Journal of the Academy of Marketing Science (2021) 49:221–235
https://doi.org/10.1007/s11747-020-00766-8

METHODOLOGICAL PAPER

The biasing effect of common method variance: some clarifications

Hans Baumgartner[1] · Bert Weijters[2] · Rik Pieters[3]

Received: 9 December 2020 / Published online: 6 January 2021

Abstract

There are enduring misconceptions in the marketing and management literature about the potential biasing effects of Common Method Variance (CMV). One belief is that the biasing effect of CMV is of greater theoretical than practical importance; another belief is that if CMV is a potential problem, it can be easily identified with the Harman one-factor test. In this article, we show that both beliefs are ill founded and need correction. To demonstrate our key points with greater generality, we use analytical derivations rather than empirical simulations. First, we examine the effects of CMV on correlations between observed variables as a function of measure unreliability and the sign and size of the "true" trait correlation. We demonstrate that, for negative trait correlations, CMV leads to a substantial upward bias in observed correlations (i.e., observed correlations are less negative than the trait correlation), and under certain conditions observed correlations may even have the wrong sign (assuming that the method loadings are both positive or both negative). We also show that, for positive trait correlations, the downward bias due to measurement unreliability does not always mitigate the upward bias due to CMV (again assuming that the method loadings are either both positive or both negative). Importantly, our results indicate that the inflationary effect of CMV is larger at *lower* levels of (positive) trait correlations, whereas the deflationary effect of unreliability is larger at *higher* levels of trait correlations. Second, we demonstrate analytically the serious deficiencies of the popular Harman one-factor test for detecting common method variance and strongly recommend against its use in future research.

Keywords Common method variance · Common method bias · Harman one-factor test · Systematic error · Unreliability of measurement

The potential bias caused by method variance has been of enduring interest to marketing and management researchers (Baumgartner and Steenkamp 2001; Brannick et al. 2010; Cote and Buckley 1987, 1988; MacKenzie and Podsakoff 2012; Podsakoff et al. 2003, 2012; Richardson et al. 2009; Spector 1987, 2006; Williams and Brown 1994; Williams et al. 1989). Broadly defined, method variance refers to any systematic, non-substantive influence on measures of substantive constructs that is due to the method of measurement used. Although method effects can occur for all kinds of measurement, we follow the literature by limiting our discussion to single and multi-item self-report measures of constructs. Method variance is one source of correlational error, which occurs "when individual responses vary consistently to different degrees over and above true differences in the construct being measured; that is, it is a result of different individuals responding in consistently different ways over and above true differences in the construct" (Viswanathan 2005, p. 108). Common method variance (CMV) is a specific type of method variance pertaining to situations where multiple measures of the same construct or different constructs share the same measurement method. Podsakoff et al. (2003) (see also Baumgartner and Weijters 2019, Section 4.1.2) distinguish three sets of factors that may lead to method effects: (a) characteristics of the respondent providing the self-reports (called common rater effects by Podsakoff et al.), particularly response styles such as

John Hulland served as Editor for this article.

✉ Hans Baumgartner
HansBaumgartner@psu.edu

Bert Weijters
bert.weijters@ugent.be

Rik Pieters
pieters@uvt.nl

[1] Smeal College of Business, Pennsylvania State University, University Park, PA, USA

[2] Department of Work, Organization and Society, Ghent University, Ghent, Belgium

[3] Business and Economics, Tilburg School of Economics and Management, Tilburg University, Tilburg, The Netherlands

acquiescent, extreme, and midpoint responding and response sets such as social desirability (Baumgartner and Steenkamp 2001; Steenkamp et al. 2010); (b) properties of the items used to assess a construct (called item characteristic effects in Podsakoff et al.), such as the keying direction of the items or the response scale used to collect the ratings (Baumgartner et al. 2018; Weijters et al. 2010); and (c) general features of the questionnaire and the questionnaire context influencing the measurements (called item and measurement context effects in Podsakoff et al.), such as the positioning of items in the questionnaire or the mode of data collection (Weijters et al. 2008, 2009).[1] To illustrate, imagine a survey in which individual differences in extraversion are measured by asking respondents to report whether they are "talkative." Any systematic measurement influence on the response to the "talkative" item, other than people's standing on the trait of extraversion, such as their preference for a certain response category ("agree" or "strongly agree") or the position of the question in the questionnaire (e.g., beginning, middle, end), would be a potential method effect. CMV is problematic because it induces dependencies between items independently of the substance of the items (as noted by Viswanathan 2005). This is in contrast to random errors of measurement, which can be assumed to be independent across different measures of the same construct or measures of different constructs. As a consequence of CMV, items designed to measure the same trait (or related traits) share variance not only because the underlying trait is the same (or the traits are related), but also because the measurement method is the same (or the methods are similar).

Although the problem of CMV has been debated for decades, important unresolved issues remain, two of which are addressed in this article. A first question is whether CMV distorts empirical findings at all. Conventional wisdom holds that CMV is pervasive and can seriously bias research findings. Podsakoff et al. (2012) provide a review of the evidence. However, some researchers view the damaging effects of CMV as an urban legend. For example, in an early paper on self-reported affect and perceptions at work, Spector (1987) surmised that "the problem [of CMV] may in fact be mythical" (p. 442). More recently, Lance et al. (2010) concluded that "method variance occurs frequently, but when substantive measurement facets are excluded as alleged methods, common method variance does not appear to be as robust and threatening as many seem to think" (p. 448). One problem is that prior conclusions about the effects of CMV often rely on illustrative examples, Monte Carlo simulations assuming particular parameter values, or meta-analyses of data sets (sometimes limited in scope) for which methods effects can be estimated. In contrast, this article demonstrates that under simplifying but realistic assumptions, it is straightforward to derive the effect of different levels of CMV on observed correlations analytically, while taking into account other relevant determinants of observed correlations besides CMV. This allows us to demonstrate the conditions under which CMV will distort observed correlations relative to the "true" trait correlations as well as the severity of the distortion.

A second question is how to detect the presence of CMV and the severity of potential negative consequences in one's own research. Podsakoff et al. (2003) and Podsakoff et al. (2012) catalogued and critically evaluated the available approaches. Arguably the most commonly used, but also the most deficient, method is the so-called Harman one-factor test. For example, based on evidence from a related field, 65% of 145 articles published between 2011 and 2015 in leading Information Systems journals relied on the test to detect common method bias (Aguirre-Urreta and Hu 2019). Despite forceful conceptual critiques of the Harman one-factor test (e.g., Hulland et al. 2018, Appendix), some researchers continue to recommend the test and it remains popular. As a case in point, Fuller et al. (2016, p. 3197) acknowledge that "Harman's one-factor test cannot consistently produce an accurate conclusion about biasing levels of CMV in data". Nonetheless, these authors proceed to conclude that "this study indicates that the most commonly used post-hoc approach to managing CMV—Harman's one-factor test—can detect biasing levels of CMV under conditions commonly found in survey-based marketing research" (idem, p. 3197). Our second goal is to critically analyze and rebut such claims and to demonstrate that the Harman one-factor test (a) is invalid and (b) should not be used to detect the presence of CMV in one's data.

We should note that the purpose of this article is not to provide a catalogue of methods to prevent, detect, or cope with CMV, as several prior studies across disciplines have already done so (e.g., Podsakoff et al. 2003, 2012) and a full treatment of these issues would require another article. Our aim is more modest, namely, to show that two common beliefs related to CMV are incorrect and potentially dangerous.

Does CMV distort observed correlations between measures of constructs?

This section reviews the available evidence on the effects of CMV on observed correlations and demonstrates that the

[1] There are differences of opinion in the literature about what counts as a method (Podsakoff et al. 2012). Many researchers subscribe to the broad definition of method used here. However, Lance et al. (2010) argue that method should essentially be restricted to different measurement instruments (e.g., different items to measure the Big Five personality traits) and different response formats (e.g., Likert, Thurstone, and semantic differential scales) for measuring the same trait. They specifically claim that raters and measurement occasions should not be considered as methods because in some areas (e.g., multisource performance ratings) these have been shown to constitute substantive measurement facets, not method facets. It is undeniable that method factors may sometimes have substantive implications. For instance, social desirability might correlate with, say, true yielding to persuasive communications and self-report measures used to assess yielding. However, it is incorrect to reverse the argument and claim that if a variable could, under certain conditions, function as a substantive construct, it cannot be a method factor in a different situation. The fact that an umbrella could perhaps be used as a hockey stick does not imply that it should not be used to protect oneself against the rain.

biasing effect of CMV depends on the sign and size of the trait correlation and the degree of measure unreliability, in addition to the sign and size of the method loadings. It shows that for commonly encountered combinations of these factors the bias due to CMV can be substantial.

Prior research on the biasing effects of shared method variance

Podsakoff et al. (2012) reviewed different streams of literature investigating the severity of method bias in empirical studies in marketing and management. They consider meta-analyses of prior multitrait-multimethod (MTMM) investigations and studies assessing the effects of specific method biases on relationships between constructs. MTMM studies make it possible to decompose the total variance in observed measures into trait, method, and unique factor variance and to examine the effect of method variance on observed correlations relative to trait correlations. Podsakoff et al. (2012) show, among other things, that across five meta-analytic studies involving between 11 and 70 MTMM matrices, the variance in individual items attributable to various method factors ranged from 18 to 32%, and the estimated trait correlations were inflated between 38 and 92% by CMV. These figures suggest that the amount of CMV in observed measures is substantial and that CMV can seriously distort estimates of trait correlations based on observed correlations.

Despite such troubling findings, Lance et al. (2010) argue that certain method factors may actually be substantive factors and that the attenuation of observed correlations due to unreliability of measurement may offset the inflation of observed correlations due to CMV. This reasoning implies that CMV may not seriously distort observed correlations, particularly when researchers do not take into account measurement error. In contrast, Fuller et al. (2016) acknowledge that CMV bias might be substantial, but they claim that "no researcher can draw accurate conclusions regarding the accuracy of CMV rates in real data" (p. 3194). Therefore, they resort to simulations to determine when CMV will bias estimated correlations. We agree that the true data generating process cannot be known with real data but disagree that the logical implication is to instead rely on data simulations. Our analytical results will provide more detailed and novel insights into the biasing effects of CMV than is possible with simulations.

Consider the model in Fig. 1, which has two observed measures x_1 and x_2 that are each a function of a substantive factor (F_1 and F_2), a method factor (M_1 and M_2), and a unique factor (δ_1 and δ_2). In equation form:

$$x_1 = \lambda_1 F_1 + \mu_1 M_1 + \delta_1 \tag{1}$$

$$x_2 = \lambda_2 F_2 + \mu_2 M_2 + \delta_2, \tag{2}$$

where λ_1 and λ_2 are the trait (substantive) loadings and μ_1 and μ_2 the method loadings. Let the variances of F_i and M_i be one and assume that M_i is uncorrelated with F_i. Furthermore, assume that the unique factors δ_i are uncorrelated with each other and with F_i and M_i.

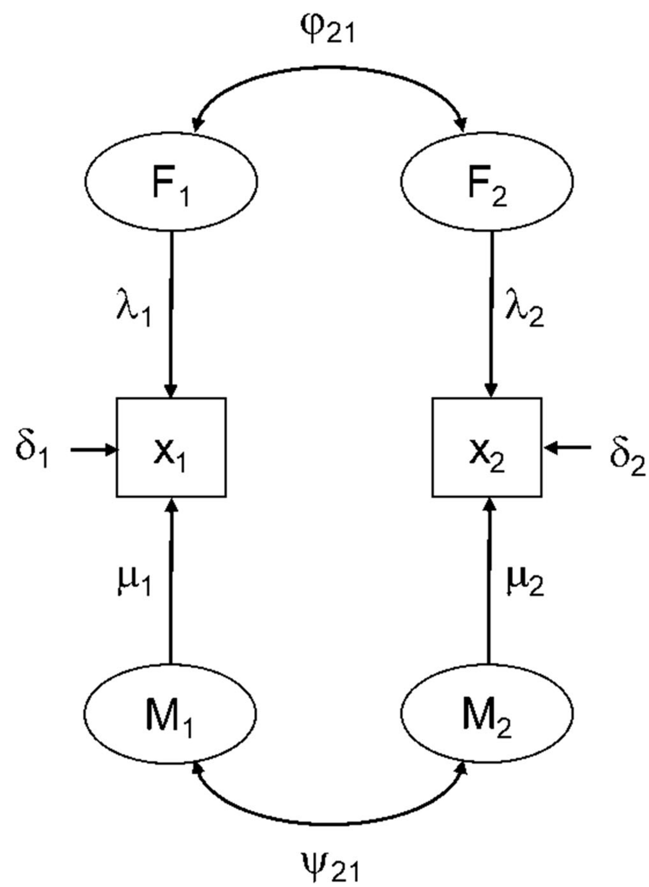

Fig. 1 Model of two observed variables that are each a function of a substantive (trait) factor, a method factor, and a unique factor. Note: The common method factor model is a special case of this model, which is obtained when the method correlation (ψ) equals 1 (unity)

The covariance between the observed variables x_1 and x_2 (i.e., the observed covariance) is then:

$$Cov(x_1, x_2) = \lambda_1\lambda_2\varphi_{21} + \mu_1\mu_2\psi_{21}, \tag{3}$$

where φ_{21} is the trait correlation between F_1 and F_2 (i.e., the true correlation between the two constructs that researchers are interested in) and ψ_{21} is the correlation between the method factors M_1 and M_2 (i.e., the method correlation). Similarly, the total variance in each observed variable (or measure) is:

$$Var(x_i) = \lambda_i^2 + \mu_i^2 + \theta_{ii}, \tag{4}$$

where θ_{ii} is the variance of δ_i. The proportions of trait, method, and unique factor variance in each observed measure are, respectively,

$$\frac{\lambda_i^2}{\lambda_i^2 + \mu_i^2 + \theta_{ii}}, \tag{5}$$

$$\frac{\mu_i^2}{\lambda_i^2+\mu_i^2+\theta_{ii}}, \tag{6}$$

and

$$\frac{\theta_{ii}}{\lambda_i^2+\mu_i^2+\theta_{ii}}. \tag{7}$$

When the observed variables are standardized (which we assume from now on), the denominator equals 1, and λ_i^2, μ_i^2, and θ_{ii} indicate the proportion of the total variance in each observed measure that is due to, respectively, the substantive construct (trait variance), the method factor (method variance), and unique influences on x_i.

Cote and Buckley (1988, pp. 580–581) already presented Eq. (3) and noted the following: "As can be seen from Eq. 1 [the well-known Spearman correction for attenuation formula] and 3 [same as Eq. (3) here], random error variance attenuates the observed bivariate correlation between measures. However, method effects can either attenuate estimates of the relationship, inflate estimates of the relationship, or have no effect depending on the correlation between the methods." Although Cote and Buckley (1988), as well as other experts (e.g., Podsakoff et al. 2003, 2012), have long recognized that method bias can either inflate, deflate, or have no effect on observed correlations, many researchers are still under the impression that method effects generally increase observed correlations or that the net effect is zero or small because measurement unreliability offsets the upward bias due to CMV (e.g., Fuller et al. 2016).

Cote and Buckley (1988, pp. 580–581) provide an empirical example to illustrate the effects of method variance, which we repeat here:

> Suppose that measures of two constructs both contain 60 percent trait variance, 20 percent method variance, 20 percent random error, and the true correlation between the traits is 0.5. Further, suppose that the true correlation between the methods is either 0.2, 0.375, or 0.8. From Eq. 3, the corresponding observed correlation between two measures would be 0.34, 0.375, and 0.46, respectively. If the method effects were removed, the measures would contain 75 percent trait variance and 25 percent random error. Under these measurement conditions, the observed correlation between the measures would be 0.375 (using Eq. 1). The example illustrates that method effects attenuate the relationship between two measures when the correlation between the methods is lower than the observed correlation between the measures with method effects removed. With method effects removed, the observed relationship will be unaffected when the correlation between the methods is the same as the observed correlation between the measures. Lastly, method effects will inflate the observed relationship when the correlation between the methods is higher than the observed correlation between the measures with method effects removed. Therefore, the degree to which two measures are correlated depends not only on the traits being measured, but also on the size of the method variance, random error variance, and the correlation between the methods used to measure the constructs.

Cote and Buckley (1988) focus on the effects of shared method variance on observed correlations attenuated by measurement error, but researchers are mostly interested in the trait correlation (.5 in their example) rather than the attenuated trait correlation (i.e., observed correlation). Note that the trait correlation is underestimated in all cases considered by the authors. Cote and Buckley (1988) proceed to use Eq. (3) to simulate observed correlations between two measures assuming (a) trait correlations varying from 0 to 1 in .05 increments and (b) average trait, method and random error variances as well as average method correlations derived from a meta-analysis by Cote and Buckley (1987) of 70 MTMM studies across the social sciences (both overall and separately for personality and attitude measures). The authors' simulations show that for trait correlations of .3 or higher, observed correlations underestimate the true correlation, whereas for trait correlations of .2 or lower, observed correlations overestimate true correlations.

Lance et al. (2010) expand on Cote and Buckley (1988) and conclude, based on theoretical arguments and an analysis of 18 MTMM matrices, that "there is a kernel of truth to the urban legend that common method effects inflate monomethod correlations but it is a myth that monomethod correlations are larger than correlations among the constructs themselves, and this is because of the offsetting and attenuating effects of measurement error. Thus, monomethod correlations are generally not inflated as compared to their true score counterparts" (p. 448). We believe that it is premature and potentially dangerous to conclude, as Lance et al. (2010) do—based on a small sample of studies where this was the case—that the inflationary effect of shared method variance will routinely be offset by the deflationary effect of measurement error. Moreover, it seems ironic to rely on one data deficiency (i.e., measurement error) to offset another data deficiency (i.e., method effects). Instead, it is important to understand when and to what extent shared method variance will bias research findings, as we do below.

When and how will shared method variance bias research findings?

Cote and Buckley (1988) noted that method effects can inflate or deflate observed correlations, or leave observed correlations unchanged, depending on the size of the method variance, random error, and the correlation between the methods, and they presented illustrative examples to demonstrate this point. Using Eq. (3), more specific conclusions about the biasing effects of shared method variance can be derived, and the results of Cote and Buckley (1988) can be extended by considering the sign and size of the trait and method correlations as well as the sign and size of the method loadings.[2]

In Eq. (3), we can assume without loss of generality that both trait loadings (λ_1, λ_2) are nonnegative because the observed measures x_1 and x_2 can always be recoded such that a higher score on x_1 and x_2 indicates a higher standing on the underlying constructs. However, the correlation of the two traits (φ_{21}) could be positive or negative (or zero). We will distinguish several qualitatively distinct situations of bias due to method effects, depending on the sign of the second term on the right-hand side of Eq. (3). Consider first the case where $\mu_1\mu_2\psi_{21} > 0$. This occurs when (a) the correlation between the methods is positive and the method loadings are either both positive or both negative or (b) the method correlation is negative and one method loading is positive while the other is negative.

1. If the trait correlation is *positive*, method effects will always inflate observed correlations relative to $\lambda_1\lambda_2\varphi_{21}$ (where $\lambda_1\lambda_2\varphi_{21}$ is the attenuated trait correlation in the absence of method effects). That is, the observed correlation will be more positive than it would be if there were no method effects.
2. If the trait correlation is *negative*, method effects will also inflate the observed correlation, but in this case the observed correlation will be less negative than it would be in the absence of method effects. More specifically, when $\mu_1\mu_2\psi_{21} < |\lambda_1\lambda_2\varphi_{21}|$, the observed correlation will be biased toward zero, and when $\mu_1\mu_2\psi_{21} > |\lambda_1\lambda_2\varphi_{21}|$, the observed correlation will have the wrong sign (i.e., it will be positive rather than negative).

Consider next the case where $\mu_1\mu_2\psi_{21} < 0$. This occurs when (a) the correlation between the methods is negative and the method loadings are either both positive or both negative or (b) the method correlation is positive and one method loading is positive while the other is negative.

1. If the trait correlation is *negative*, method effects will always deflate the observed correlations relative to $\lambda_1\lambda_2\varphi_{21}$. That is, the observed correlation will be more negative than it would be if there were no method effects.
2. If the trait correlation is *positive*, method effects will also deflate the observed correlation, but in this case the observed correlation will be less positive than it would be in the absence of method effects. More specifically, when $|\mu_1\mu_2\psi_{21}| < \lambda_1\lambda_2\varphi_{21}$, the observed correlation will be biased toward zero, and when $|\mu_1\mu_2\psi_{21}| > \lambda_1\lambda_2\varphi_{21}$, the observed correlation will have the wrong sign (i.e., it will be negative rather than positive).

So far, we have assumed that the two measurement methods are correlated but not identical. Shared method variance becomes *common method variance* (CMV) when the method correlation equals one, that is, when a single common method factor contributes to the correlation between the observed measures. In this case, the distortion of observed correlations depends only on the sign and size of the method loadings (i.e., since $\psi_{21} = 1$, $\mu_1\mu_2\psi_{21} = \mu_1\mu_2$). If both method loadings are positive or both method loadings are negative, method effects will inflate the observed correlation (i.e., make it less negative or more positive) relative to the attenuated trait correlation in the absence of method effects. If one method loading is positive and the other negative, the observed correlation will be deflated (i.e., more negative or less positive than the attenuated trait correlation). The qualitatively distinct situations discussed earlier apply with the method correlation set to one (i.e., $\psi_{21} = 1$).

The effects of common method variance when the measures are equally reliable and the method loadings are equal in absolute magnitude

Equation (3) shows that the correlation between two observed measures depends on six variables: the size of the two trait loadings (λ_1, λ_2), which we assume to be positive without loss of generality; the sign and size of the two method loadings (μ_1, μ_2); and the sign and size of the trait (φ_{21}) and method (ψ_{21}) correlations. This makes it challenging to evaluate the magnitude of method effects and discern possible interactions between the influencing factors. However, we can gain additional insights into method effects with some simplifying but reasonable assumptions.

First, assume that the unique factor variance θ_{ii} is equal to random error variance so that $(1-\theta_{ii}) = \left(\lambda_i^2 + \mu_i^2\right) = \rho_{x_i x_i}$ is the (individual-item) reliability of x_i. Substituting $\lambda_i^2 = \rho_{x_i x_i} - \mu_i^2$ into equation (32) yields:

$$Cov(x_1, x_2) = \left(\sqrt{\left(\rho_{x_1x_1} - \mu_1^2\right)}\right)\left(\sqrt{\left(\rho_{x_2x_2} - \mu_2^2\right)}\right)\varphi_{21} + \mu_1\mu_2\psi_{21}. \quad (8)$$

Note that when x_1 and x_2 are standardized, $Cov(x_1, x_2)$ becomes the correlation between the observed variables x_1 and x_2 (i.e., the observed correlation).

Second, if we assume that x_1 and x_2 are equally reliable and the method loadings are equal in absolute magnitude although the signs might differ (i.e., $\rho_{x_1x_1} = \rho_{x_2x_2} = \rho_{xx}$, and

[2] We thank the reviewers for nudging us to consider these extensions.

$|\mu_1| = |\mu_2| = \mu$), Eq. (8) simplifies to either

$$Cov(x_1, x_2) = \left(\rho_{xx} - \mu^2\right)\varphi_{21} + \mu^2\psi_{21} \tag{9}$$

when μ_1 and μ_2 are both positive or both negative, or

$$Cov(x_1, x_2) = \left(\rho_{xx} - \mu^2\right)\varphi_{21} - \mu^2\psi_{21}, \tag{10}$$

when μ_1 and μ_2 differ in sign. We can then rewrite Eqs. (9) and (10) as follows:

$$Cov(x_1, x_2) = \rho_{xx}\varphi_{21} + \mu^2(\psi_{21} - \varphi_{21}), \tag{11}$$

when μ_1 and μ_2 are both positive or both negative, and

$$Cov(x_1, x_2) = \rho_{xx}\varphi_{21} - \mu^2(\psi_{21} + \varphi_{21}). \tag{12}$$

when μ_1 and μ_2 differ in sign.

Third and finally, assume that the two observed measures are based on the same method of measurement, which implies that the two methods are perfectly positively correlated (i.e., $\psi_{21} = 1$). We then get

$$Cov(x_1, x_2) = \rho_{xx}\varphi_{21} + \mu^2(1 - \varphi_{21}) \tag{13}$$

when μ_1 and μ_2 are both positive or both negative, and

$$Cov(x_1, x_2) = \rho_{xx}\varphi_{21} - \mu^2(1 + \varphi_{21}) \tag{14}$$

when μ_1 and μ_2 differ in sign.

The first term on the right-hand side of Eqs. (13) and (14) indicates the attenuation of the trait correlation by unreliability of measurement (in the special case where both measures are equally (un)reliable). In the bivariate case, observed correlations (whether positive or negative) are increasingly biased toward zero as random measurement error (unreliability) gets larger.[3] The second term on the right-hand side is the contribution of common method variance to the observed correlation.

Consider Eq. (13) first, which holds when μ_1 and μ_2 are both positive or both negative.It is apparent that unless the two substantive (trait) factors lack discriminant validity (i.e., when $\varphi_{21} = 1$ so that $(1 - \varphi_{21}) = 0$), the second term will always be positive (regardless of whether the trait correlation is positive or negative) and will therefore increase the observed correlation (i.e., make it more positive or less negative relative to the attenuated trait correlation). In particular, using Eqs. (13) and (14), we can distinguish the following cases:

1. If the trait correlation is *positive* ($\varphi_{21} > 0$), CMV will always bias the observed correlation upward (i.e., make it more positive).
2. If the trait correlation is *negative* ($\varphi_{21} < 0$), CMV will:
    a. Bias the observed correlation toward zero when the CMV effect is smaller than the absolute magnitude of the attenuated trait correlation (i.e., when $\mu^2(1 - \varphi_{21}) < |\rho_{xx}\varphi_{21}|$).
    b. Reverse the sign of the observed correlation relative to the trait correlation when the CMV effect is larger than the absolute magnitude of the attenuated trait correlation (i.e., when $\mu^2(1 - \varphi_{21}) > |\rho_{xx}\varphi_{21}|$).

It is worth noting that when the trait correlation is positive, $(1 - \varphi_{21})$ is between 0 and 1, whereas when the trait correlation is negative, $(1 - \varphi_{21})$ is greater than one. Therefore, the contribution of method effects to the observed correlation, that is, $\mu^2(1 - \varphi_{21})$, is larger for negative than for positive trait correlations. Furthermore, when the trait correlation is positive, attenuation due to unreliability biases the observed correlation toward zero (i.e., makes it less positive), whereas CMV biases the observed correlation away from zero (i.e., makes it more positive). Thus, unreliability and CMV affect observed correlations in opposite directions and the two effects may balance each other out if they happen to be the equal in magnitude.

However, the net effect of CMV and measurement unreliability is very different when the trait correlation is negative. In that case, attenuation due to unreliability biases the observed correlation toward zero (i.e., makes it less negative), and CMV also makes the observed correlation less negative or even positive. Overall, this means that CMV will distort observed correlations more strongly when the trait correlation is negative rather than positive.

Consider Eq. (14) next, which holds when the absolute values of μ_1 and μ_2 are equal, but one method loading is positive and the other is negative. When a method of measurement has opposite effects on two observed measures, method effects will generally deflate observed correlations (i.e., make them less positive or more negative relative to the attenuated trait correlation). In particular, the following cases can be distinguished:

1. If the trait correlation is *positive* ($\varphi_{21} > 0$), CMV will:
    a. Bias the observed correlation toward zero when the CMV effect is smaller than the attenuated trait correlation (i.e., when $\mu^2(1 - \varphi_{21}) < \rho_{xx}\varphi_{21}$).
    b. Reverse the sign of the observed correlation relative to the trait correlation when the CMV effect is larger than the attenuated trait correlation (i.e., when $\mu^2(1 - \varphi_{21}) > \rho_{xx}\varphi_{21}$).
2. If the trait correlation is *negative* ($\varphi_{21} < 0$), CMV will always bias the observed correlation downward (i.e., make it more negative).

[3] In the multivariate case, measurement error does not necessarily attenuate the partial correlations.

The situation described in Eq. (13) (i.e., the method loadings have the same sign) seems more likely than the situation described in Eq. (14) (i.e., the method loadings have different signs).[4] To appreciate the implications of Eq. (13), Fig. 2 graphs the effects of direction and size of the trait (substantive) correlation, amount of CMV, and degree of measure unreliability on the observed correlation between x_1 and x_2 (when both method loadings are nonnegative). Trait correlation was varied at 8 levels (−.7, −.5, −.3, −.1, .1, .3, .5 and .7) and is shown on the x-axis. CMV was varied at 7 levels ranging from 0 to 60% in 10% increments and is depicted as the grouping variable. There are separate panels for the four levels of measure reliability (.6, .7, .8 and .9). The findings in the figures can be readily extrapolated to additional levels of the design factors.

An ANOVA of the observed correlations on the size of the trait (substantive) correlation, amount of CMV, and measure (un)reliability, and all interactions, reveals large main effects of trait correlation and CMV and two-way interactions of trait correlation with CMV and trait correlation with reliability.[5] The two-way interaction of trait correlation with CMV is clearly visible in Fig. 2. The interaction between trait correlation and reliability is less pronounced.

Figure 2 reveals the following. First, although an increase in CMV always increases the observed correlation, the magnitude of the increase depends on the sign and size of the trait correlation. The effect is strongest for the most negative trait correlation and becomes successively smaller as the trait correlation first becomes less negative and then more positive. This pattern holds regardless of the level of measure unreliability. Second, the attenuating effect of measure unreliability is stronger for more extreme positive or negative trait correlations (particularly in the 0% CMV condition).

For negative trait correlations, measurement unreliability biases the observed correlations toward zero and CMV amplifies the bias. In fact, since the latter effect is stronger, the observed correlations actually become positive (and thus have the wrong sign) even for modest amounts of CMV. For positive trait correlations, measurement unreliability biases the observed correlations toward zero, while CMV has the opposite effect. However, the attenuation effect is more pronounced for stronger trait correlations, whereas the inflationary effect due to CMV is more pronounced for weaker trait correlations.

Under which conditions will observed correlations deviate significantly from the true trait correlation? This question can be answered by constructing a confidence interval around the observed correlation and determining whether the confidence interval contains the true correlation (for a given level of measure unreliability, size of the trait correlation, and amount of CMV). The width of the confidence interval depends on the chosen confidence level (95% in the present case) and the sample size. For purposes of illustration, Figs. 3 and 4 show the confidence intervals for sample sizes of 100 (relatively small) and 500 (relatively large). The figures display the observed correlation for different levels of CMV and measure reliability, with separate graphs for the four negative trait correlations (Figs. 3a and 4a) and the four positive trait correlations (Figs. 3b and 4b). We used Fisher's r-to-z transformation to construct the confidence intervals.

Figures 3a and 4a show that when the trait correlation is negative, even relatively modest amounts of CMV (i.e., 10%) lead to an *upward* bias in observed correlations: the observed correlation is less negative than it should be, and the confidence interval around the observed correlation generally does not include the true trait correlation. At higher levels of CMV the bias can be substantial and observed correlations can have the wrong sign. For example, consider the situation in which a measure contains 30% method variance. In this case, the observed correlation will be significantly positive (at a sample size of 500) for a reliability of .6 even though the true trait correlation is actually −.7; the observed correlation will be significantly positive for reliabilities of .6 and .7 when the true trait correlation is actually −.5; and the observed correlation will always be significantly positive (regardless of measurement reliability) when the true trait correlation is −.3 or − .10. The only cases in which observed correlations will not deviate significantly from the true trait correlation occur at the smaller sample size of 100 when CMV is only 10% and the negative trait correlation is relatively small (i.e., −.3 or − .1).

Figures 3b and 4b show that when the trait correlation is positive, both upward and downward bias in observed correlations are possible. First, at the smaller sample size of 100, downward bias occurs primarily when the trait correlation is high (.7) and when measure reliability is low (i.e., .6 or .7). In contrast, there is an upward bias in observed correlations when the trait correlation is low (i.e., ≤ .3) and CMV is at least 30%.

Second, at the higher sample size of 500, observed correlations exhibit a downward bias primarily when trait correlations are relatively high (i.e., .5 or .7) and CMV is low (0 or 10%), although underestimation can occur for CMV as high as 60% when the trait correlation is high (i.e., .7) and measure reliability is low (i.e., .6). In contrast, upward bias occurs for even modest levels of CMV (i.e., ≥ 20%) when the trait correlation is low (i.e., .1), regardless of measure unreliability.

[4] Lance et al. (2010) re-analyzed 18 MTMM matrices and found that the average method loading was .427 and the average method correlation was .520. These findings support our contention that positive method loadings are more common than mixed positive and negative method loadings (or, more generally, that $\mu_1\mu_2\psi_{21} > 0$ is more common than $\mu_1\mu_2\psi_{21} < 0$).

[5] Since there is only one observation per cell, it is not possible to report statistical tests. However, the sums of squares of reliability, the interaction of reliability with CMV, and the triple interaction are zero. The percentages of the total variation in observed correlations accounted for by the remaining design factors (in decreasing order of importance) are as follows: trait correlation 45 percent; CMV 43 percent; trait correlation by CMV 9 percent; and trait correlation by reliability 3 percent.

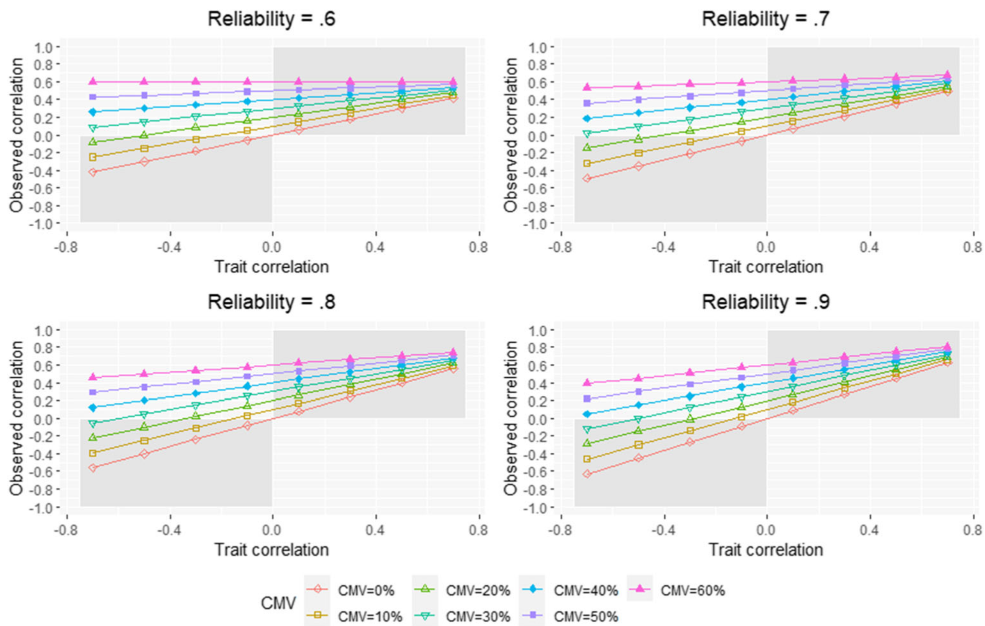

Fig. 2 Observed correlation as a function of trait correlation and CMV (for each level of measure reliability). Note: Grey areas indicate observed correlations for which the sign is the same as for the trait correlation

However, at higher trait correlations (i.e., ≥ .5), overestimation is less common, and CMV has to be substantial (i.e., ≥ 50%) and reliability has to be high (i.e., ≥ .8) in order for observed correlations to be larger than true trait correlations.

What are the main implications of the foregoing analysis, focusing again on the common situation in which the method loadings are both positive or both negative? First, the upward bias due to CMV and the downward bias due to measure unreliability do *not* cancel out for negative trait correlations. In fact, under most scenarios, common method variance biases the observed correlations and this bias can be substantial. That is, even though the trait correlation is *negative*, the observed correlation may be strongly *positive (*a sign reversal).

Second, the upward bias in correlations due to CMV and the downward bias due to measurement unreliability do *not always* cancel each other out for positive trait correlations, even though observed correlations are biased in opposite directions by CMV and measure unreliability. In fact, depending on the magnitude of the trait correlation, level of reliability, and amount of CMV, observed correlations can under- or overestimate trait correlations.

Third, overestimation of trait correlations, which is primarily due to CMV, is more severe at lower levels of trait correlation (.3 or lower), and it occurs for typical levels of CMV that have been found in prior meta-analyses (between 20 and 30%). However, underestimation of trait correlations, which is primarily due to unreliability of measurement, is more severe at higher levels of trait correlations.

Taken together, our analysis demonstrates that CMV *can* cause a marked bias in the size and even the sign of observed correlations. Researchers should not assume or hope that CMV and measurement unreliability will cancel each other out: a negative observed correlation could result from a larger or lower negative trait correlation, whereas a positive observed correlation could result from a larger or lower positive trait correlation or even a negative trait correlation. These results do not lend support to the optimistic belief that the effects of CMV tend to be small or that they will be offset by the effects of (untreated) measurement unreliability (Fuller et al. 2016; Lance et al. 2010). Nor do these results lend support to the (erroneous) belief that CMV can only inflate observed correlations as assumed by some widely used methods that purport to correct for CMV, such as the marker variable approach (Lindell and Whitney 2001).

Is Harman's one-factor test useful for detecting the presence of CMV?

This section demonstrates that the Harman one-factor "test" is not useful for detecting the presence of common method variance because it has both low sensitivity and low specificity (i.e., it is likely to produce both false negatives and false positives). Despite its known fallibility as a test of common method bias (e.g., Hulland et al. 2018), the Harman test survives, perhaps because it is so easy to implement and usually yields the result that researchers desire, namely that there is no evidence of CMV. The idea of the test is that if CMV unduly contaminates data, the first (unrotated) factor in either a principal component analysis (PCA) or an exploratory factor analysis (EFA) will account for a "substantial" portion of the variance in the items measuring the constructs of interest. This criterion is vague, but in practice researchers often assume that

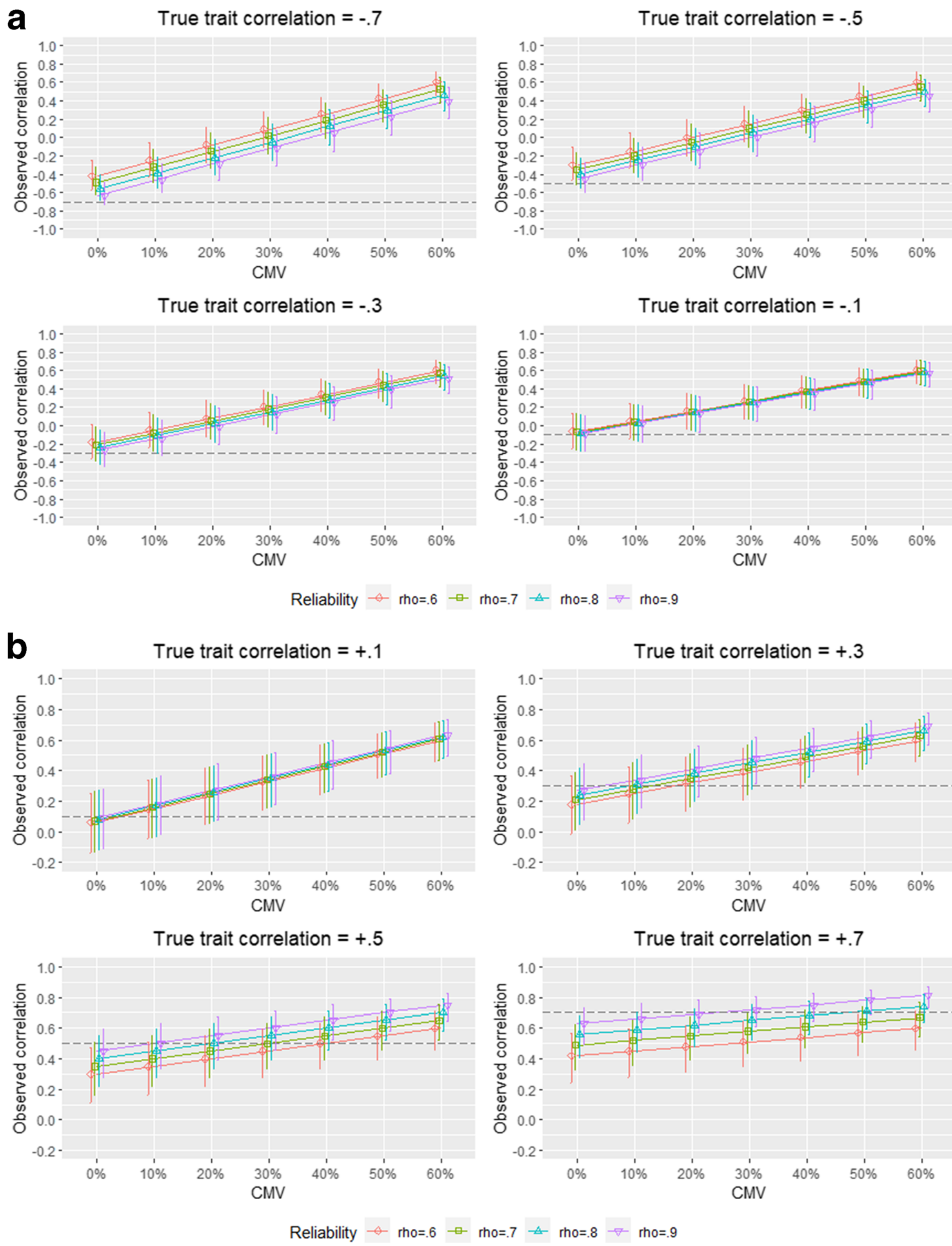

Fig. 3 **a** Confidence intervals around correlation between two observed variables for different levels of measure unreliability, negative trait correlation, and CMV (sample size of 100). Note: The horizontal lines in these graphs show the true trait correlation **b** Confidence intervals around correlation between two observed variables for different levels of measure unreliability, positive trait correlation, and CMV (sample size of 100). Note: The horizontal lines in these graphs show the true trait correlation

there is evidence of CMV if a single eigenvalue is greater than one or if the first (unrotated) factor accounts for 50% or more of the variance in the items (e.g., Cohen and Ehrlich 2019, p. 1432; Fuller et al. 2016). As a case in point, Dai et al. (2020) compared three multi-item trust scales with 25, 20, and 18 items, respectively, and concluded that because 18 eigenvalues exceeded 1 and the factor with the largest eigenvalue accounted for only 11% of the variance in the data, there was no evidence of common method bias.

The Harman one-factor test is liable to two types of error. On the one hand, the Harman test may fail to detect CMV when it is present (a false negative error). There are two sources for this lack of sensitivity of the Harman test: measurement unreliability and the dependence of the test on the number of observed variables analyzed. First, the Harman test is applied to the correlation matrix of the observed variables, and since observed bivariate correlations are attenuated by measure unreliability, the test is affected by a source of variance (i.e., random measurement error) that is not directly related to method variance (i.e., systematic measurement error). Importantly, unreliability makes it less likely that a single, dominant factor will emerge in a principal component or exploratory factor analysis. Second, for a given level of correlation between the observed measures, the likelihood that the first factor will account for more than 50% of the total variance in the data goes down as the number of variables

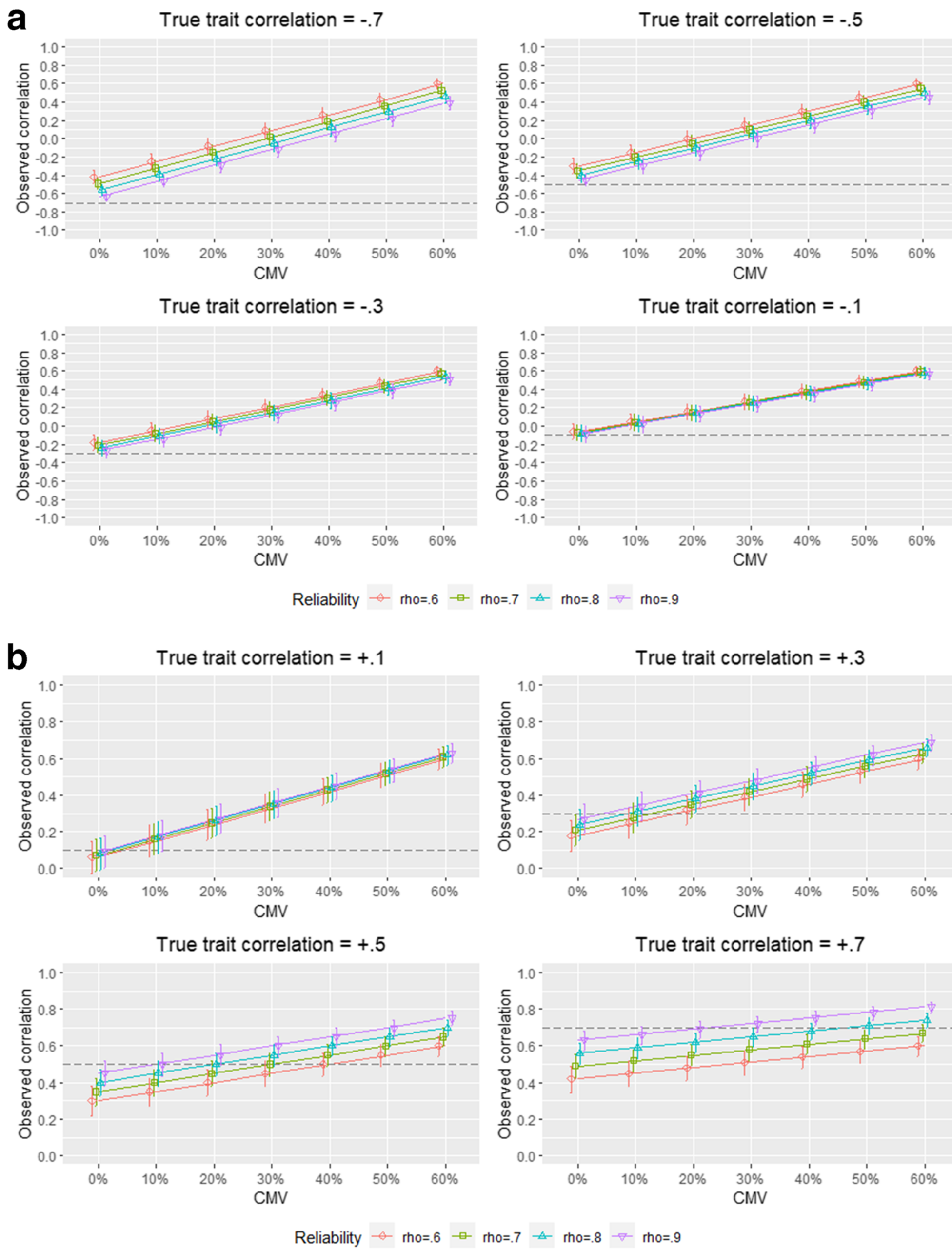

Fig. 4 **a** Confidence intervals around correlation between two observed variables for different levels of measure unreliability, negative trait correlation, and CMV (sample size of 500). Note: The horizontal lines in these graphs show the true trait correlation. **b** Confidence intervals around correlation between two observed variables for different levels of measure unreliability, positive trait correlation, and CMV (sample size of 500). Note: The horizontal lines in these graphs show the true trait correlation

increases. Note that Dai et al. (2020) applied the Harman one-factor test to three trust measures containing a total of 63 items.

On the other hand, the Harman test may suggest the presence of CMV when none exists (a false positive error). The test lacks specificity because it incorrectly assumes that a high degree of communality among the observed variables necessarily reflects common method variance. Particularly when the underlying traits are positively correlated, such as when the Harman one-factor test is applied to three multi-item trust measures, the shared variance among the measures of these traits may be due to substantive overlap between the underlying traits (constructs).

Even if the first factor were to capture common method variance, the criterion that a problem with CMV exists when the first factor accounts for the majority of the variance in the data is misleading. Recall that observed correlations may, on the one hand, exhibit significant upward bias even for CMV levels below 20% and, on the other hand, not deviate significantly from trait correlations even when CMV exceeds 50%.

Let us examine the false positive and false negative errors of the Harman one-factor test more formally. Consider again the model in Fig. 1 with two observed variables. The first eigenvalue v_1 of a 2×2 correlation matrix can be shown to be: $v_1 = (1+|\ r\ |)$, where $|\ r\ |$ is the absolute value of the correlation between x_1 and x_2. Since the sum of the eigenvalues equals the number of variables (2 in the present case), the first eigenvalue is the single eigenvalue greater than one as long as $r \neq 0$. Moreover, because the total observed variance for two standardized variables is also 2, the first eigenvalue accounts for more than 50% of the variance (i.e., $\frac{v_1}{2} = \frac{(1+|r|)}{2} > .5$), unless r is zero.

If we assume that the method loadings are both positive or both negative, we can express the first eigenvalue using Eq. (13) as follows:

$$v_1 = \begin{cases} 1-\rho_{xx}\varphi_{21}-\mu^2(1-\varphi_{21}), & r<0 \\ 1, & r=0 \\ 1+\rho_{xx}\varphi_{21}+\mu^2(1-\varphi_{21}), & r>0 \end{cases} \tag{15}$$

Note that even when the observed correlation is positive, the trait correlation can be negative. Specifically, for $\frac{-\mu^2}{(\rho_{xx}-\mu^2)} < \varphi_{21} < 0$, the trait correlation will be negative while the observed correlation is positive.[6]

Based on Eq. (15), we can distinguish three qualitatively different scenarios with respect to the composition of the first eigenvalue:

> Scenario 1. All the common variance is method variance: $v_1 = 1 + \mu^2$. The magnitude of the first eigenvalue is solely a function of CMV. This occurs when the observed measures contain zero trait variance (i.e., $\lambda^2 = 0$ so that $\rho_{xx} = \mu^2$), in which case Eq. (15) simplifies to $1 + \mu^2$ (r is always positive in this situation).
> Scenario 2. All the common variance is trait variance: $v_1 = 1 + |\rho_{xx}\varphi_{21}|$. The magnitude of the first eigenvalue is solely a function of the (attenuated) trait correlation φ_{21} (where the degree of attenuation is indicated by ρ_{xx}). This occurs when the observed measures contain zero method variance (i.e., $\mu^2 = 0$).
> Scenario 3. The common variance is a mixture of trait and method variance: $v_1 = 1 \pm \rho_{xx}\varphi_{21} \pm \mu^2(1-\varphi_{21})$ (depending on whether $r > 0$ or $r < 0$). This is a combination of scenarios 1 and 2 and occurs when the observed measures contain both trait and method variance.

The Harman one-factor test assumes that only scenario 1 (all common variance is method variance) can occur in practice, which is, in general, an unjustified assumption. In fact, in many situations where positively-correlated traits are measured with multi-item measures, the proportion of trait variance will be high (scenario 2) and the magnitude of the first eigenvalue will strongly depend on shared trait variance rather than CMV.

Let us return to the study by Dai et al. (2020), who compared three multi-item trust scales with 25, 20, and 18 items, respectively. They found that 18 eigenvalues exceeded one and that the first unrotated factor accounted for only 11% of the variance in the data. These authors examined three multi-item trust scales, which presumably contain a substantial amount of shared trait variance, but the Harman one-factor test did not indicate worrying levels of CMV. One explanation is the large number of items used in each of the trust scales (63 items in total). As already mentioned, an increase in the number of items reduces the likelihood of observing a dominant first factor, which can be shown with a simple example. Assume that there are p variables with a uniform positive correlation of r. The first eigenvalue equals $v_1 = 1 + (p-1)r$ in this case. For example, for three variables and r = .5, the first eigenvalue is 2. The proportion of the total variance accounted for by the first factor is $\frac{1+(p-1)r}{p}$, which is easily shown to decrease as the number of observed variables increases (i.e., the derivative of $\frac{1+(p-1)r}{p}$ with respect to p is negative). In other words, for a given uniform correlation, the Harman one-factor test is less likely to indicate a problem with CMV for a larger number of observed variables. In Dai et al. (2020), a uniform, very low correlation of .096 between all 63 items purportedly measuring trust produced the 11% variance accounted for by the first factor. If only two variables were available, the same small correlation of .096 would produce a first factor that accounts for more than 50% of the variance, as we showed previously.

In summary, the Harman one-factor test has so many conceptual deficiencies that it is surprising that some researchers (e.g., Fuller et al. 2016) continue to recommend its use, simply because it performed reasonably well in a simulation, primarily for reasons unrelated to whether or not method variance was present (e.g., because unreliability of measurement offset the bias caused by CMV). The implications of the present analysis are clear: The size of the first eigenvalue is uninformative about the presence and magnitude of CMV in data, and researchers should not use the Harman one-factor test to check for CMV. There is a high probability that a Harman one-factor test flagging a problem with CMV is a false positive, and a high probability that a Harman one-factor test indicating no problem with CMV is a false negative.

[6] The effects of μ^2, ρ_{xx}, and φ_{21} on v_1 can be evaluated by taking the partial derivative of v_1 with respect to each of these terms. This analysis shows the following. First, increasing CMV will increase the magnitude of the first eigenvalue when r > 0 (or $\varphi_{21} > \frac{-\mu^2}{(\rho_{xx}-\mu^2)}$), whereas it will decrease the magnitude of the first eigenvalue when r < 0 (or $\varphi_{21} < \frac{-\mu^2}{(\rho_{xx}-\mu^2)}$). A positive correlation increases the communality of the items, a negative correlation decreases it. Second, increasing unreliability of measurement will decrease the magnitude of the first eigenvalue when $\varphi_{21} < \frac{-\mu^2}{(\rho_{xx}-\mu^2)}$ or $\varphi_{21} > 0$, whereas it will increase the magnitude of the first eigenvalue when $\frac{-\mu^2}{(\rho_{xx}-\mu^2)} < \varphi_{21} < 0$. Third, an increase in the trait correlation (i.e., a less negative or more positive trait correlation) will increase the magnitude of the first eigenvalue when r > 0 (or $\varphi_{21} > \frac{-\mu^2}{(\rho_{xx}-\mu^2)}$), whereas it will decrease the magnitude of the first eigenvalue when r < 0 (or $\varphi_{21} < \frac{-\mu^2}{(\rho_{xx}-\mu^2)}$).

Discussion and conclusion

Some scholars (including reviewers and journal editors) seem to assume that data based on the same method of measurement (including data collected from the same source) are necessarily and possibly fatally distorted by common method variance. Our analytic results do not support this position. However, our analytic results also reject the opposite position, namely, that the presence of common method variance in empirical data and its harmful effects on research conclusions are urban legends or myths.

When a method factor influences two observed variables in the same direction (i.e., μ_1 and μ_2 are both positive or μ_1 and μ_2 are both negative), common method variance will always inflate observed correlations relative to the attenuated trait correlation in the absence of method effects (i.e., observed correlations will be more positive or less negative). However, the magnitude of the inflationary effect depends on other factors, such as the sign and size of the trait correlation. As we have shown, the upward bias caused by CMV is strongest for high negative trait correlations and weakest for high positive trait correlations. In addition, unreliability of measurement can exacerbate or mitigate the inflationary effect due to CMV. In general, greater unreliability will make negative trait correlations less negative and positive trait correlations less positive. However, whereas for positive trait correlations, CMV and unreliability have countervailing effects, for negative trait correlations CMV and unreliability will reinforce each other's effect by making observed correlations less negative and maybe even positive. While errors-of-size are already worrisome enough, errors-of-sign could completely reverse policy evaluations and decisions based on the observed correlations between traits obtained from survey research.

It is possible that a method factor has opposite effects on two observed measures. For example, the tendency to respond in a socially desirable fashion may bias responses upward on one measure and downward on another. If this is the case, method effects will always deflate observed correlations relative to the attenuated correlation in the absence of method effects. When the trait correlation is negative, the observed correlation will be even more negative than it would be if method effects were absent; if the trait correlation is positive, method effects will bias the observed correlation toward zero and may even result in the observed correlation having the wrong sign (i.e., the observed correlation is negative when the trait correlation is actually positive).

Some authors (e.g., Fuller et al. 2016; Lance et al. 2010) have argued that because of the countervailing effects of CMV and measure unreliability, which in some cases neutralize each other, CMV does not have a net biasing effect on observed correlations. This might be comforting to researchers and may lull them into ignoring the potential bias due to common method variance. We believe that this feeling of comfort is misguided for several reasons. First and most importantly, when the effects of a method factor on two observed measures is directionally the same, the opposing effects of CMV and unreliability are restricted to positive trait correlations, a point which prior research has failed to recognize. Second, even when trait correlations are positive, researchers cannot assume that because the two opposing effects sometimes cancel out, this will always be the case. Since the effect of CMV is greater at lower levels of (positive) trait correlation and the effect of unreliability is greater at higher levels of (positive) trait correlations, cancellation is unlikely to be the norm. Furthermore, as seen in Figs. 3b and 4b, there are numerous, commonly encountered conditions in which even a high degree of measure unreliability does not offset the inflationary effects of CMV. To illustrate, when reliability is only .6 and observed correlations may be expected to be substantially attenuated, 20 (30) percent CMV will yield an observed correlation of .24 (.39) when the true correlation is only .1 (.3), an overestimate of 140 (30) percent. Third, conclusions derived from meta-analyses or Monte Carlo simulations, which attest to the counterbalancing effects of CMV and measure unreliability, fail to provide insights into the precise dynamics that give rise to these compensating effects. Monte Carlo simulations can be extremely useful when it is hard to analytically derive relevant results. However, in the present situation analytical derivations are straightforward and they provide deeper insights than reanalyses of existing data (for which the true data generating process is unknown) and Monte Carlo simulation (which depend on an assumed data generating process and post-hoc interpretations of the results obtained). As a case in point, our analytical findings show that the statement of Fuller et al. (2016, p. 3196) that "[r]esults from the simulation indicate that lower to moderate levels of CMV do not inflate correlations and in some cases may deflate correlations" is inaccurate. Moderate levels of CMV *can* lead to inflated observed correlations (relative to the trait correlation), and CMV will never lower correlations (unless the method loadings differ in sign, which was not the case in their simulation), although the use of unreliable measures can have a deflationary (and thus compensatory) effect. Finally, previous simulation studies have only considered the situation of positive trait correlations (e.g., in the simulation by Fuller et al., all pairwise correlations between the seven constructs were positive). This fails to recognize that when method effects have the same sign, the biasing effects of CMV are far more pronounced for negative trait correlations, because (a) the CMV effect dominates the attenuation effect in this case and (b) the biases caused by CMV and measure unreliability are consistent rather than competitive.

How, then, should researchers deal with the problem of CMV? Prior research in multiple domains has already identified various potential solutions and a full treatment of the issues involved would require a separate article. However, we would like to offer several recommendations. The best approach seems to be to take preventative steps to minimize the potential for bias when designing a study. MacKenzie and

Podsakoff (2012) discuss three sets of factors that can lead to method bias depending on whether the bias is due to the difficulty of responding, reduced motivation to respond accurately, or respondent satisficing (see also Viswanathan and Kayande 2012). Based on these factors, they suggest a variety of procedural remedies to counter the problem, which can be very useful (see also Baumgartner and Weijters 2019). Still, despite researchers' best efforts, common method variance cannot always be avoided, and even if a priori remedies are implemented, it is beneficial to check for the presence of CMV in data post hoc and, if necessary, control for potential biases. Researchers should not assume that unreliability of measurement will routinely offset the biasing effects of CMV and that observed correlations will thus approximate trait correlations. That would lead to the ironic recommendation that researchers who expect a substantial amount of CMV in their data should use unreliable scales to measure their constructs to ensure that the two effects cancel each other out (which, at any rate, would only occur for positive trait correlations). Instead, researchers should assess the reliability of their measures, ascertain the extent of common method variance (if there are reasons to believe that common method variance is a problem), and test their substantive hypotheses taking into account both measure unreliability and CMV. Baumgartner and Weijters (2019) present an overview of measurement models that take into account systematic errors of measurement, and they provide an example data set and code to estimate these models.

We acknowledge that it may not always be possible to obtain a good estimate of the amount of CMV with typical mono-method data that researchers have available. Explicit measures of potential method effects (e.g., acquiescent response tendencies, social desirability) may not be available, and implicitly defined method factors, which try to infer method effects from the substantive measures, unfortunately do not always yield accurate estimates of CMV (Richardson et al. 2009). As an alternative, we recommend that researchers perform a sensitivity analysis of the likely effects of systematic measurement error. As an illustration, we will demonstrate a sensitivity analysis using correlation estimates and construct reliabilities for customer satisfaction, customer loyalty, and customer social media use as reported in a recent study by Bill et al. (2020). Our intent is not to criticize the original findings (particularly since their focus differs), but to use actual data to illustrate the proposed sensitivity analysis. Using Eq. (8), the trait correlation can be expressed as a function of the observed correlation ($r_{x_1x_2}$), the measure reliabilities, the method loadings, and the method correlation as follows:

$$\varphi_{21} = (r_{x_1x_2} - \mu_1\mu_2\psi_{21})\left(\sqrt{(\rho_{x_1x_1} - \mu_1^2)}\right)\left(\sqrt{(\rho_{x_2x_2} - \mu_2^2)}\right) \tag{16}$$

Based on Table 2 in Bill et al. (2020), the reliabilities of customer satisfaction, customer loyalty, and customer social media are .88, .72, and .78, and the correlations between satisfaction and loyalty, satisfaction and social media use, and loyalty and social media use are .63, .13, and .07, respectively (calculated from the responses of 334 customers). Figure 5 displays the estimated trait correlation for each combination of constructs as a function of method loadings varied orthogonally from −.54 to +.54 (i.e., 30% method variance) in increments of 10% (assuming the correlation between the methods, ψ_{21}, to be 1). In Podsakoff et al.'s (2012) meta-analytic review the amount of CMV across studies ranged from 18 to 32%, so the range of method variance examined in our scenarios should be a realistic reflections of what might happen in practice.

The graphs show that when the observed correlation is positive and relatively large (i.e., .63), the estimated trait correlation is .79 in the absence of method effects, but when method effects are present, it could be as low as .67 and as high as .97. Regardless of whether the method loadings are positive or negative, the observed correlation always underestimates the true correlation. In contrast, when the observed correlation is positive and relatively low (.07 or .13), the estimated trait correlations in the absence of method effects are .08 and .17, respectively. However, when method effects are considered, the estimated trait correlation could be as high as .70 and .96, respectively, or as low as −.44 and − .38, respectively. In other words, the trait correlation could be much more positive than suggested by the observed correlation, or it could be negative and sizable, in which case the observed correlation is entirely misleading.

When multi-method data are unavailable, which is common in many disciplines, our proposed sensitivity analysis helps researchers gauge the potential harm that CMV can cause in their data. It should be noted that while we varied the size of the method loadings orthogonally from −.54 to +.54, domain experts may be able to impose a narrower range of possible values. For example, it appears unlikely that social desirability will have opposite effects on satisfaction, loyalty, and social media use. However, since all items measuring the three constructs in Bill et al. (2020) were measured with 7-point strongly disagree to strongly agree scales, it is possible that individual differences in agreement tendency across customers contributed to the observed correlation, which would suggest that the method loadings are positive. This means that negative μ_1 and μ_2 values in Fig. 5 can be ignored. If it is reasonable to assume that the method effects on different observed measures are approximately equal, the range of possible method effects can be further reduced.

Our analysis once more demonstrates the importance of evaluating the presence and magnitude of CMV in survey data. However, it is hard to over-emphasize that Harman's one-factor test should *not* be used for this purpose. Researchers cannot assume that the first factor extracted from data necessarily represents method variance, rather than

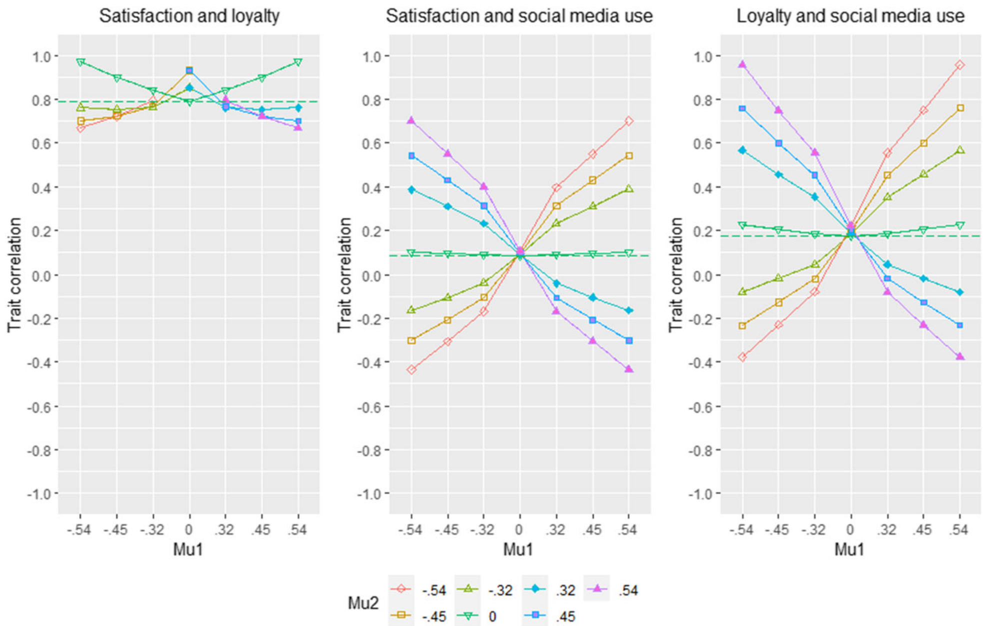

Fig. 5 Estimated trait correlation as a function of the size and sign of method effects for three constructs. Note: The y-axis shows the estimated trait correlation for different levels of method effects; the x-axis shows the variation in method loadings for satisfaction, satisfaction, and loyalty (from left to right), and the different lines in each panel correspond to the method loading for the other construct in the pair. Method loadings of .54 (−.54), .45 (−.45), .32 (−.32) and 0 correspond to percentages of method variance of 30, 20, 10 and 0%. In the graph for the correlation between satisfaction and loyalty, some estimated trait correlations exceed 1, which are omitted. The horizontal lines in each graph show the estimated trait correlation in the absence of method effects

substantive variance. Furthermore, even if the first factor were to represent method variance, the criterion that the first factor should not account for the majority of the variance in observed variables is meaningless. The variance in observed variables is influenced by numerous factors that are unrelated to method variance, and relying on the decision heuristic that only the eigenvalue of the first factor should exceed one and/or that the first factor should account for 50% or more of the variance for method variance to be present may produce both false positives and false negatives. In brief, the Harman one-factor test is an ineffective tool for detecting CMV; researchers should stop using this likely misleading technique; and reviewers and editors should insist that it not be reported in published articles.

Acknowledgements The first author gratefully acknowledges support from the Smeal Chair endowment. The authors would like to sincerely thank the reviewers for their extremely helpful and constructive comments.

References

Aguirre-Urreta, M., & Hu, J. (2019). Detecting common method bias: performance of the Harman's single-factor test. *Advances in Information Systems, 50*(2), 45–70.

Baumgartner, H., & Steenkamp, J. B. E. (2001). Response styles in marketing research: a cross-national investigation. *Journal of Marketing Research, 38*(2), 143–156.

Baumgartner, H., & Weijters, B. (2019). Measurement in marketing. *Foundations and Trends® in Marketing, 12*(4), 278–400.

Baumgartner, H., Weijters, B., & Pieters, R. (2018). Misresponse to survey questions: a conceptual framework and empirical test of the effects of reversals, negations, and polar opposite core concepts. *Journal of Marketing Research, 55*(6), 869–883.

Bill, F., Feurer, S., & Klarmann, M. (2020). Salesperson social media use in business-to-business relationships: an empirical test of an integrative framework linking antecedents and consequences. *Journal of the Academy of Marketing Science, 48*, 734–752.

Brannick, M. T., Chan, D., Conway, J. M., Lance, C. E., & Spector, P. E. (2010). What is method variance and how can we cope with it? A panel discussion. *Organizational Research Methods, 13*(3), 407–420.

Cohen, A., & Ehrlich, S. (2019). Exchange variables, organizational culture and their relationship with constructive deviance. *Management Research Review, 42*(12), 1423–1446.

Cote, J. A., & Buckley, M. R. (1987). Measurement error and theory testing in consumer research: an illustration of the importance of construct validation. *Journal of Consumer Research, 14*(March), 579–582.

Cote, J. A., & Buckley, M. R. (1988). Estimating trait, method, and error variance: generalizing across 70 construct validation studies. *Journal of Marketing Research, 24*(3), 315–318.

Dai, B., Zhang, W., Wang, Y., & Jian, X. (2020). Comparison of trust assessment scales based on item response theory. *Frontiers in Psychology, 11*, 1–8.

Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research, 69*, 3192–3198.

Hulland, J., Baumgartner, H., & Smith, K. M. (2018). Marketing survey research best practices: evidence and recommendations from a review of JAMS articles. *Journal of the Academy of Marketing Science, 46*(January), 92–108.

Lance, C. E., Dawson, B., Birkelbach, D., & Hoffman, B. J. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods, 13*(3), 435–455.

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional designs. *Journal of Applied Psychology, 86*, 114–121.

MacKenzie, S. B., & Podsakoff, P. M. (2012). Common method bias in marketing: causes, mechanisms, and procedural remedies. *Journal of Retailing, 88*(4), 542–555.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539–569.

Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives: examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods, 12*(4), 762–800.

Spector, P. E. (1987). Method variance as an artifact in self-reported affect and perceptions at work: myth or significant problem? *Journal of Applied Psychology, 72*(3), 438–444.

Spector, P. E. (2006). Method variance in organizational research. *Organizational Research Methods, 9*(2), 221–232.

Steenkamp, J.-B. E. M., de Jong, M., & Baumgartner, H. (2010). Socially desirable response tendencies in survey research. *Journal of Marketing Research, 47*(April), 199–214.

Viswanathan, M. (2005). *Measurement error and research design*. Thousand Oaks: Sage.

Viswanathan, M., & Kayande, U. (2012). Common method bias, types of measurement error, and nature of measures. *Journal of Retailing, 88*(4), 556–562.

Weijters, B., Schillewaert, N., & Geuens, M. (2008). Assessing response styles across modes of data collection. *Journal of the Academy of Marketing Science, 36*(3), 409–422.

Weijters, B., Geuens, M., & Schillewaert, N. (2009). The proximity effect: the role of inter-item distance on reverse-item bias. *International Journal of Research in Marketing, 26*, 2–12.

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: the number of response categories and response category labels. *International Journal of Research in Marketing, 27*(September), 236–247.

Williams, L. J., & Brown, B. K. (1994). Method variance in organizational behavior and human resources research: effects on correlations, path coefficients, and hypothesis testing. *Organizational Behavior and Human Decision Processes, 57*, 185–209.

Williams, L. J., Cote, J. A., & Buckley, M. R. (1989). Lack of method variance in self-reported affect and perceptions at work: reality of artifact? *Journal of Applied Psychology, 74*(3), 462–468.

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