Original Research Article

Assessment
1–16

DOI: 10.1177/10731911251319933
journals.sagepub.com/home/asm
Sage

Assessing Conspiracist Ideation Reliably, Validly, and Efficiently: A Psychometric Comparison of Five Short-Form Measures

Cameron S. Kay[1] and Paul Slovic[2,3]

Abstract

Choosing a short-form measure of conspiracist ideation (i.e., the tendency to believe in conspiracy theories) is fraught. Despite there being numerous scales to choose from, little work has been done to compare their psychometric properties. To address this shortcoming, we compared the internal consistency, 2-week test–retest reliability, criterion validity, and construct validity of five short-form conspiracist ideation measures: the Generic Conspiracist Beliefs Scale–5 (GCB-5), the Conspiracy Mentality Questionnaire (CMQ), the General Measure of Conspiracism (GMC), the American Conspiracy Thinking Scale (ACTS), and the One-Item Conspiracy Measure (1CM). The results of our investigation indicated that all five scales are reliable and valid measures of conspiracist ideation. That said, the GCB-5 tended to perform the best, while the 1CM tended to perform the worst. We conclude our investigation by discussing trade-offs among the five scales, as well as providing recommendations for future research.

Keywords

conspiracy theories, brief measures, psychometrics, internal consistency, test–retest reliability, criterion validity, construct validity

Introduction

Over the past decade, there has been increasing recognition of the negative consequences of conspiracist ideation (i.e., the tendency to believe in conspiracy theories; see Brotherton et al., 2013; Goreis & Voracek, 2019; Imhoff et al., 2022; Uscinski, 2020; see also Nera, 2024; Sutton et al., 2024). Research has suggested that those who believe in conspiracy theories are less likely to vote (Butler et al., 1995), vaccinate (Jolley & Douglas, 2014), and engage in behaviors intended to combat climate change (Van der Linden, 2015), as well as being more likely to harbor prejudicial beliefs (Kay, 2024a; Sapountzis & Condor, 2013) and hold favorable attitudes toward the use of nuclear weapons (Imhoff & Bruder, 2014; Kay & Slovic, 2023). To further investigate the consequences of conspiracist beliefs (and, ultimately, develop interventions to combat these beliefs), it is crucial for researchers to have measures that are psychometrically sound.

Unfortunately, little work has been done to compare the psychometric properties of existing conspiracist ideation measures, making it exceptionally difficult for researchers to make informed decisions about which measures to use in their studies. The present project is intended to address this limitation by providing a psychometric comparison of five short-form conspiracist ideation measures: the Generic Conspiracist Beliefs Scale–5 (GCB-5; Kay & Slovic, 2023), the Conspiracy Mentality Questionnaire (CMQ; Bruder et al., 2013), the General Measure of Conspiracism (GMC; Drinkwater et al., 2012), the American Conspiracy Thinking Scale (ACTS; Uscinski & Parent, 2014), and the One-Item Conspiracy Measure (1CM; Lantian et al., 2016).[1]

Background

Before turning to our investigation, it is important to describe what we mean by a measure's "psychometric

[1]Union College, Schenectady, NY, USA
[2]University of Oregon, Eugene, OR, USA
[3]Decision Research, Springfield, OR, USA

Corresponding Author:
Cameron S. Kay, Psychology Department, Union College, 807 Union Street, Schenectady, NY 12308, USA.
Email: cameronstuartkay@gmail.com

properties." Here, we use the term to refer to three qualities of a scale: its reliability, validity, and efficiency.

The reliability of a scale refers to whether it produces consistent measurements (Cronbach, 1947; John & Soto, 2007; Revelle & Condon, 2019). Here, we consider two approaches to reliability. The first is the internal consistency approach, which assesses the consistency of responses to a scale at a single time point. For better or worse, Cronbach's alpha is the prototypical measure of internal consistency (John & Soto, 2007). It represents the average correlation between all split halves of a scale, corrected to full test length (Cronbach, 1951; Cortina, 1993). The second approach is the test–retest approach, which involves administering a measure across multiple time points. The prototypical measure for the test–retest approach is the simple correlation of scores from a first administration of a measure with scores from a second administration of the measure. When it comes to assessing dependability (i.e., a lack of transient measurement error) rather than stability (i.e., a true change in an underlying construct), two weeks is the typical interval used between both administrations (Chmielewski & Watson, 2009; Nunnally & Bernstein, 1994). It is ostensibly long enough to minimize memory effects while being short enough to not allow true changes in trait-like constructs (such as conspiracist ideation) to occur.

The validity of a scale concerns whether it produces accurate measurements (Cronbach, 1990). As with reliability, we consider two forms of validity here. The first is criterion validity, which concerns whether a measure is associated with a theoretically relevant outcome (Allen & Yen, 1979). In the case of conspiracist ideation, the natural theoretically relevant outcome is the belief in specific conspiracy theories, such as the belief that the earth is flat or that the Apollo moon landings were faked. If a measure of conspiracist ideation is able to predict the belief in specific conspiracy theories, it is evidence for the criterion validity of the scale.[2] The second form of validity that we are going to consider is construct validity, which concerns whether a measure manifests in associations that are consistent with the underlying construct (Cronbach & Meehl, 1955). As but one example, a measure of conspiracist ideation should, theoretically, be associated with paranoia, given that conspiracist ideation and paranoia both involve feelings of distrust and a lack of autonomy (Imhoff & Lamberty, 2018). Finding a positive association between a measure of conspiracist ideation and a measure of paranoia is, therefore, evidence for the construct validity of the scale.

Compared to reliability and validity, efficiency is rarely discussed, but it is an important psychometric property to consider all the same. The efficiency of a scale refers to its length or, more specifically, its brevity. As has been discussed elsewhere (Kay & Slovic, 2023), efficient measures have several advantages over inefficient measures. For one, efficient measures can save researchers time. All else being equal, the shorter a survey is, the more participants who will be willing to complete it (Galesic & Bosnjak, 2009). Consequently, if a researcher is using convenience sampling, shorter, more efficient scales will reduce the time it takes the researcher to achieve their desired sample size. This is also the case when a researcher uses a human subjects pool, since a researcher typically only has a limited number of research credits to award to participants, and the research credit awarded to participants is typically tied to the length of a survey. Efficient measures can also save researchers money. More often than not, the amount of money a participant is awarded for completing a survey on a data-collection platform (e.g., Prolific; CloudResearch Connect) is directly tied to the length of the survey. As such, decreasing the length of a survey by using measures that are more efficient can reduce the total amount of money researchers spend on participant payments. Furthermore, efficient measures can improve data quality. As the length of a survey increases, so does the likelihood that a participant will lapse into careless responding (e.g., Bowling et al., 2021). The use of efficient measures can, therefore, decrease the incidence of careless responding and, by extension, improve data quality (e.g., Cornell et al., 2012; Credé, 2010; Kay, 2024b; Schmitt & Stults, 1985). Finally, there are ethical reasons to use efficient measures. Namely, if the same information can be ascertained using a shorter scale, the shorter scale should be used because it reduces the burden imposed on participants.

The five short-form conspiracist ideation measures of interest in the present study are all remarkably efficient. The GCB-5 (Kay & Slovic, 2023), CMQ (Bruder et al., 2013), and GMC (Drinkwater et al., 2012) are each five items long, taking participants about 35 seconds to complete. The ACTS (Uscinski & Parent, 2014) is four items long, taking participants about 28 seconds to complete. Finally, the 1CM (Lantian et al., 2016) is only one item long, taking participants about 7 seconds to complete. If a researcher is only concerned about maximizing efficiency, the 1CM is clearly the best choice. However, researchers are often not only concerned about a measure's efficiency. Often, they are also concerned about a measure's reliability and validity, and the evidence for the reliability and validity of the five short-form conspiracist ideation measures considered here is varied, perhaps owing to their different origins.

The GCB-5 (Kay & Slovic, 2023; Table 1) has its origins in the GCB-15, a highly popular 15-item measure of conspiracist ideation (Brotherton et al., 2013). To create the GCB-15, Brotherton and colleagues started by conducting a factor analysis of 75 different generic

Table 1. The Items From the Generic Conspiracist Beliefs Scale–5, Conspiracy Mentality Questionnaire, General Measure of Conspiracism, American Conspiracy Thinking Scale, and One-Item Conspiracy Measure.

Item
Generic Conspiracist Beliefs Scale–5 (GCB-5)
The government permits or perpetrates acts of terrorism on its own soil, disguising its involvement.
Evidence of alien contact is being concealed from the public.
New and advanced technology which would harm current industry is being suppressed.
Certain significant events have been the result of the activity of a small group who secretly manipulate world events.
Experiments involving new drugs or technologies are routinely carried out on the public without their knowledge or consent.
Conspiracy Mentality Questionnaire (CMQ)
Many very important things happen in the world, which the public is never informed about.
Politicians usually do not tell us the true motives for their decisions.
Government agencies closely monitor all citizens.
Events which superficially seem to lack a connection are often the result of secret activities.
There are secret organizations that greatly influence political decisions.
General Measure of Conspiracism (GMC)
Conspiracy theories accurately depict real life events.
The information contained within conspiracy theories is generally true.
When I hear conspiracy theories, I feel they are untrue.
Conspiracy theories contain information, which has proved to be false.
I have heard several conspiracy theories, which I believe to be true.
American Conspiracy Thinking Scale (ACTS)
Even though we live in a democracy, a few people will always run things anyway.
The people who really "run" the country are not known to the voters.
Big events like wars, the recent recession, and the outcomes of elections are controlled by small groups of people who are working in secret against the rest of us.
Much of our lives are being controlled by plots hatched in secret places.
One-Item Conspiracy Measure (1CM)
I think that the official version of the events given by the authorities very often hides the truth.

conspiracist beliefs (e.g., "The government has employed people in secret to assassinate others"). The results of their factor analysis indicated that there are five themes that underlie conspiracist beliefs. Namely, there are beliefs that (a) the government is engaged in wrongdoing, (b) the public is being lied to about the existence of aliens, (c) small malevolent groups of people influence global events, (d) plots and schemes threaten the public's well-being and liberty, and (e) important information is being suppressed, manipulated, or otherwise controlled by powerful actors. The GCB-15 was created by writing three items to assess each of these five themes. Kay and Slovic, in turn, created the GCB-5 by extracting the highest loading item from each theme. The purpose of selecting the highest loading item was to ensure that the GCB-5 would still retain the conceptual breadth of the GCB-15, despite being substantially shorter.

Although it was only introduced recently, a fair amount of evidence has been collected in favor of the GCB-5's reliability and validity. Across five studies, Kay and Slovic (2023) found support for the GCB-5's internal consistency (αs = .71–.80); criterion validity, as evidenced by its sizable associations with 21 specific conspiracy theories (*r*s = .59–.68); and construct validity, as evidenced by its sizable associations with constructs such as delusional ideation (*r*s = .46–.52), paranoia (r = .38), and anomie (*r*s = .30–.39). Kay and Slovic also found that informants were more likely to label a person a "conspiracy theorist" if they scored high on the GCB-5 (r = .23), providing additional evidence for the scale's criterion validity. A follow-up study by Dagnall and colleagues (2023) found further evidence for the GCB-5's internal consistency, assessed via omega (ω = .83), and construct validity, as evidenced by its sizable associations with a composite of eight specific conspiracy theories (r = .72). As of yet, no work has examined the test–retest reliability of the GCB-5, although prior work indicates that the GCB-15 has high test–retest reliability (e.g., Brotherton et al., 2013; Majima & Nakamura, 2020; Siwiak et al., 2019).

The CMQ (Bruder et al., 2013; Table 1) was developed as an efficient measure of conspiracy mentality, defined as the propensity to explain important societal phenomena as being the result of conspiracies among malevolent individuals or groups. Its exact provenance is a bit unclear (M. Bruder, personal communication, July 8, 2024), but it seems to have been created by combining three items from a scale assessing suspicious thought patterns (Sjöberg, 2005) with two novel items written based on the content from websites about

conspiracy theories. The five items were originally combined with 33 specific conspiracy theories to form the Conspiracy Theory Questionnaire (Bruder & Manstead, 2009; see also Darwin et al., 2011), but it was later separated to form its own scale.

The CMQ is the most popular short-form measure of conspiracist ideation to date and, as a result, there are many sources of evidence for its reliability and validity. Some of this evidence is incidental, coming from researchers examining other empirical questions. For example, researchers have recently shown that the CMQ is associated with both greater credulity (r = .20; Brauner et al., 2023) and greater use of one's intuition to determine what is true (r = .29; Abels & Lewandowsky, 2024). Other evidence has come by way of dedicated psychometric investigations. For instance, across four studies, Bruder and colleagues (2013) found support for the CMQ's internal consistency (αs = .72–.85); test–retest reliability (r_{TwoWeek} = .84); criterion validity, as evidenced by its large associations with 33 specific conspiracy theories (rs = .20–.81); and construct validity, as evidenced by its sizable associations with constructs like paranoia (r = .45), schizotypy (rs = .18–.36), right-wing authoritarianism (r = .28), and anomie (rs = .22). Likewise, subsequent research has found similar results for the CMQ's criterion validity and construct validity, demonstrating that it is associated with belief in specific 9/11 conspiracy theories (r = .53) and anti-vaccination beliefs (r = .33; Swami et al., 2017). That said, Kay and Slovic (2023) have found that people who score high on the CMQ are no more likely to be labeled "conspiracy theorists" by informants (r = .07). They also found that the GCB-5 demonstrated larger associations than the CMQ with both specific conspiracy theories (rs = .59–.68 versus rs = .35–.55) and odd beliefs (r = .46 versus r = .24).

The GMC (Drinkwater et al., 2012; Table 1) was created ad hoc for a study investigating the association between conspiracist ideation and the critical evaluation of one's sensory experiences. Of the five measures considered here, it is, by far, the most face-valid. Each of the five items from the GMC asks participants to report, in one way or another, how believable they find conspiracy theories (e.g., "I have heard several conspiracy theories, which I believe to be true"). Of course, this approach may seem flawed. Given the term's negative connotations, would anyone really be willing to admit to finding truth in "conspiracy theories?" Interestingly, the answer appears to be "yes." Describing a theory as a "conspiracy theory" seems to do little to dissuade people from endorsing the theory (Wood, 2016; but see Dentith et al., 2023), suggesting that there is little direct harm in using this language. However, there is also a second potential flaw that is not so easily addressed. The definition of what is and is not a conspiracy theory presumably varies from person to person, which can lead to the items being interpreted in different ways, interfering with the scale's ability to provide meaningful comparisons across groups.

This potential generalizability issue aside, the evidence for the GMC's reliability and validity is promising. Drinkwater and colleagues (2012) found that the GMC has respectable internal consistency (α = .72); criterion validity, as evidenced by its large negative association with endorsing official explanations for historical events (r = −.52) and large positive association with endorsing alternative explanations for historical events (r = .42); and construct validity, as evidenced by its sizable associations with constructs like paranormal beliefs (r = .31), urban legends (r = .31), and new-age philosophy (r = .28). To our knowledge, no work has been conducted to investigate the test–retest reliability of the GMC.

The ACTS (Uscinski & Parent, 2014; Table 1) was created ad hoc to investigate the sociopolitical correlates of conspiracist beliefs. The researchers wanted to develop a measure that was able to assess what they considered to be at the heart of conspiratorial thinking: the belief that powerful groups of people secretly manipulate events to undermine the common good. To create this measure, they selected three items from McClosky and Chong's (1985) research into radicalization, with one of the items originally coming from the California Fascism Scale (Sanford et al., 1950), and two of the items originally coming from a scale assessing distrust of the government (McClosky, 1964). An item that had originally been used to validate the scale was later added as a fourth item (Uscinski et al., 2016).

As with the GMC, the ACTS did not go through a formal validation process, but the evidence that is available for its reliability and validity is promising. Recent work from Uscinski and colleagues (Enders et al., 2023; Uscinski et al., 2022) has indicated that the ACTS has good internal consistency (αs = .84–.86); criterion validity, as evidenced by its sizable associations with 39 specific conspiracy theories ($\bar{r}$ = .35); and construct validity, as evidenced by its sizable associations with constructs like anomie (r = .40), distrust of the government (r = .22), and the willingness to share false information online (r = .30). Moreover, Han and colleagues (2022) found that the scale is associated with anti-expert sentiment (r = .45), as well as distrust of a wide range of groups, including the police (r = .40), the World Health Organization (r = .37), and scientists (r = .40), providing further support for the scale's construct validity. To our knowledge, no research has examined the test–retest reliability of the ACTS.

Finally, the 1CM (Lantian et al., 2016; Table 1) was created as part of an ambitious (and, by all accounts,

successful) effort to develop a one-item measure of conspiracist ideation. Unlike the GCB-5, the 1CM was not created by extracting items or, in this case, an item from a longer scale. Lantian and colleagues were concerned that, if they drew the item from an existing scale, the 1CM would only be able to assess one aspect of conspiracist ideation. So, instead, Lantian and colleagues wrote a novel item assessing the belief that authorities often mislead the public, which they noted is a central feature of many conspiracist beliefs.

Prior research has found good evidence for the 1CM's reliability and validity. Since it only contains one item, researchers cannot investigate the scale's internal consistency, but, across three studies, Lantian and colleagues (2016) found support for the 1CM's test–retest reliability (r = .75); criterion validity, as evidenced by its large positive correlations with a set of specific conspiracy theories (rs = .50–.66) and willingness to sign up to receive a newsletter about conspiracy theories (r = .16); and construct validity, as evidenced by its sizable negative association with interpersonal trust (r = −.27). Subsequent research has found additional support for the 1CM's criterion and construct validity, demonstrating that it is associated with belief in specific 9/11 conspiracy theories (r = .40) and anti-vaccination beliefs (r = .36; Swami et al., 2017).

Taken together, the prior studies indicate that the five short-form conspiracist ideation measures are reliable and valid. However, these investigations largely considered the scales in isolation (but see Kay & Slovic, 2023; Swami et al., 2017). Unfortunately, this means the findings from these investigations cannot be used to directly compare the reliability and validity of the measures, since any observed differences could simply be due to the fact that the studies used different methodologies. The purpose of the present study is to address this limitation by simultaneously evaluating the five scales.

Current Study

In the present study, we evaluate the reliability of the short-form conspiracist ideation measures in two ways. First, we evaluate their internal consistencies by producing a Cronbach's alpha for each scale. We hypothesize that all of the alphas will be greater than the traditional cutoff threshold of .70 (Nunnally, 1978; but see also Lance et al., 2006). We also hypothesize that all of the alphas will be comparable in size. Second, we evaluate the test–retest reliabilities of the scales by correlating scores from an initial administration of the scales with scores from an administration of the same scales approximately 2 weeks later. We hypothesize that all of the scales will evince strong positive correlations. We also hypothesize that all of the correlations will be comparable in size.

As with reliability, we evaluate the validity of the short-form conspiracist ideation measures in two ways. First, we evaluate their criterion validities by correlating the scales with a set of 21 specific conspiracy theories. We hypothesize that each scale will evince a strong positive correlation. Given the GCB-5 showed a larger association than the CMQ with a set of specific conspiracy theories in a prior study (Kay & Slovic, 2023), we also hypothesize that the GCB-5 will demonstrate a larger association with the set of specific conspiracy theories than the other measures of conspiracist ideation in the present study. Second, we evaluate the scales' construct validities by producing correlations of the scales' scores with scores from measures of six theoretically relevant constructs, including paranoia (Imhoff & Lamberty, 2018), the tendency to entertain odd beliefs (Barron et al., 2018; Dagnall et al., 2015; Darwin et al., 2011; Furnham & Grover, 2021; Swami et al., 2011, 2016; Van der Tempel & Alcock, 2015), anomie (Abalakina-Paap et al., 1999; Goertzel, 1994), a desire for chaos (Farhart et al., 2023), a denial of expert information (Uscinski & Klofstad, 2024), and illusory pattern perception (Van Prooijen et al., 2018). We hypothesize that the scales will demonstrate moderate-to-large positive associations with the six construct validity measures. We also hypothesize that the associations will be comparable across the five conspiracist ideation measures.

All of these hypotheses are pre-registered[3] (https://osf.io/mzaup/?view_only=697f04a0a10d4337abbcdd6df69a1316), as are all of the methods described in the following section. The materials, data, and analytic code for the present study are also provided on OSF (https://osf.io/uzrgk/?view_only=aca403a5146240b-da740e1e6d640 751f).

Method

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

Participants and Procedures

The data for this study were collected anonymously using two Qualtrics surveys posted to Prolific. The surveys were posted approximately 2 weeks apart. Only those participants who completed the first survey were invited back to complete the second survey. Participants were paid approximately $2.40 for completing the first survey and approximately $ 0.43 for completing the second survey, rates roughly equivalent to $8.00 per hour.

Participants had to be 18 years of age or older and currently living in the United States to participate.

To determine the necessary sample size for the present study, we conducted two power analyses. The first power analysis was used to calculate the necessary sample size to test the correlational hypotheses. The power analysis indicated that 419 participants would be required to detect a moderate correlation ($r = .20$; see Funder & Ozer, 2019; Gignac & Szodorai, 2016) 80% of the time that such an effect existed in the population with an alpha level of .001. We opted for an alpha level of .001 to account for Type 1 error rate inflation resulting from testing multiple associations. The second power analysis was used to calculate the necessary sample size to test the hypotheses that involved *comparing* correlations. The power analysis indicated that 366 participants would be required to detect a difference between a .40 correlation and a .20 correlation (the smallest difference deemed to be of practical interest) 80% of the time that such an effect existed in the population with an alpha level of .001 and a correlation of .50 between the two predictor variables. The rationale for using an alpha level of .001 is provided earlier; the rationale for specifying a .50 correlation between the predictor variables was based on the smallest correlation (rounded down to the nearest tenth) between the GCB-5 and CMQ in a prior study (i.e., Kay & Slovic, 2023). Taking into account these two power analyses, we aimed to collect 500 participants for the first survey. We opted for this larger sample size to account for exclusions and for the potential misspecification of our power analyses. For the second survey, we optimistically aimed to collect 500 participants but, given that we expected some level of attrition, we estimated the actual number of respondents to be somewhere around 450.

Over 500 ($N = 504$) participants responded to the first survey. After excluding participants who failed two or more of the six instructed response items included in the survey ($n = 4$; Curran, 2016; Kay & Saucier, 2023), responded faster than one-third of the median response time ($n = 3$; Bedford-Petersen & Saucier, 2021), provided the same response to over half of the items in the survey in a row ($n = 3$; Johnson, 2005), demonstrated a response standard deviation of less than .50 ($n = 0$; Thalmayer & Saucier, 2014; see also Dunn et al., 2018), or provided an average response greater than zero to the six infrequency/frequency items embedded in the survey ($n = 2$; Kay, 2024c), the sample included 492 participants ($M_{Age} = 43.79$, $SD_{Age} = 14.77$). In line with preset demographic quotas, approximately half of the participants identified as women (48.98%), and approximately half of the participants identified as men (49.80%). Likewise, approximately half of the participants identified as Democrats (46.75%), and approximately half of the participants identified as Republicans (46.54%). Additional demographic information (e.g., the participants' states of residence; the participants' education levels) can be found in the data-processing document on OSF (https://osf.io/uzrgk/?view_only=aca403a5146240bda740e1e6d640751f).

Over 400 ($N = 410$) of the participants who completed the first survey (and were not excluded under our first round of screening) responded to the second survey. After excluding participants who failed both the instructed response items in the second survey ($n = 0$; Curran, 2016; Kay & Saucier, 2023), responded faster than one-third of the median response time ($n = 5$; Bedford-Petersen & Saucier, 2021), provided the same response to over half of the items in the second survey in a row ($n = 5$; Johnson, 2005), demonstrated a response standard deviation of less than .50 ($n = 1$; Thalmayer & Saucier, 2014; see also Dunn et al., 2018), or provided an average response greater than zero to the two infrequency/frequency items embedded in the second survey ($n = 10$; Kay, 2024c), the sample included 389 participants.[4] The retained participants completed the second survey between 8.57 and 27.48 days after the first survey ($M = 15.25$, $SD = 3.44$).

Materials

The Five Short-Form Conspiracist Ideation Measures. The participants completed the GCB-5 (Kay & Slovic, 2023), CMQ (Bruder et al., 2013), GMC (Drinkwater et al., 2012), ACTS (Uscinski & Parent, 2014), and 1CM (Lantian et al., 2016) as part of the first and second surveys. The GCB-5 includes five items ($\alpha = .82$; $\bar{r}_{ij} = .47$); the CMQ includes five items ($\alpha = .83$; $\bar{r}_{ij} = .50$); the GMC includes five items ($\alpha = .89$; $\bar{r}_{ij} = .62$); the ACTS includes four items ($\alpha = .85$; $\bar{r}_{ij} = .58$); and the 1CM includes one item. The items for each measure can be found in Table 1. Participants responded to the measures using a 7-point Likert scale (1 = "strongly disagree"; 7 = "strongly agree").

Criterion Validity Measure. To assess the criterion validity of the five short-form conspiracist ideation measures, the participants completed the Belief in Conspiracy Theories Inventory-21 (BCTI-21; Kay & Slovic, 2023; see also Swami et al., 2011) as part of the first survey. The BCTI-21 includes 21 specific conspiracy theories (e.g., "The Apollo moon landings never happened and were staged in a Hollywood film studio"; $\alpha = .93$; $\bar{r}_{ij} = .40$). The scale was formed by adding six conspiracy theories (e.g., "Some airplanes release chemical/biological agents intended to control the human population; Kay & Slovic, 2023) to the 15 conspiracy theories from the

Table 2. A Comparison of the Internal Consistency, Test–Retest Reliability, Criterion Validity, and Construct Validity of the Generic Conspiracist Beliefs Scale–5 (GCB-5), Conspiracy Mentality Questionnaire (CMQ), General Measure of Conspiracism (GMC), American Conspiracy Thinking Scale (ACTS), and One-Item Conspiracy Measure (1CM).

Psychometric test	GCB-5	CMQ	GMC	ACTS	1CM
Internal consistency					
Cronbach's alpha (α)	$.82_b$	$.83_b$	$.89_a$	$.85_b$	-
Test-retest reliability					
Test-retest correlation (*r*)	$.90*_a$	$.88*_a$	$.86*_a$	$.87*_a$	$.66*_b$
Criterion validity					
BCTI-21	$.78*_a$	$.69*_b$	$.74*_{ab}$	$.73*_{ab}$	.57*c
Construct validity					
Paranoia	$.48*_a$	$.47*_{ab}$	$.38*_{bc}$	$.50*_a$	.36*c
Odd beliefs	$.48*_a$	$.35*_b$	$.40*_{ab}$	$.38*_b$	.31*b
Anomie	$.56*_{ab}$	$.59*_a$	$.40*_c$	$.55*_{ab}$	.50*bc
Need for chaos	$.41*_a$	$.35*_a$	$.33*_a$	$.41*_a$	.31*a
Denial of expert information	$.70*_b$	$.76*_a$	$.58*_c$	$.72*_{ab}$	.71*ab
Illusory pattern perception	$.18*_a$	$.20*_a$	$.14_a$	$.12_a$	.13a

Note. Different subscripted letters in a row indicate the Cronbach's alphas or correlations are significantly different at $p < .001$. A Cronbach's alpha was not produced for the 1CM because it only contains one item.
*$p < .001$.

BCTI-15 (e.g., "Area 51 in Nevada, US, is a secretive military base that contains hidden alien spacecraft and/or alien bodies"; Swami et al., 2011). Participants responded to the conspiracy theories on a 9-point scale (1 = "completely false"; 9 = "completely true").

Construct Validity Measures. To assess the construct validity of the five short-form conspiracist ideation scales, participants completed the Persecution and Deservedness Scale (Melo et al., 2009), the Odd Beliefs subscale from the Schizotypal Personality Questionnaire (Raine, 1991), Agnew's Anomie Scale (Agnew, 1980), the Need for Chaos Scale (Arceneaux et al., 2021), and the Denialism Scale (Uscinski et al., 2020) as part of the first survey. The Persecution and Deservedness Scale is a 10-item measure of paranoia (e.g., "There are times when I worry that others might be plotting against me"; $\alpha = .89$; $\bar{r}_{ij} = .44$). The Odd Beliefs subscale from the Schizotypal Personality Questionnaire is a seven-item measure of a person's tendency to hold odd beliefs (e.g., "I believe in clairvoyancy [psychic forces, fortune telling]"; $\alpha = .87$; $\bar{r}_{ij} = .50$). Agnew's Anomie Scale is an eight-item measure of anomie (e.g., "These days a person really doesn't know who they can trust"; $\alpha = .72$; $\bar{r}_{ij} = .24$). The Need for Chaos Scale is a seven-item measure of how much a person desires chaos (e.g., "I think society should be burned to the ground"; $\alpha = .76$; $\bar{r}_{ij} = .32$). The Denialism Scale is a four-item measure of a person's tendency to deny information provided by experts (e.g., "Major events are not always what they seem"; $\alpha = .76$; $\bar{r}_{ij} = .44$). The participants responded to the scales using a seven-point Likert scale (1 = "strongly disagree"; 7 = "strongly agree").

As an additional test of the construct validity of the five short-form conspiracist ideation measures, participants completed Van Prooijen's Coin Toss Task (Van Prooijen et al., 2018; see also Dagnall et al., 2007) as part of the first survey. The Coin Toss Task assesses a person's tendency to see patterns where none exist (i.e., illusory pattern perception). Participants began by rating the randomness of 10 different sequences of 10 coin flips (e.g., "THHTTHHHHH," with "H" referring to heads and "T" referring to tails). They were then told to imagine the 10 sequences were part of a longer 100-flip sequence and asked to provide an additional rating. The participants responded to the 11 sequences using a 7-point scale (1 = "completely random"; 7 = "completely determined"). Their 11 responses were averaged together to generate a total score ($\alpha = .90$; $\bar{r}_{ij} = .45$). In the present study, the sequences were generated by simulating 100 coin flips using the 'rbinom' function from the '{stats}' package in *R* (R Core Team, 2024).

Results

To account for Type I error rate inflation resulting from the testing of multiple associations, we have used a more conservative alpha level of .001 for all tests reported here.

Internal Consistency Reliability

To evaluate the internal consistency of the four multi-item short-form conspiracist ideation measures, we started by calculating a Cronbach's alpha for each

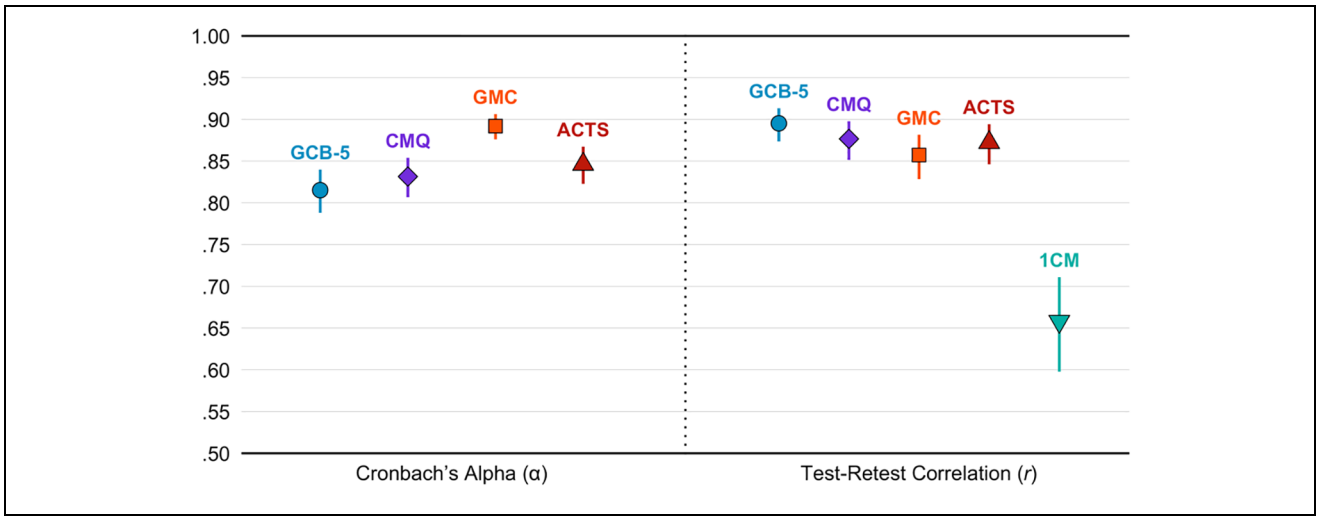

Figure 1. Cronbach's Alphas and Two-Week Test–Retest Correlations With 95% Confidence Intervals for the Generic Conspiracist Beliefs Scale–5 (GCB-5), Conspiracy Mentality Questionnaire (CMQ), General Measure of Conspiracism (GMC), American Conspiracy Thinking Scale (ACTS), and One-Item Conspiracy Measure (1CM).
Note. A Cronbach's alpha was not produced for the 1CM because it only contains one item.

measure. As shown in Table 2 and Figure 1, the Cronbach's alphas for the GCB-5 ($\alpha = .82$), CMQ ($\alpha = .83$), GMC ($\alpha = .89$), and ACTS ($\alpha = .85$) all exceeded .70. We then used the method outlined by Feldt and colleagues (1987) to compare the four Cronbach's alphas. The results indicated that there was a significant difference among the alphas, $\chi^2(3, N = 492) = 64.66$, $p < .001$. Specifically, the GMC had a higher Cronbach's alpha than the other three measures.

Test–Retest Reliability

To evaluate the test–retest reliabilities of the five short-form conspiracist ideation measures, we calculated correlations of the scales' scores from the first survey with the scales' scores from the second survey. As shown in Table 2 and Figure 1, the 2-week test-rest correlations of the GCB-5 ($r = .90$), CMQ ($r = .88$), GMC ($r = .86$), and ACTS ($r = .87$) were all above .70. The 2-week test–retest correlation of the 1CM ($r = .66$) landed slightly below this threshold. Follow-up comparisons using Silver and colleagues' (2004) procedure revealed that the 1CM had a significantly lower 2-week test–retest correlation than the other four measures.

Criterion Validity

To evaluate the criterion validity of the five short-form conspiracist ideation measures, we produced zero-order correlations of the five measures with the BCTI-21. As shown in Table 2, the GCB-5 ($r = .78$), CMQ ($r = .69$), GMC ($r = .74$), ACTS ($r = .73$), and 1CM ($r = .66$) were all highly positively correlated with the BCTI-21. Comparing the associations using Hittner and colleagues' (2003) procedure revealed that the GCB-5 exhibited a larger association with the BCTI-21 than both the CMQ and 1CM. The comparison also revealed that the CMQ, GMC, and ACTS exhibited larger associations with the BCTI-21 than the 1CM. These trends are also apparent in the associations of the five scales with the 21 individual conspiracy theories from the BCTI-21 (Figure 2).

Construct Validity

To evaluate the construct validity of the five short-form conspiracist ideation measures, we produced zero-order correlations of each measure with the measures of paranoia, odd beliefs, anomie, need for chaos, denial of expert information, and illusory pattern perception. As shown in Table 2 and Figure 3, the GCB-5, CMQ, GMC, ACTS, and 1CM were all highly positively correlated with paranoia, odd beliefs, anomie, a desire for chaos, and a denial of expert information. Only the GCB-5 and CMQ were significantly positively correlated with illusory pattern perception, although the correlations were modest. Comparing the associations using Hittner and colleagues' (2003) procedure revealed several notable differences in the associations. For paranoia, the GCB-5 and ACTS exhibited larger associations than the GMC and 1CM. The CMQ also exhibited a larger association than the 1CM. For odd beliefs, the GCB-5 exhibited a larger association than the CMQ, ACTS, and 1CM. For anomie, the GCB-5, CMQ, and ACTS exhibited larger associations

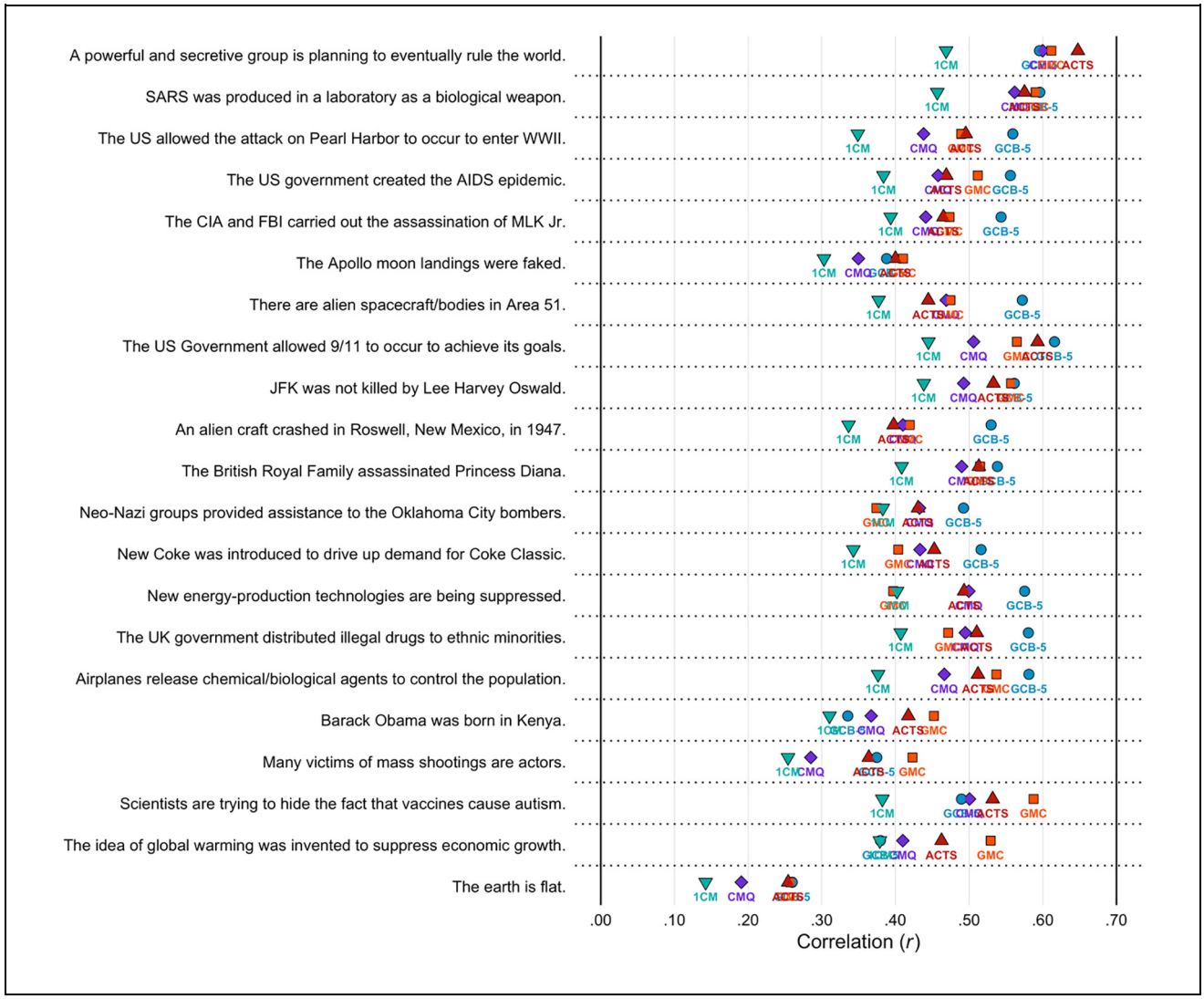

Figure 2. Correlations of the Generic Conspiracist Beliefs Scale–5 (GCB-5), Conspiracy Mentality Questionnaire (CMQ), General Measure of Conspiracism (GMC), American Conspiracy Thinking Scale (ACTS), and One-Item Conspiracy Measure (1CM) With Each of the Conspiracy Theories From the Belief in Conspiracy Theories Inventory–21.

than the GMC. The CMQ also exhibited a larger association than the 1CM. For the denial of expert information, the CMQ exhibited a larger association than the GCB-5 and 1CM. The GCB-5, CMQ, ACTS, and 1CM also all exhibited larger associations than the GMC. For the need for chaos and illusory pattern perception, all of the associations were comparable.

Discussion

Prior investigations into the reliability and validity of the GCB-5 (Kay & Slovic, 2023), CMQ (Bruder et al., 2013), GMC (Drinkwater et al., 2012), ACTS (Uscinski & Parent, 2014), and 1CM (Lantian et al., 2016) have mostly considered the measures individually (but see Kay & Slovic, 2023; Swami et al., 2017). As a result, the prior studies are largely unable to tell us which measures are more reliable or more valid. To provide an answer to these questions, we simultaneously tested the internal consistency, test–retest reliability, criterion validity, and construct validity of the five measures.

Turning to our results, we found evidence that the scales are, in fact, internally consistent. Specifically, as hypothesized, the four multi-item scales (i.e., the GCB-5, CMQ, GMC, and ACTS) all had Cronbach's alphas above the oft-cited threshold of .70 (Nunnally, 1978; but see also Lance et al., 2006). Inconsistent with our hypotheses, however, the Cronbach's alphas for the four

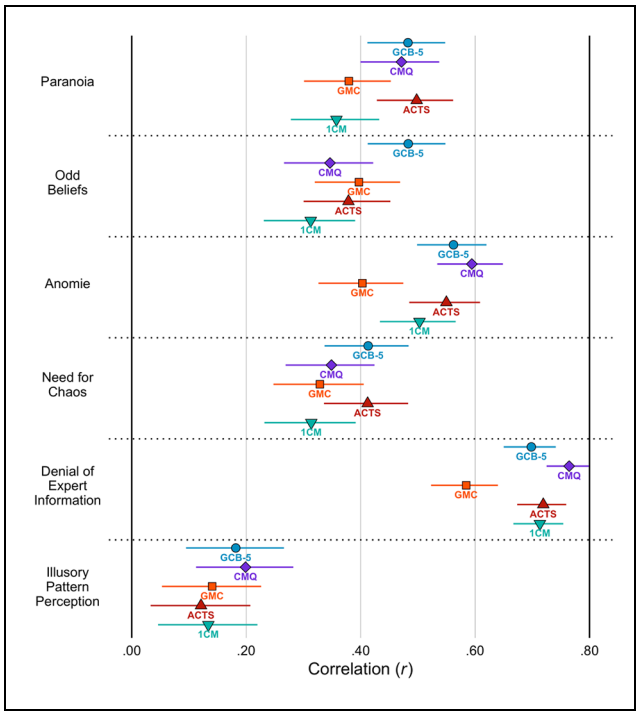

Figure 3. Correlations of the Generic Conspiracist Beliefs Scale–5 (GCB-5), Conspiracy Mentality Questionnaire (CMQ), General Measure of Conspiracism (GMC), American Conspiracy Thinking Scale (ACTS), and One-Item Conspiracy Measure (1CM) With the Six Construct Validity Scales With 95% Confidence Intervals.

scales were significantly different: the Cronbach's alpha for the GMC was larger than the Cronbach's alphas for the GCB-5, CMQ, and ACTS. Although this did not accord with our expectations, it is not particularly surprising. A determinant of Cronbach's alpha is the degree of interrelatedness among a scale's items, and a determinant of the interrelatedness among a scale's items is how similar the items are. Of the four multi-item scales considered here, the items from the GMC are, by far, the most homogeneous. All of the items reference "conspiracy theories" and ask participants to, in one way or another, rate the veracity of theories classified as conspiracy theories. Although this homogeneity likely increased the GMC's reliability, it is also possible that it narrowed the breadth of the content captured by the scale, reducing its criterion validity and construct validity (discussed below).

We also found good evidence for the test–retest reliability of the five scales. As hypothesized, scores from an initial administration of the scales were highly positively correlated with scores from those same scales approximately 2 weeks later. Nevertheless, we also hypothesized that the five scales would have comparable levels of test–retest reliability but instead found that the test–retest reliability of the 1CM was significantly lower than that of the other four scales. In fact, its 2-week test–retest reliability was below the .70 threshold. Although this was unexpected, it, again, is not particularly surprising. Responses to all items include measurement error. However, this error can be reduced by aggregating across multiple items (see Allen & Yen, 1979). Through aggregation, positive errors—those that would cause a participant to score higher than their true score—cancel out negative errors—those that would cause a participant to score lower than their true score. A longer scale should, therefore, have less error and show greater dependability than a shorter scale, especially one that only includes a single item like the 1CM.

The five scales tested here also all appear to be criterion valid. As hypothesized, we found that the five scales were all highly positively correlated with a set of 21 specific conspiracy theories. Given prior work comparing the criterion validity of the GCB-5 and CMQ (Kay & Slovic, 2023), we also hypothesized that the GCB-5 would evince a stronger association with the set of specific conspiracy theories than the other measures. We found mixed support for this hypothesis. The GCB-5 demonstrated the largest association among the five scales, but the association was only significantly greater than the associations seen for the CMQ and 1CM. It was not larger than that seen for the GMC and ACTS. The reason that the GCB-5 demonstrated a larger association than the CMQ and 1CM may be because, even though it is a generic measure of conspiracist beliefs, it is more specific than the CMQ and 1CM. By way of illustration, the GCB-5 includes the item "Evidence of alien contact is being concealed from the public," which is generic compared to the item "Area 51 in Nevada, US, is a secretive military base that contains hidden alien spacecraft and/or alien bodies" from the BCTI-15 but specific compared to the item "Many very important things happen in the world, which the public is never informed about" from the CMQ and the item "I think that the official version of the events given by the authorities very often hides the truth" from the 1CM. The GCB-5 may, therefore, be better than the CMQ and 1CM at predicting agreement with specific conspiracy theories because it is closer to being a measure of specific conspiracy theories than either the CMQ or 1CM. Whatever the case may be, the results of the present study indicate that the criterion validity of the GCB-5 is greater than some (but not all) of the other short-form conspiracist ideation measures considered here.

Finally, we found good evidence that the five scales are construct valid. Specifically, consistent with our hypotheses, we found that the five measures of conspiracist ideation demonstrated moderate-to-large associations with paranoia, odd beliefs, anomie, a desire for

chaos, and the denial of expert information. Inconsistent with our hypotheses, however, only the GCB-5 and CMQ were significantly associated with illusory pattern perception. The GMC, ACTS, and 1CM were not significantly associated with illusory pattern perception. One interpretation of this finding is that these three scales do not accurately capture aspects of conspiracist ideation related to illusory pattern perception. However, the consistency in the effects across the five measures (rs = .12–.20) support a second, potentially more defensible interpretation. Namely, the link between conspiracist ideation and illusory pattern perception may be weaker than suggested by the prior literature (e.g., r = .37; Van Prooijen et al., 2018). We also hypothesized that the five measures would demonstrate comparable associations with the six construct validity measures. The support for this notion was, again, mixed. The associations were comparable for a desire for chaos and illusory pattern perception, but not for paranoia, odd beliefs, anomie, and the denial of expert information. These results indicate that the five short-form conspiracist ideation measures, for the most part, tap theoretically relevant constructs but to varying degrees.

Taken in concert, the aforementioned findings suggest that the five measures tested here are reliable and valid measures of conspiracist ideation. As such, we feel relatively comfortable recommending researchers use whichever of the five measures they prefer. That said, we do have two more specific recommendations.

Our first specific recommendation is to use the GCB-5. Although there was good evidence for the reliability and validity of all the measures tested here, the GCB-5 generally outperformed the other measures. The GCB-5 had a stronger test–retest correlation than the 1CM; a stronger correlation with specific conspiracy theories than the CMQ and 1CM; and a stronger association with paranoia, odd beliefs, anomie, and a denial of expert information than a number of the other measures. Moreover, even in those cases where the GCB-5 did not outperform the other measures, it tended to perform at least comparably. In fact, there were only two cases where the GCB-5 performed worse than the other measures. The first was that the GCB-5 had a smaller Cronbach's alpha than the GMC. As noted earlier, the elevated Cronbach's alpha for the GMC may be due to the similarity among its item. One reaction to this finding is to suggest that the GCB-5 should be updated to be more homogeneous. However, the heterogeneity of the GCB-5's items may actually be why it outperforms some of the other measures in terms of its criterion validity and construct validity. Namely, the GCB-5 was specifically designed to capture the five themes of conspiracist beliefs identified by Brotherton and colleagues (2013), which necessarily made the scale less homogeneous but potentially better able to capture the full breadth of the conspiracist ideation construct. The second case where the GCB-5 underperformed was in its association with the denial of expert information. Specifically, the CMQ showed a larger association with the denial of expert information than the GCB-5. This is potentially due to the fact that a number of the items from the CMQ specifically reference content related to the sharing and processing of information (e.g., "Politicians usually do not tell us the true motives for their decisions"; "Many very important things happen in the world, which the public is never informed about"). Of course, the GCB-5 could be updated to better capture the denial of expert information, but, again, increasing fidelity in this domain runs the risk of decreasing its ability to assess conspiracist ideation across multiple domains.

Our second specific recommendation is to use a measure other than the 1CM unless maximizing efficiency is a top priority. The 1CM generally performed worse than the other measures. It had a weaker test–retest correlation than the other measures; a weaker correlation with specific conspiracy theories than the other measures; and a weaker association with paranoia, odd beliefs, and anomie than a number of other measures. In fact, it only demonstrated a larger association than another measure once. The 1CM demonstrated a larger association with the denial of expert information than the GMC. Similar to the CMQ, this is presumably because the one item from the 1CM essentially assesses whether a person denies expert information (i.e., "I think that the official version of the events given by the authorities very often hides the truth."). However, despite its relatively poor reliability and validity, the 1CM is the most efficient of the five measures discussed here. It can, therefore, be used in exceptionally time- and resource-constrained situations. If a researcher is able to administer more than one item, we recommend using a different measure, but, if the choice is between administering the 1CM and not assessing conspiracist ideation at all, we fully endorse using the 1CM.

Limitations and Future Directions

The present study had a number of limitations that are worth noting. First, we only considered five short-form conspiracist ideation measures. To our knowledge, this is the largest psychometric comparison of conspiracist ideation measures to date, but there are other measures that could have been considered (e.g., the Conspiracy Mentality Scale; Stojanov & Halberstadt, 2019). The present results should, therefore, not be taken to mean that the measures considered here would perform better (or worse) than all other short-form conspiracist ideation measures nor

that the measures considered here should be used to the exclusion of all other short-form conspiracist ideation measures. We fully encourage future work considering the psychometric properties of other short-form conspiracist ideation measures.

Second, we used Cronbach's alpha as our measure of internal consistency in the present study. We used Cronbach's alpha because, for better or worse, it is the most popular measure of internal consistency. Nevertheless, we appreciate that it is a flawed index (McNeish, 2018) and encourage researchers to consider additional indices in future work.

Third, we only assessed the 2-week test–retest reliability of the scales. When it comes to assessing a scale's dependability, 2 weeks is something of a gold standard (Nunnally & Bernstein, 1994), but, as noted by Revelle and Condon (2019), using multiple retest intervals can provide additional insight. We encourage researchers to consider retest intervals of varying lengths in future work.

Fourth, the set of specific conspiracy theories used to assess criterion validity in the present study represents only a small subset of all of the possible conspiracy theories that could have been used. We did use a fairly large set of conspiracy theories, but it is possible that using a different set would have yielded different results. As a case in point, it is possible that the GCB-5 was particularly well-suited to predict the specific conspiracy theories considered in the present study because, in addition to capturing variance related to believing in conspiracy theories, it also captures rogue variance (see Saucier & Iurino, 2020) related to these specific conspiracy theories. Namely, given how it was constructed, the GCB-5 may include rogue variance related to Western conspiracy theories, which may have improved its ability to predict the largely Western set of specific conspiracy theories considered here. Critically, this could mean that the GCB-5 would generate worse predictions if used to predict conspiracy theories that do not share this rogue variance. We, therefore, encourage future work considering the criterion validity of the short-form conspiracist ideation measures included here in relation to a broader and preferably novel array of conspiracy theories, such as Swami and colleagues' (2011) fictitious Red Bull conspiracy theories.

Fifth (and relatedly), we only considered a subset of all of the possible measures that could have been used to assess the construct validity of the five measures. The measures we used here are all theoretically related to (and empirically associated with) conspiracist ideation, but, again, the results could have differed if a different set of measures was used. We, therefore, encourage future work considering other theoretically relevant constructs, including, for example, trust (e.g., Wagner-Egger & Bangerter, 2007).

Finally, our investigation only included a single sample, which was drawn from a Western, Educated, Industrialized, Rich, and Democratic society (Henrich et al., 2010) and composed mostly of people who identified as white (69.51%). It is, therefore, unclear whether the results we observed in the present study would generalize to other samples. We encourage future work to examine the psychometric properties of these measures across different temporal and spatial contexts.

Conclusion

When it comes to assessing conspiracist ideation, researchers have a wealth of measures to choose from. Unfortunately, relatively little work has been done to compare the psychometric properties of these measures. The present study addressed this limitation by comparing five short-form conspiracist ideation measures. All of the measures performed well, but the GCB-5 tended to perform the best, while the 1CM tended to perform the worst. We believe these findings can help researchers make informed decisions about which measures they include in their future studies.

Acknowledgments

The authors thank Dr. Martin Bruder for answering our questions about the development of the CMQ.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The authors of the present manuscript developed one of the tools evaluated here. There is currently no way for the authors to profit financially from this tool, and they do not intend to ever change this arrangement.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding from the Union College Faculty Research Fund is gratefully acknowledged.

Ethics Approval

The study reported here was determined to be exempt by the Human Subjects Review Committee at Union College (E23033). It was conducted in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Consent to Participate

Participants provided informed consent at the beginning of the surveys reported here.

Consent for Publication

Not applicable.

Data Availability

The data for the present study are provided at https://osf.io/uzrgk/?view_only = aca403a5146240bda740e1 e6d640751f.

Open Practices

The pre-registration for the present study is provided at https://osf.io/mzaup/?view_only = 697f04a0a10d4337abbcdd6 df69a1316. The materials, data, and analytic code are provided at https://osf.io/uzrgk/?view_only = aca403a5146240bda740e1 e6d640751f.

ORCID iDs

Cameron S. Kay https://orcid.org/0000-0002-5210-427X
Paul Slovic https://orcid.org/0000-0002-7473-6403

Notes

1. The five short-form measures of conspiracist ideation considered here were selected based on two main criteria. First, they are all generic measures of conspiracist ideation (see Goreis & Voracek, 2019). In other words, instead of assessing whether a person believes in specific conspiracy theories, they aim to assess the mind-set, worldview, or overarching belief system that leads a person to believe in specific conspiracy theories. Generic measures have become increasingly popular among conspiracy theory researchers because, by divorcing the scales from specific people, places, and events, they are, at least in theory, better able to assess conspiracist beliefs across different temporal and spatial contexts. Second, the measures considered here are all, at most, five items long. This threshold was chosen largely based on the fact that the most popular short-form measure of conspiracist ideation—the CMQ (Bruder et al., 2013)—is five items long. Researchers who typically use the CMQ could, therefore, use any of the measures considered here without having to worry about expending additional research funds.
2. Importantly, specific conspiracy theories are criteria for *measures* of conspiracist ideation. They are not criteria for the *individual items* from measures of conspiracist ideation. The individual items that comprise measures of conspiracist ideation are often generalized versions of specific conspiracy theories, which would make specific conspiracy theories not so much *criteria* of the individual items as much as *narrower manifestations* of the individual items. That said, even when averaged together, the individual items can contribute rogue variance (see Saucier & Iurino, 2020) to a measure's score that is not directly related to the tendency to believe in conspiracy theories, with the amount of this variance being tied to how generalized the items are to begin with. For example, if all of the generalized items contributing to a measure of conspiracist ideation reference extraterrestrials, the resulting measure of conspiracist ideation is going to be better able to predict the tendency for people to believe in conspiracy theories about extraterrestrials. This is not because the measure is a more criterion-valid measure of the tendency to believe in conspiracy theories generally but because the measure is a more criterion-valid measure of the tendency to believe in conspiracy theories about extraterrestrials specifically. This is all to say that, although specific conspiracy theories are likely the *best* criteria for testing the criterion validity of measures of conspiracist ideation, they are also imperfect.
3. As per our pre-registration, we also collected data on 20 items similar to those used by Slovic and colleagues (Kay & Slovic, 2023; Slovic et al., 2020) to assess so-called "virtuous violence" (see Fiske & Rai, 2014; see also Slovic et al., 2020). The goal of collecting these data was to examine the association between conspiracist ideation and virtuous violence in an exploratory fashion. Given these results are not directly relevant to the present manuscript, we have not reported them here.
4. Notably, we ended up collecting fewer responses to our second survey (389) than was suggested by the first power analysis (419). However, this should not be an issue. The responses to the second survey will only be used to evaluate the test–retest reliability hypotheses, which specify large effects (.30). Only 182 participants are required to detect a large effect 80% of the time that such an effect exists in the population with an alpha level of .001.

References

Abalakina-Paap, M., Stephan, W. G., Craig, T., & Gregory, W. L. (1999). Beliefs in conspiracies. *Political Psychology*, *20*(3), 637–647.

Abels, C. M., & Lewandowsky, S. (2024). *Development and validation of the evidence intuition scale*. PsyArXiv Preprint, 1–38.

Agnew, R. S. (1980). Success and anomie: A study of the effect of goals on anomie. *Sociological Quarterly*, *21*(1), 53–64.

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Waveland Press.

Arceneaux, K., Gravelle, T. B., Osmundsen, M., Petersen, M. B., Reifler, J., & Scotto, T. J. (2021). Some people just want to watch the world burn: The prevalence, psychology, and politics of the "Need for Chaos." *Philosophical Transactions of the Royal Society B*, *376*, 1–9.

Barron, D., Furnham, A., Weis, L., Morgan, K. D., Towell, T., & Swami, V. (2018). The relationship between schizotypal facets and conspiracist beliefs via cognitive processes. *Psychiatry Research*, *259*, 15–20.

Bedford-Petersen, C., & Saucier, G. (2021). Identifying contrasting themes that orchestrate personality expression across situations. *Personality and Individual Differences*, *171*, 110495.

Bowling, N. A., Gibson, A. M., Houpt, J. W., & Brower, C. K. (2021). Will the questions ever end? Person-level increases in careless responding during questionnaire completion. *Organizational Research Methods*, *24*(4), 718–738.

Brauner, F., Fonagy, P., Campbell, C., Griem, J., Storck, T., & Nolte, T. (2023). "Trust me, do not trust anyone": How epistemic mistrust and credulity are associated with conspiracy mentality. *Research in Psychotherapy: Psychopathology, Process and Outcome*, *26*(3), 55–67.

Brotherton, R., French, C. C., & Pickering, A. D. (2013). Measuring belief in conspiracy theories: The Generic Conspiracist Beliefs scale. *Frontiers in Psychology*, *4*, 1–15.

Bruder, M., Haffke, P., Neave, N., Nouripanah, N., & Imhoff, R. (2013). Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy Mentality Questionnaire. *Frontiers in Psychology*, *4*(225), 11–15.

Bruder, M., & Manstead, A. (2009). *Questionnaire on conspiracy theories*. Cardiff University. http://www.conspiracytheory.martinbruder.com/en/

Butler, L. D., Koopman, C., & Zimbardo, P. G. (1995). The psychological impact of viewing the film "JFK": Emotions, beliefs, and political behavioral intentions. *Political Psychology*, *16*(2), 237–257.

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

Cornell, D., Klein, J., Konold, T., & Huang, F. (2012). Effects of validity screening items on adolescent survey data. *Psychological Assessment*, *24*(1), 21–35.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104.

Credé, M. (2010). Random responding as a threat to the validity of effect size estimates in correlational research. *Educational and Psychological Measurement*, *70*(4), 596–612.

Cronbach, L. J. (1947). Test "reliability": Its meaning and determination. *Psychometrika*, *12*(1), 1–16.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Cronbach, L. J. (1990). How to judge tests: Validation. In L. J. Cronbach (Ed.), *Essentials of psychological testing* (5th ed., pp. 144–189). Harper Collins.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302.

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19.

Dagnall, N., Denovan, A., Drinkwater, K. G., & Escolà-Gascón, A. (2023). The Generic Conspiracist Beliefs Scale-5: Further psychometric evaluation using a United Kingdom-based sample. *Frontiers in Psychology*, *14*, 1–8.

Dagnall, N., Drinkwater, K., Parker, A., Denovan, A., & Parton, M. (2015). Conspiracy theory and cognitive style: A worldview. *Frontiers in Psychology*, *6*, 1–9.

Dagnall, N., Parker, A., & Munley, G. (2007). Paranormal belief and reasoning. *Personality and Individual Differences*, *43*, 1406–1415.

Darwin, H., Neave, N., & Holmes, J. (2011). Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Personality and Individual Differences*, *50*(8), 1289–1293.

Dentith, M. R. X., Husting, G., & Orr, M. (2023). Does the phrase "conspiracy theory" matter? *Society*, *61*, 189–196.

Drinkwater, K., Dagnall, N., & Parker, A. (2012). Reality testing, conspiracy theories, and paranormal beliefs. *Journal of Parapsychology*, *76*(1), 57–77.

Dunn, A. M., Heggestad, E. D., Shanock, L. R., & Theilgard, N. (2018). Intra-individual response variability as an indicator of insufficient effort responding: Comparison to other indicators and relationships with individual differences. *Journal of Business and Psychology*, *33*(1), 105–121.

Enders, A. M., Diekman, A., Klofstad, C., Murthi, M., Verdear, D., Wuchty, S., & Uscinski, J. (2023). On modeling the correlates of conspiracy thinking. *Scientific Reports*, *13*, 8325.

Farhart, C. E., Fitz, E. B., Miller, J. M., & Saunders, K. L. (2023). By any memes necessary: Belief- and chaos-driven motives for sharing conspiracy theories on social media. *Research and Politics*, *10*(3), 1–8.

Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, *11*(1), 93–103.

Fiske, A. P., & Rai, T. S. (2014). *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge University Press.

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168.

Furnham, A., & Grover, S. (2021). Do you have to be mad to believe in conspiracy theories? Personality disorders and conspiracy theories. *International Journal of Social Psychiatry*, *68*(7), 1454–1461.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78.

Goertzel, T. (1994). Belief in conspiracy theories. *Political Psychology*, *15*(4), 731–742.

Goreis, A., & Voracek, M. (2019). A systematic review and meta-analysis of psychological research on conspiracy beliefs: Field characteristics, measurement instruments, and associations with personality traits. *Frontiers in Psychology*, *10*, 1–13.

Han, H., Blackburn, A. M., Jeftić, A., Tran, T. P., Stöckli, S., Reifler, J., & Vestergren, S. (2022). Validity testing of the conspiratorial thinking and anti-expert sentiment scales during the COVID-19 pandemic across 24 languages from a large-scale global dataset. *Epidemiology and Infection*, *150*, e167.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–135.

Hittner, J. B., May, K., & Silver, N. C. (2003). A monte carlo evaluation of tests for comparing dependent correlations. *Journal of General Psychology*, *130*(2), 149–168.

Imhoff, R., Bertlich, T., & Frenken, M. (2022). Tearing apart the "evil" twins: A general conspiracy mentality is not the

same as specific conspiracy beliefs. *Current Opinion in Psychology*, *46*, 101349.

Imhoff, R., & Bruder, M. (2014). Speaking (Un-)truth to power: Conspiracy mentality as a generalised political attitude. *European Journal of Personality*, *28*(1), 25–43.

Imhoff, R., & Lamberty, P. (2018). How paranoid are conspiracy believers? Toward a more fine-grained understanding of the connect and disconnect between paranoia and belief in conspiracy theories. *European Journal of Social Psychology*, *48*(7), 909–926.

John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461–494). The Guilford Press.

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from Web-based personality inventories. *Journal of Research in Personality*, *39*, 103–129.

Jolley, D., & Douglas, K. M. (2014). The effects of anti-vaccine conspiracy theories on vaccination intentions. *PLOS ONE*, *9*(2), e89177.

Kay, C. S. (2024a). *Conspiracist ideation is associated with generalized prejudice: An investigation with three measures of conspiracist ideation and twenty measures of prejudice*. PsyArXiv Preprint, 1–38.

Kay, C. S. (2024b). Why you shouldn't trust data collected on MTurk. PsyArXiv Preprint, 1–23.

Kay, C. S. (2024c). Validating the IDRIS and IDRIA: Two infrequency/frequency scales for detecting careless and insufficient effort survey responders. *Behavior Research Methods*, *56*, 7790–7813.

Kay, C. S., & Saucier, G. (2023). The Comprehensive Infrequency/Frequency Item Repository (CIFR): An online database of items for detecting careless/insufficient-effort responders in survey data. *Personality and Individual Differences*, *205*, 112073.

Kay, C. S., & Slovic, P. (2023). The Generic Conspiracist Beliefs Scale - 5: A short-form measure of conspiracist ideation. *Journal of Research in Personality*, *102*, 104315.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, *9*(2), 202–220.

Lantian, A., Muller, D., Nurra, C., & Douglas, K. M. (2016). Measuring belief in conspiracy theories: Validation of a French and English Single-Item Scale. *International Review of Social Psychology*, *29*(1), 1–14.

Majima, Y., & Nakamura, H. (2020). Development of the Japanese version of the Generic Conspiracist Beliefs Scale (GCBS-J). *Japanese Psychological Research*, *62*(4), 254–267.

McClosky, H. (1964). Consensus and ideology in American Politics. *American Political Science Review*, *58*(2), 361–382.

McClosky, H., & Chong, D. (1985). Similarities and differences between left-wing and right-wing radicals. *British Journal of Political Science*, *15*(3), 329–363.

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433.

Melo, S., Corcoran, R., Shryane, N., & Bentall, R. P. (2009). The persecution and deservedness scale. *Psychology and Psychotherapy: Theory, Research, and Practice*, *82*, 247–260.

Nera, K. (2024). Analyzing the causation between conspiracy mentality and belief in conspiracy theories: Potential pitfalls and leads to address them. *Zeitschrift Für Psychologie*, *232*(1), 44–49.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Raine, A. (1991). The SPQ: A scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophrenia Bulletin*, *17*(4), 555–564.

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment*, *31*(12), 1395–1411.

Sanford, R. N., Adorno, T. W., Frenkel-Brunswik, E., & Levinson, D. J. (1950). The measurement of implicit antidemocratic trends. In T. W. Adorno, E. Frenkel-Brunswik, D. J. Levinson & R. N. Sanford (Eds.), *The authoritarian personality* (pp. 222–279). Harper & Row.

Sapountzis, A., & Condor, S. (2013). Conspiracy accounts as intergroup theories: Challenging dominant understandings of social power and political legitimacy. *Political Psychology*, *34*(5), 731–752.

Saucier, G., & Iurino, K. (2020). High-dimensionality personality structure in the natural language: Further analyses of classic sets of English-language trait-adjectives. *Journal of Personality and Social Psychology*, *199*(5), 1188–1219.

Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, *9*(4), 367–373.

Silver, N. C., Hittner, J. B., & May, K. (2004). Testing dependent correlations with nonoverlapping variables: A Monte Carlo simulation. *Journal of Experimental Education*, *73*(1), 53–69.

Siwiak, A., Szpitalak, M., & Polczyk, R. (2019). Generic Conspiracist Beliefs Scale—Polish adaptation of the method. *Polish Psychological Bulletin*, *50*(3), 259–269.

Sjöberg, L. (2005). The perceived risk of terrorism. *Risk Management*, *7*(1), 43–61.

Slovic, P., Mertz, C. K., Markowitz, D. M., Quist, A., & Västfjäll, D. (2020). Virtuous violence from the war room to death row. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(34), 1–9.

Stojanov, A., & Halberstadt, J. (2019). The Conspiracy Mentality Scale: Distinguishing between irrational and rational suspicion. *Social Psychology*, *50*(4), 215–232.

Sutton, R. M., Douglas, K. M., & Trella, C. (2024). Conspiracy mentality versus belief in conspiracy theories: Response to Nera and some recommendations for researchers. *Zeitschrift Fur Psychologie / Journal of Psychology*, *232*(1), 50–54.

Swami, V., Barron, D., Weis, L., & Voracek, M. (2017). An examination of the factorial and convergent validity of four measures of conspiracist ideation, with recommendations for researchers. *PLOS ONE*, *12*(2), 1–27.

Swami, V., Coles, R., Stieger, S., Pietschnig, J., Furnham, A., Rehim, S., & Voracek, M. (2011). Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories. *British Journal of Psychology*, *102*(3), 443–463.

Swami, V., Weis, L., Lay, A., Barron, D., & Furnham, A. (2016). Associations between belief in conspiracy theories and the maladaptive personality traits of the personality inventory for DSM-5. *Psychiatry Research*, *236*, 86–90.

Thalmayer, A. G., & Saucier, G. (2014). The questionnaire Big Six in 26 nations: Developing cross-culturally applicable Big Six, Big Five and Big Two inventories. *European Journal of Personality*, *28*, 482–496.

Uscinski, J. E. (2020). *Conspiracy theories: A primer*. Rowman & Littlefield.

Uscinski, J. E., Enders, A. M., Diekman, A., Funchion, J., Klofstad, C., Kuebler, S., Murthi, M., Premaratne, K., Seelig, M., Verdear, D., & Wuchty, S. (2022). The psychological and political correlates of conspiracy theory beliefs. *Scientific Reports*, *12*, 21672.

Uscinski, J. E., Enders, A. M., Klofstad, C. A., Seelig, M., Funchion, J., Everett, C., Wuchty, S., Premaratne, K., & Murthi, M. (2020, April 28). Why do people believe COVID-19 conspiracy theories? *Harvard Kennedy School Misinformation Review*, pp. 1–24.

Uscinski, J. E., Klofstad, C., & Atkinson, M. D. (2016). What drives conspiratorial beliefs? The role of informational cues and predispositions. *Political Research Quarterly*, *69*(1), 57–71.

Uscinski, J. E., & Klofstad, C. A. (2024). Denialism or conspiracism? *The causes and consequences of rejecting official accounts [Unpublished manuscript]*.

Uscinski, J. E., & Parent, J. M. (2014). *American conspiracy theories*. Oxford University Press.

Van der Linden, S. L. (2015). The conspiracy-effect: Exposure to conspiracy theories (about global warming) decreases pro-social behavior and science acceptance. *Personality and Individual Differences*, *87*, 171–173.

Van der Tempel, J., & Alcock, J. E. (2015). Relationships between conspiracy mentality, hyperactive agency detection, and schizotypy: Supernatural forces at work? *Personality and Individual Differences*, *82*, 136–141.

Van Prooijen, J. W., Douglas, K. M., & De Inocencio, C. (2018). Connecting the dots: Illusory pattern perception predicts belief in conspiracies and the supernatural. *European Journal of Social Psychology*, *48*(3), 320–335.

Wagner-Egger, P., & Bangerter, A. (2007). La vérité est ailleurs: Corrélats de l'adhésion aux théories du complot. Revue Internationale de Psychologie Sociale, 20(4), 31–61.

Wood, M. J. (2016). Some dare call it conspiracy: Labeling something a conspiracy theory does not reduce belief in it. *Political Psychology*, *37*(5), 695–705.