25	-.35	.24
		Compulsions	.11	.12	.11	-.10	.33	.07	-.06	-.34
	CAARS	Inattention/memory	.64	.60	.56	.40	.69	.21	-.20	-.07
		Hyperactivity/restlessness	.41	.48	.33	.22	.41	.73	-.63	-.11
		Impulsivity/emotional lability	.47	.63	.42	.27	-.08	1.18	-.54	-.02
		Problems w/self-concept	.54	.47	.40	.29	1.26	.13	-.73	-.25
	UPPS	Positive urgency	.43	.58	.41	.24	-.22	1.07	-.33	-.03
		Negative urgency	.20	.30	.26	.11	-.58	.50	.23	.03
		Lack of premeditation	.44	.41	.45	.26	.29	-.04	.26	-.07
		Lack of perseverance	.01	.17	.01	-.09	-.39	.93	-.39	-.12
		Sensation seeking	.21	.39	.33	.05	-.97	.82	.49	-.02

Note. D-KEFS = Delis-Kaplan Executive Function System; RAVLT = Rey Auditory Verbal Learning Test; DS = digit span; BDI = Beck Depression Inventory; GDS = Geriatric Depression Scale; CAARS = Conners' Adult ADHD Rating Scales; UPPS = Urgency, Premeditation, Perseverance, Sensation Seeking Impulsive Behavior Scale; STAI = State-Trait Anxiety Inventory. See the online article for the color version of this table.

(Chen et al., 2001; Rindskopf, 1984) when embedded in a larger model. Additionally, the modified single-factor model includes a residual covariance between the two items comprising the *Names* subscale, appropriately modeling the shared variance due to item content without relying on an underidentified factor.

Taken together, the CFQ appears to be essentially unidimensional (Embretson & Reise, 2000; Reise et al., 2011) and likely represents a single underlying construct rather than substantially meaningful subdimensions. Based solely on the internal structure of the CFQ, the use of a total score is supported, whereas subscale scores are discouraged.

A measure's validity should be determined by carefully considering external relationships that emerge in isolation as well as in the presence of other highly correlated predictors of interest (AERA, APA, & NCME, 2014; Lynam et al., 2006). Discrepancies in relationships across models help to accurately understand the construct validity of a measure (Lynam et al., 2006). The single-factor CFQ latent variable was weakly associated with three executive functioning tasks, suggesting *better* objective cognitive flexibility and letter-guided fluency were associated with more frequent cognitive failures, but was unrelated to verbal learning, recall, attention, and working memory. Likewise, subfactors in isolation were weakly and inconsistently associated with set-shifting, learning, and recall, but were largely

unrelated to other neuropsychological tasks. Disconcerting changes occurred when transitioning to correlated-factor models, casting further doubt on the validity of these subscales as reflective of cognitive abilities. In both the three-factor and four-factor solutions, factors with nonsignificant effects in isolation became significant and strong predictors in combined models. Such inconsistencies serve as a clear indicator of the spurious effects likely to arise with continued use of CFQ subscales. We believe these results help to explain the continued ambiguity in elucidating what the CFQ truly measures (Carrigan & Barkus, 2016; Smilek et al., 2010), and align with clinical research suggesting subjective cognitive failures may not be predictive of objective impairment (Balash et al., 2013; Jorm et al., 2001). Further research leveraging the single-factor model of the CFQ in neurologic and neuropsychiatric populations with known cognitive deficits or psychopathology will be instrumental in further resolving this ambiguity.

Lastly, the external validity of the CFQ was explored by relating the single-factor latent variable to several measures of internalization and externalization. Consistent with past work from the stress vulnerability perspective of cognitive failures (Brück et al., 2019; Matthews et al., 1990; Wagle et al., 1999; Weintraub et al., 2018) and subjective cognitive decline more broadly

(Hill et al., 2016), CFQ total scores were strongly associated with depressive and anxious symptoms, and obsessive-compulsive symptoms to a lesser extent. Although CFQ scores were not significantly related to task-based executive functioning performance, strong and positive relationships emerged with self-reported ADHD symptoms and facets of impulsivity. This combination of findings is similar to research revealing that individuals with ADHD report more subjective difficulties, despite similar performance on working memory tasks (Gu et al., 2018). Consequently, we believe the CFQ may serve as a measure of psychological distress specific to cognitively oriented experiences, but that does not map onto objective deficits expected in neurocognitive disorders.

An important consideration is the contrast between task-based and self-reported executive dysfunction (Toplak et al., 2013). The lack of convergence across methodologies may reflect that task-based and self-reported measures of cognition are assessing fundamentally different constructs (Toplak et al., 2013), or that neuropsychological assessments are not ecologically valid representations of cognitive difficulties (Chaytor & Schmitter-Edgecombe, 2003). However, lab-based neuropsychological assessments demonstrate moderate correlations with daily life deficits as documented by patient relatives or caregivers (Burgess et al., 1998). Subjective declines in cognition have been associated with cognitive decline—particularly in memory domains—though conclusions should be tempered given the high variability across studies in the presence, size, and direction of associations (Burmester et al., 2016). Although we concur with the sentiment that subjective experiences are likely clinically relevant experiences worthy of consideration, the results of this study clearly demonstrate that the CFQ does not offer a convenient proxy for task-based assessments of cognitive functioning and that the CFQ should not be assumed to represent cognitive domains as traditionally conceptualized. Moreover, uniformity in the conceptualization and measurement of perceived cognitive difficulties, including the CFQ, is necessary to elucidate the clinical utility of subjective cognitive decline (Molinuevo et al., 2017).

There are limitations of note in the present study, which help to contextualize our findings. This study relies on a normative sample of adults, primarily without psychiatric or neurologic conditions. Cognitive failures may become more pronounced and impairing in clinical populations, and the relationship between these subjective experiences and more objective deficits may manifest or change. Although we did not observe this effect in our sample of generally healthy older adults (see Footnote 1), patient populations may experience cognitive failures differently. Measurement invariance techniques can help to elucidate these potential deviations in construct validity of the CFQ in future research. Additionally, a handful of studies have indicated subjective complaints predate the onset of a neurological condition (St John & Montgomery, 2002; Wang et al., 2004). Longitudinal validation efforts are necessary to determine if the CFQ provides clinical utility in predicting neurodegenerative processes. Despite these limitations, our normative sample does allow us to conclude that in general, subjective cognitive complaints measured via the CFQ cannot serve as a proxy for cognitive domains traditionally assessed via neuropsychological tasks in research contexts.

In the initial psychometric analysis of the CFQ, Broadbent et al. (1982) perceptively noted that “it should be emphasized there is no evidence for separate categories of perceptual, memory, and action failures” (p. 6). Our study not only echoes that sentiment, demonstrating that the CFQ is a predominantly unidimensional construct,

but also highlights the danger of ignoring this advice in favor of CFQ subscales which give the impression of cognitive domain specificity. Continued use of these subscales serves to obscure the construct validity of subjective, self-reported cognitive failures and the potential implications of these perceived mental slips in clinical psychology, neuropsychology, and neuroscience. Our study demonstrates that the CFQ does not serve as a convenient proxy for objective measures of cognitive deficits, but rather relates to symptoms of psychological distress and subjective symptoms of inattention, hyperactivity, and impulsivity. Although the CFQ may continue to be a useful measure of how individuals perceive impairment in their daily lives, it should be noted that this perception is unlikely to map onto more objective deficits typically assessed in clinical settings.

References

- Adler, L. A., Faraone, S. V., Spencer, T. J., Michelson, D., Reimherr, F. W., Glatt, S. J., Marchant, B. K., & Biederman, J. (2008). The reliability and validity of self- and investigator ratings of ADHD in adults. *Journal of Attention Disorders, 11*(6), 711–719. <https://doi.org/10.1177/1087054707308503>
- American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME]. (2014). *The standards for educational and psychological testing*.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- Ashendorf, L., Jefferson, A. L., O'Connor, M. K., Chaisson, C., Green, R. C., & Stern, R. A. (2008). Trail making test errors in normal aging, mild cognitive impairment, and dementia. *Archives of Clinical Neuropsychology, 23*, 129–137. <https://doi.org/10.1016/j.acn.2007.11.005>
- Balash, Y., Mordechovich, M., Shabtai, H., Giladi, N., Gurevich, T., & Korczyn, A. D. (2013). Subjective memory complaints in elders: Depression, anxiety, or cognitive decline? *Acta Neurologica Scandinavica, 127*(5), 344–350. <https://doi.org/10.1111/ane.12038>
- Balota, D. A., Tse, C.-S., Hutchison, K. A., Spieler, D. H., Duchek, J. M., & Morris, J. C. (2010). Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in stroop color naming. *Psychology and Aging, 25*(1), 208–218. <https://doi.org/10.1037/a0017474>
- Balsamo, M., Romanelli, R., Innamorati, M., Ciccarese, G., Carlucci, L., & Saggino, A. (2013). The state-trait anxiety inventory: Shadows and lights on its construct validity. *Journal of Psychopathology and Behavioral Assessment, 35*(4), 475–486. <https://doi.org/10.1007/s10862-013-9354-5>
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the beck depression inventory-II*. Psychological Corporation.
- Bridger, R. S., Johnsen, S. Å. K., & Brasher, K. (2013). Psychometric properties of the cognitive failures questionnaire. *Ergonomics, 56*(10), 1515–1524. <https://doi.org/10.1080/00140139.2013.821172>
- Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*(1), 1–16. <https://doi.org/10.1111/j.2044-8260.1982.tb01421.x>
- Brück, E., Larsson, J. W., Lasselin, J., Bottai, M., Hirvikoski, T., Sundman, E., Eberhardson, M., Sackey, P., & Olofsson, P. S. (2019). Lack of clinically relevant correlation between subjective and objective cognitive function in ICU survivors: A prospective 12-month follow-up study. *Critical Care (London, England), 23*(1), Article 253. <https://doi.org/10.1186/s13054-019-2527-1>
- Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the*

- International Neuropsychological Society*, 4(6), 547–558. <https://doi.org/10.1017/S1355617798466037>
- Burmester, B., Leathem, J., & Merrick, P. (2016). Subjective cognitive complaints and objective cognitive function in aging: A systematic review and meta-analysis of recent cross-sectional findings. *Neuropsychology Review*, 26(4), 376–393. <https://doi.org/10.1007/s11065-016-9332-2>
- Carrigan, N., & Barkus, E. (2016). A systematic review of cognitive failures in daily life: Healthy populations. *Neuroscience and Biobehavioral Reviews*, 63, 29–42. <https://doi.org/10.1016/j.neubiorev.2016.01.010>
- Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, 13(4), 181–197. <https://doi.org/10.1023/B:NERV.0000009483.91468.fb>
- Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods & Research*, 29(4), 468–508. <https://doi.org/10.1177/0049124101029004003>
- Clément, F., Belleville, S., & Gauthier, S. (2008). Cognitive complaint in mild cognitive impairment and Alzheimer's disease. *Journal of the International Neuropsychological Society*, 14(2), 222–232. <https://doi.org/10.1017/S1355617708080260>
- Conners, C. K., Erhardt, D., Epstein, J. N., Parker, J. D. A., Sitarenios, G., & Sparrow, E. (1999). Self-ratings of ADHD symptoms in adults I: Factor structure and normative data. *Journal of Attention Disorders*, 3(3), 141–151. <https://doi.org/10.1177/108705479900300303>
- Wallace, J. C. (2004). Confirmatory factor analysis of the cognitive failures questionnaire: Evidence for dimensionality and construct validity. *Personality and Individual Differences*, 37(2), 307–324. <https://doi.org/10.1016/j.paid.2003.09.005>
- Davison, M. L., Davenport, E. C., Jr., Chang, Y.-F., Vue, K., & Su, S. (2015). Criterion-related validity: Assessing the value of subscores. *Journal of Educational Measurement*, 52(3), 263–279. <https://doi.org/10.1111/jedm.12081>
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis–Kaplan executive function system: Examiner's manual*. The Psychological.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Psychology Press.
- Feinberg, R. A., & Jurich, D. P. (2017). Guidelines for interpreting and reporting subscores. *Educational Measurement: Issues and Practice*, 36(1), 5–13. <https://doi.org/10.1111/emip.12142>
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. (SCID-I/P)*. Biometrics Research, New York State Psychiatric Institute.
- Forster, S., & Lavie, N. (2007). High perceptual load makes everybody equal: Eliminating individual differences in distractibility with load. *Psychological Science*, 18(5), 377–381. <https://doi.org/10.1111/j.1467-9280.2007.01908.x>
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., & Charney, D. S. (1989). The yale-brown obsessive compulsive scale. II. Validity. *Archives of General Psychiatry*, 46(11), 1012–1016. <https://doi.org/10.1001/archpsyc.1989.01810110054008>
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The yale-brown obsessive compulsive scale. I. Development, use, and reliability. *Archives of General Psychiatry*, 46(11), 1006–1011. <https://doi.org/10.1001/archpsyc.1989.01810110048007>
- Gu, C., Liu, Z.-X., Tannock, R., & Woltering, S. (2018). Neural processing of working memory in adults with ADHD in a visuospatial change detection task with distractors. *PeerJ*, 6, Article e5601. <https://doi.org/10.7717/peerj.5601>
- Haberman, S. J. (2008). When can subscores have value? *Journal of Educational and Behavioral Statistics*, 33(2), 204–229. <https://doi.org/10.3102/1076998607302636>
- Hart, T., Whyte, J., Kim, J., & Vaccaro, M. (2005). Executive function and self-awareness of “real-world” behavior and attention deficits following traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, 20(4), 333–347. <https://doi.org/10.1097/00001199-200507000-00005>
- Hill, N. L., Mogle, J., Wion, R., Munoz, E., DePasquale, N., Yevchak, A. M., & Parisi, J. M. (2016). Subjective cognitive impairment and affective symptoms: A systematic review. *The Gerontologist*, 56(6), e109–e127. <https://doi.org/10.1093/geront/gnw091>
- Hohman, T. J., Beason-Held, L. L., Lamar, M., & Resnick, S. M. (2011). Subjective cognitive complaints and longitudinal changes in memory and brain function. *Neuropsychology*, 25(1), 125–130. <https://doi.org/10.1037/a0020859>
- Homack, S., Lee, D., & Riccio, C. A. (2005). Test review: Delis–Kaplan executive function system. *Journal of Clinical and Experimental Neuropsychology*, 27(5), 599–609. <https://doi.org/10.1080/13803390490918444>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jessen, F., Amariglio, R. E., Buckley, R. F., van der Flier, W. M., Han, Y., Molinuevo, J. L., Rabin, L., Rentz, D. M., Rodriguez-Gomez, O., Saykin, A. J., & Sikkes, S. A. (2020). The characterisation of subjective cognitive decline. *Lancet Neurology*, 19(3), 271–278. [https://doi.org/10.1016/S1474-4422\(19\)30368-0](https://doi.org/10.1016/S1474-4422(19)30368-0)
- Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Henderson, A. S. (2001). Memory complaints as a precursor of memory impairment in older people: A longitudinal analysis over 7–8 years. *Psychological Medicine*, 31(3), 441–449. <https://doi.org/10.1017/S0033291701003245>
- Kanai, R., Dong, M. Y., Bahrami, B., & Rees, G. (2011). Distractibility in daily life is reflected in the structure and function of human parietal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(18), 6620–6626. <https://doi.org/10.1523/JNEUROSCI.5864-10.2011>
- Karr, J. E., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2019). Examining the latent structure of the Delis–Kaplan executive function system. *Archives of Clinical Neuropsychology*, 34(3), 381–394. <https://doi.org/10.1093/arclin/acy043>
- Kessler, R. C., Green, J. G., Adler, L. A., Barkley, R. A., Chatterji, S., Faraone, S. V., Finkelstein, M., Greenhill, L. L., Gruber, M. J., Jewell, M., & Russo, L. J. (2010). Structure and diagnosis of adult attention-deficit/hyperactivity disorder: Analysis of expanded symptom criteria from the Adult ADHD Clinical Diagnostic Scale. *Archives of General Psychiatry*, 67(11), 1168–1178. <https://doi.org/10.1001/archgenpsychiatry.2010.146>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Publications.
- Kolenikov, S., & Bollen, K. A. (2012). Testing negative error variances: Is a heywood case a symptom of misspecification? *Sociological Methods & Research*, 41(1), 124–167. <https://doi.org/10.1177/0049124112442138>
- Kørner, A., Lauritzen, L., Abelskov, K., Gulmann, N., Marie Brodersen, A., Wedervang-Jensen, T., & Marie Kjeldgaard, K. (2006). The geriatric depression scale and the cornell scale for depression in dementia. A validity study. *Nordic Journal of Psychiatry*, 60(5), 360–364. <https://doi.org/10.1080/08039480600937066>
- Larson, G. E., Alderton, D. L., Neideffer, M., & Underhill, E. (1997). Further evidence on dimensionality and correlates of the cognitive failures questionnaire. *British Journal of Psychology*, 88(1), 29–38. <https://doi.org/10.1111/j.2044-8295.1997.tb02618.x>
- Lynam, D. R., Hoyle, R. H., & Newman, J. P. (2006). The perils of partialling: Cautionary tales from aggression and psychopathy. *Assessment*, 13(3), 328–341. <https://doi.org/10.1177/1073191106290562>
- Mahoney, A. M., Dalby, J. T., & King, M. C. (1998). Cognitive failures and stress. *Psychological Reports*, 82(Suppl. 3), 1432–1434. <https://doi.org/10.2466/pr0.1998.82.3c.1432>

- Matthews, G., Coyle, K., & Craig, A. (1990). Multiple factors of cognitive failure and their relationships with stress vulnerability. *Journal of Psychopathology and Behavioral Assessment*, 12(1), 49–65. <https://doi.org/10.1007/BF00960453>
- Mitrushina, M., Satz, P., Chervinsky, A., & D'Elia, L. (1991). Performance of four age groups of normal elderly on the Rey auditory-verbal learning test. *Journal of Clinical Psychology*, 47(3), 351–357. [https://doi.org/10.1002/1097-4679\(199105\)47:3<351::AID-JCLP2270470305>3.0.CO;2-S](https://doi.org/10.1002/1097-4679(199105)47:3<351::AID-JCLP2270470305>3.0.CO;2-S)
- Molinuevo, J. L., Rabin, L. A., Amariglio, R., Buckley, R., Dubois, B., Ellis, K. A., Ewers, M., Hampel, H., Klöppel, S., Rami, L., Reisberg, B., & the Subjective Cognitive Decline Initiative (SCD-I) Working Group. (2017). Implementation of subjective cognitive decline criteria in research studies. *Alzheimer's & Dementia*, 13(3), 296–311. <https://doi.org/10.1016/j.jalz.2016.09.012>
- Nooner, K. B., Colcombe, S. J., Tobe, R. H., Mennes, M., Benedict, M. M., Moreno, A. L., Panek, L. J., Brown, S., Zavitz, S. T., Li, Q., Sikka, S., Gutman, D., Bangaru, S., Schlachter, R. T., Kamiel, S. M., Anwar, A. R., Hinz, C. M., Kaplan, M. S., Rachlin, A. B., . . . Milham, M. P. (2012). The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry. *Frontiers in Neuroscience*, 6, Article 152. <https://doi.org/10.3389/fnins.2012.00152>
- Poliakoff, E., & Smith-Spark, J. H. (2008). Everyday cognitive failures and memory problems in Parkinson's patients without dementia. *Brain and Cognition*, 67(3), 340–350. <https://doi.org/10.1016/j.bandc.2008.02.004>
- Pollina, L. K., Greene, A. L., Tunick, R. H., & Puckett, J. M. (1992). Dimensions of everyday memory in young adulthood. *British Journal of Psychology*, 83(3), 305–321. <https://doi.org/10.1111/j.2044-8295.1992.tb02443.x>
- Rast, P., Zimprich, D., Van Boxtel, M., & Jolles, J. (2009). Factor structure and measurement invariance of the cognitive failures questionnaire across the adult life span. *Assessment*, 16(2), 145–158. <https://doi.org/10.1177/1073191108324440>
- Redoblado, M. A., Grayson, S. J., & Miller, L. A. (2003). Lateralized-temporal-lobe-lesion effects on learning and memory: Examining the contributions of stimulus novelty and presentation mode. *Journal of Clinical and Experimental Neuropsychology*, 25(1), 36–48. <https://doi.org/10.1076/jcen.25.1.36.13625>
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2), 129–140. <https://doi.org/10.1080/00223891.2012.725437>
- Reise, S. P., Ventura, J., Keefe, R. S., Baade, L. E., Gold, J. M., Green, M. F., Kern, R. S., Mesholam-Gately, R., Nuechterlein, K. H., Seidman, L. J., & Bilder, R. (2011). Bifactor and item response theory analyses of interviewer report scales of cognitive impairment in schizophrenia. *Psychological Assessment*, 23(1), 245–261. <https://doi.org/10.1037/a0021501>
- Rindskopf, D. (1984). Structural equation models: Empirical identification, heywood cases, and related problems. *Sociological Methods & Research*, 13(1), 109–119. <https://doi.org/10.1177/0049124184013001004>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Schmidt, M. (1996). *Rey auditory and verbal learning test: A handbook*. Western Psychological.
- Shakeel, M. K., & Goghari, V. M. (2017). Measuring fluid intelligence in healthy older adults. *Journal of Aging Research*, 2017, Article 8514582. <https://doi.org/10.1155/2017/8514582>
- Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Failures of sustained attention in life, lab, and brain: Ecological validity of the SART. *Neuropsychologia*, 48(9), 2564–2570. <https://doi.org/10.1016/j.neuropsychologia.2010.05.002>
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-trait anxiety inventory*. Consulting Psychologists Press.
- St John, P., & Montgomery, P. (2002). Are cognitively intact seniors with subjective memory loss more likely to develop dementia? *International Journal of Geriatric Psychiatry*, 17(9), 814–820. <https://doi.org/10.1002/gps.559>
- Tierney, M. C., Nores, A., Snow, W. G., Fisher, R. H., Zoritto, M. L., & Reid, D. W. (1994). Use of the Rey auditory verbal learning test in differentiating normal aging from Alzheimer's and Parkinson's dementia. *Psychological Assessment*, 6(2), 129–134. <https://doi.org/10.1037/1040-3590.6.2.129>
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 54(2), 131–143. <https://doi.org/10.1111/jcpp.12001>
- van der Werf-Eldering, M. J., Burger, H., Jabben, N., Holthausen, E. A. E., Aleman, A., & Nolen, W. A. (2011). Is the lack of association between cognitive complaints and objective cognitive functioning in patients with bipolar disorder moderated by depressive symptoms? *Journal of Affective Disorders*, 130(1–2), 306–311. <https://doi.org/10.1016/j.jad.2010.10.005>
- Wagle, A. C., Berrios, G. E., & Ho, L. (1999). The cognitive failures questionnaire in psychiatry. *Comprehensive Psychiatry*, 40(6), 478–484. [https://doi.org/10.1016/S0010-440X\(99\)90093-7](https://doi.org/10.1016/S0010-440X(99)90093-7)
- Wallace, J. C., Kass, S. J., & Stanny, C. J. (2002). The cognitive failures questionnaire revisited: Dimensions and correlates. *The Journal of General Psychology*, 129(3), 238–256. <https://doi.org/10.1080/00221300209602098>
- Wang, L., van Belle, G., Crane, P. K., Kukull, W. A., Bowen, J. D., McCormick, W. C., & Larson, E. B. (2004). Subjective memory deterioration and future dementia in people aged 65 and older. *Journal of the American Geriatrics Society*, 52(12), 2045–2051. <https://doi.org/10.1111/j.1532-5415.2004.52568.x>
- Wang, Y.-P., & Gorenstein, C. (2013). Psychometric properties of the beck depression inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria (Sao Paulo, Brazil)*, 35(4), 416–431. <https://doi.org/10.1590/1516-4446-2012-1048>
- Wechsler, D. (2008). *Wechsler adult intelligence scale* (4th ed.). Pearson.
- Wechsler, D. (2011). *Wechsler abbreviated scale of intelligence* (2nd ed.). Pearson.
- Weintraub, M. J., Brown, C. A., & Timpano, K. R. (2018). The relationship between schizotypal traits and hoarding symptoms: An examination of symptom specificity and the role of perceived cognitive failures. *Journal of Affective Disorders*, 237, 10–17. <https://doi.org/10.1016/j.jad.2018.04.121>
- Whiteside, S. P., & Lynam, D. R. (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, 30(4), 669–689. [https://doi.org/10.1016/S0191-8869\(00\)00064-7](https://doi.org/10.1016/S0191-8869(00)00064-7)
- Whiteside, S. P., Lynam, D. R., Miller, J. D., & Reynolds, S. K. (2005). Validation of the UPPS impulsive behaviour scale: A four-factor model of impulsivity. *European Journal of Personality*, 19(7), 559–574. <https://doi.org/10.1002/per.556>
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982–1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37–49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)

Received July 9, 2021

Revision received February 2, 2022

Accepted February 4, 2022 ■