



# The Journal of Positive Psychology

Dedicated to furthering research and promoting good practice

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/rpos20>

## Accountability: Construct definition and measurement of a virtue vital to flourishing

Charlotte V.O. Witvliet, Sung Joon Jang, Byron R. Johnson, C. Stephen Evans, Jack W. Berry, Joseph Leman, Robert C. Roberts, John Peteet, Andrew B. Torrance & Ashley N. Hayden

To cite this article: Charlotte V.O. Witvliet, Sung Joon Jang, Byron R. Johnson, C. Stephen Evans, Jack W. Berry, Joseph Leman, Robert C. Roberts, John Peteet, Andrew B. Torrance & Ashley N. Hayden (2022): Accountability: Construct definition and measurement of a virtue vital to flourishing, The Journal of Positive Psychology, DOI: [10.1080/17439760.2022.2109203](https://doi.org/10.1080/17439760.2022.2109203)

To link to this article: <https://doi.org/10.1080/17439760.2022.2109203>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 17 Aug 2022.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)



This article has been awarded the Centre for Open Science 'Open Data' badge.



This article has been awarded the Centre for Open Science 'Open Materials' badge.

## Accountability: Construct definition and measurement of a virtue vital to flourishing

Charlotte V.O. Witvliet<sup>a</sup>, Sung Joon Jang<sup>b</sup>, Byron R. Johnson<sup>b</sup>, C. Stephen Evans<sup>b,c</sup>, Jack W. Berry<sup>d</sup>, Joseph Leman<sup>b</sup>, Robert C. Roberts<sup>c</sup>, John Peteet<sup>e</sup>, Andrew B. Torrance<sup>f</sup> and Ashley N. Hayden<sup>a,g</sup>

<sup>a</sup>Psychology Department, Hope College, Michigan, USA; <sup>b</sup>Institute for Studies of Religion, Baylor University, TX, USA; <sup>c</sup>Department of Philosophy, Baylor University, TX, USA; <sup>d</sup>Department of Psychology, Samford University, AL, USA; <sup>e</sup>Department of Psychiatry, Harvard Medical School, MA, USA; <sup>f</sup>School of Divinity, St Mary's College, University of St. Andrews, UK; <sup>g</sup>Department of Psychiatry, Massachusetts General Hospital, MA, USA

### ABSTRACT

Embracing accountability to others for one's responsibilities within relationships is important for flourishing, yet underexamined. An interdisciplinary team defined the construct of accountability and developed an 11-item single-factor Accountability Scale. In national samples with US census demographic representation (total  $N = 1257$ ), we conducted psychometric analyses using methods from classical test theory (exploratory and confirmatory factor analyses) and item response theory. The Accountability Scale demonstrated internal consistency, construct validity, test-retest reliability, and incremental validity. Accountability correlated positively with relational variables (agreeableness, empathy) responsibility-oriented variables (conscientiousness, self-regulation), virtues (gratitude, forgiveness, limitations-owning humility), relational repair, perceived meaning presence, and flourishing, inversely with symptoms (personality disorders, temper, anxiety, depression), and weakly with searching for meaning and social desirability. Accountability scores superseded demographic variables, conscientiousness, and agreeableness to predict relational repair, perceived presence of meaning in life, and flourishing. We offer the accountability construct and scale to advance human flourishing research and applied work.

### ARTICLE HISTORY

Received 10 May 2022  
Accepted 14 July 2022

### KEYWORDS

Accountability; character; empathy; flourishing; meaning in life; psychometrics; scale development; self-regulation; virtue

People with dispositional accountability are responsive and responsible in their relationships. The importance of accountability becomes apparent when it is missing – when people fail to fulfill their relational responsibilities by violating promises, disregarding feedback, or ignoring commitments. Such failures of accountability can cause costly fractures in relational bonds. Accordingly, whether choosing a partner, employee, mechanic, or medical professional, we value people who welcome accountability – who understand what is expected of them, incorporate feedback, and fulfill their relational responsibilities. When people are accountable, they give others what they are due, and they flourish (Evans, 2021; Peteet et al., 2022a).

Emerging work in philosophy, psychiatry, psychology, and theology conceptualizes accountability as a relational virtue that fosters social bonds, meaning, and flourishing (Evans, 2021; Peteet et al., 2022a; Torrance, 2021). We believe this forward-looking approach to accountability has relevance across a range of contexts in which responsibility and responsibility in relationships matter (e.g., diversity, equity, and inclusion; finance and philanthropy; healthcare and Twelve Step programs; politics and policing; education and civic organizations; family and work life).

Accountability research has been most developed in human resource management and organizational contexts where people feel accountable especially when they perceive the possibility of having to explain their actions to an evaluator who can reward or sanction them (Frink & Klimoski, 1998; Hall et al., 2017). Focusing on the dynamic connection between individuals and institutions, Lerner and Tetlock (1999) reviewed contextual factors in how people respond to evaluators, cautioning that too often, accountability has been invoked as a panacea. They advocated research on internal processes involved in accountability in work and everyday life. Accordingly, we offer an accountability construct and scale that could aid researchers in understanding and cultivating accountability across a range of relationships to facilitate healthy relational bonds, perceived meaning, and flourishing.

### Accountability construct definition and features

As an interdisciplinary team spanning philosophy, theology, psychology, psychiatry, sociology, and criminology, we collaborated to define the construct of accountability, clarify its overarching aspects, delineate its features, and

develop a scale to measure it. In our definition, *people with the virtue of accountability are responsive to the input of others to whom they owe a response with respect to actions those others can rightfully expect from them, and they responsibly improve their attitudes, thoughts, emotions, and actions in light of these relationships*. In brief, *accountable people are relationally responsive and responsible*.

This construct of accountability has two overarching aspects. Accountable people welcome being

- (a) *responsive to others to whom they owe a response* – receptive to others' capable, good input and openly providing them with transparent explanations of decisions and actions; and
- (b) *responsible for their attitudes, thoughts, emotions, and actions in light of these relationships* – working to correct or improve their responses so that they have a positive impact.

The following features further delineate the construct: 1) Accountable people *seek to know the expectations* of people with the standing to hold them accountable. 2) They are not reluctantly responsive, but rather *willingly answer and fulfill their responsibilities* to people who justly hold them accountable. 3) They are *transparent and honest* with people to whom they are accountable, neither withholding information nor presenting only favorable information about themselves. 4) They are neither automatically dismissive of nor servile in conforming to whatever people tell them to do, but rather *discern* whether goals are good – wisely questioning, resisting, and/or reporting when people try to hold them accountable in inappropriate ways or with inappropriate goals. 5) They *value and respect the people* to whom they are accountable and those impacted by their actions or inactions. 6) They *welcome learning* how to improve, modify, and correct their behavior based on input or feedback. 7) They are *willing to accept their responsibility* even when it is difficult or costly to them. 8) Accountable people *see it as good* that they fulfill what is properly owed to those to whom they are accountable. Collectively, these features support wise interpersonal receptivity, responsivity, and responsibility in relationships.

### **Accountability roles and goals: accountees and accountors**

Welcoming accountability to others for fulfilling one's responsibilities can be shown in a variety of relational contexts – personal and professional, hierarchical or horizontal, with contractual or covenantal commitments, shaped by individualistic or communal norms, with explicit or implicit understandings of roles and goals. In addition to the broad sense that people are accountable to

others for their impacts, people in accountable relationships benefit from clear roles and shared understanding of expectations about who answers to whom, for what responsibilities, and toward what end.

Because accountability-related roles can be fluid, even in hierarchies, we refer to the person in an accountable role as the *accountee* who answers to the *accountor* for fulfilling relevant responsibilities (Bergsteiner & Avery, 2003). The accountor needs to have the standing to ask another person to fulfill responsibilities within a specific domain (Peteet et al., 2022a). For example, a teacher in the role of accountor for a student accountee has standing or practical authority to ask the student to complete an assignment, but not to provide a personal favor. The student as accountee is answerable to the professor, but not to the resident director, for class homework (although the student will likely need to meet broader academic standards to remain a student with on-campus housing privileges).

The roles of accountee and accountor depend on particular goals and responsibilities that can vary even in hierarchical relationships. For a professor and student, some of the bi-directional expectations are communicated explicitly in a syllabus whereas others are more implicit in the culture of the school and department. As an accountee, a student is answerable to the professor to pursue the legitimate learning goals of the course by participating in class and completing assignments and exams in particular ways (e.g., with academic honesty, according to a purpose and format, on time, readily explaining their approach to the work, properly acknowledging sources, incorporating feedback to improve). Conversely, the professor as accountee is similarly accountable to the student to fulfill the objectives of the course for advancing the knowledge and skills of the students. Students are also mutually accountable to each other in the learning community, especially when group projects involve collaboration toward a shared goal – responsibly fulfilling their roles in relation to one another with respect to the project aims, elements, and timeline. Such mutual accountability to peers is then nested within accountability to the professor, and with accountability for the academic integrity standards of the institution and the ethics of the relevant discipline.

### **Related psychological constructs**

Accountability involves being both *relationally responsive* and *responsible*. Accordingly, we theorized that accountability would correlate positively with the Big Five personality traits of agreeableness and conscientiousness (John et al., 1991), while being conceptually distinct from these traits. For example, agreeable people could be highly helpful, trusting, cooperative, and

readily say ‘yes’ to others’ requests, yet they may struggle with the forethought, organization, or focus to follow-through on promised projects. Others may show agreeable qualities toward friends and those needing help, but resist being answerable to those with practical authority to ask something of them. Thus, we predicted a moderate to strong correlation between agreeableness and accountability, but not so high as to indicate redundancy.

Accountability, as we have defined it, is also related to and yet distinct from conscientiousness. Whereas accountability can be exhibited only in the types of relationships in which justice is the framework for what one person owes another, conscientiousness can be shown in tasks performed in isolation. People high in conscientiousness may work diligently and carefully on tasks, yet struggle with the relational aspects of accountability. They might, for example, be so focused and efficient that they do not want to be slowed down or sidetracked by others’ feedback, and they could prefer to work with highly individualistic autonomy, emphasizing their own point of view in a way that excludes the input of others or does not consider their impact on the community. Some may avoid or ignore feedback because they see it as unnecessary, whereas others may fear rejection. Thus, we predicted a moderate to strong correlation between conscientiousness and accountability, but without redundancy.

Because virtues cluster together and relational virtues are shown in response to someone for something, we predicted positive correlations between accountability and gratitude (McCullough et al., 2002) and forgiveness (Berry et al., 2005). One is grateful to someone for a gift or benefit. One grants forgiveness to someone for a hurtful relational injustice. One is accountable to someone for fulfilling responsibilities in the relationship. Because both gratitude and accountability can be exhibited in both positive and negative circumstances, we anticipated a stronger correlation between accountability and gratitude than between accountability and forgiveness, which can only arise in the context of a relational breach (Witvliet, 2020). Another relevant virtue is humility, which includes the inclination to think of others rather than oneself and to hold an accurate rather than inflated or diminished sense of one’s capacities (Van Tongeren et al., 2019). Because accountability involves learning from feedback and responding appropriately, we predicted that accountability would correlate positively with limitations-owning humility (Haggard et al., 2018).

We further reasoned that accountable people would also be more likely to responsibly make things right with others once realizing they have hurt them in some way (Witvliet, 2020). Work on virtues including accountability

has associated them with a sense of meaning in life and flourishing (Evans, 2021). Thus, we predicted a direct correlation between accountability and the likelihood of engaging in relational repair after realizing one’s wrongdoing against another person (Witvliet et al., 2019a), perceived presence of meaning (but not searching for meaning; Steger et al., 2006), and flourishing as feeling good and functioning well personally and relationally with purpose (Keyes, 2002). We further predicted that accountability would go beyond the personality characteristics of agreeableness and conscientiousness to account for measures of relational repair, the presence of meaning in life, and flourishing.

### Current studies

We aimed to develop a scale to assess the accountability construct. To do so, we wrote items to tap accountability features as an interdisciplinary team and incorporated positive psychology expert feedback. We conducted three studies with diverse samples (plus a supplemental study to replicate and extend findings in undergraduates). In Studies 1 and 2, we conducted psychometric analyses in separate national samples. Item and test properties were assessed using both classical test theory (including exploratory and confirmatory factor analyses) and Item Response Theory (IRT). Study 2 also assessed whether accountability was positively correlated with the Big Five personality traits of agreeableness and conscientiousness, yet went beyond them to predict the likelihood of relational repair after wrongdoing, meaning in life, and flourishing. Study 3 was a follow-up study focused on construct validity with an emphasis on empathy and self-regulation, virtues, autonomy, and mental health, as well as test-retest reliability.

The Open Science Framework (OSF) has a public time-stamped registration of the scale development project summary including planned studies with hypotheses addressing confirmatory factor analysis, as well as construct validity, incremental validity, known groups validity, and test-retest reliability (<https://doi.org/10.17605/OSF.IO/J2DES>). After conducting these studies, we registered the Accountability Scale project components and their files containing deidentified data and materials (<https://doi.org/10.17605/OSF.IO/YZHCJ>).<sup>1</sup>

### Study 1

To develop a measure of accountability as a virtue, we conducted an exploratory factor analysis (EFA) using 54 items developed and refined by an interdisciplinary team, and further winnowed based on ratings and comments by positive psychology researchers.

## Method

The interdisciplinary team who had defined the construct developed possible scale items to tap its two overarching aspects and eight features. After rating 61 items for construct fit and readability, the team agreed to further test a set of 58 items. Next, a group of seven positive psychologists provided comments and rated each item based on the combination of readability and fit with the construct. Items were eliminated if mean ratings of quality and construct fit were below the midpoint of a scale that ranged from (1) *poor* to (5) *excellent item*. This yielded 54 items that the interdisciplinary team agreed to test using classical test theory statistics, parallel analysis, and the Item Response Theory (IRT) approach of Rasch modeling in tandem.

## Design

We designed this EFA study to test a diverse sample of adults selected for US census representation based on self-identified gender, race and ethnicity, age, education, and census region. Qualtrics Panels screened out respondents who failed honesty or awareness checks, engaged in speeding, answered randomly, or gave the same rating for items across accountability and other items.

## Participants

Study 1 tested 484 adults (241 F, 242 M, 1 other), an adequate sample size for the EFA (DeVellis, 2017). Their ages ranged from 18 to 91 ( $M = 46.66$ ,  $SD = 17.23$ ). When participants self-selected the identity that ‘most describes you,’ 62.8% responded White ( $n = 304$ ), and 37.2% self-identified as Hispanic, Latino, or Spanish origin ( $n = 77$ ), Black or African American ( $n = 65$ ), Asian or Asian American ( $n = 23$ ), American Indian or Alaska Native ( $n = 3$ ), Middle Eastern or North African ( $n = 1$ ), Native Hawaiian or other Pacific Islander ( $n = 2$ ), and other ( $n = 9$ ). Regarding self-reported religious and/or spiritual characterization, the sample was 43.6% religious and spiritual, 7.6% religious but not spiritual, 25.2% spiritual but not religious, and 23.6% neither religious nor spiritual.

## Measures

Participants were instructed to ‘Think about how you usually respond to people who hold you accountable. Think about people to whom you owe a response – whether they are peers, those who supervise you, or those you supervise.’ Participants rated 54 statements that tapped each of the two overarching aspects and construct features of accountability with multiple items using a response scale (1 = *Disagree strongly*, 2 = *Disagree somewhat*, 3 = *Neither agree nor disagree*, 4 = *Agree somewhat*, 5 = *Agree strongly*).

## Procedure

We administered an Institutional Review Board-approved informed consent, measurement items, and debriefing form using Qualtrics software, with participant recruitment and payment managed via Qualtrics Panels.

## Results<sup>2</sup>

We conducted a maximum likelihood (ML) EFA of the 54 items. The initial solution showed eight eigenvalues larger than one with a substantial drop between the first (17.648) and second eigenvalues (3.543), suggesting a single factor extraction. We also conducted a parallel analysis to see whether the eigenvalue for the retained factor was larger than the eigenvalue obtained utilizing random data under otherwise comparable conditions (DeVellis, 2017; Hayton et al., 2004). The results from analyzing 50 random datasets—for which each contained 54 variables for a sample of 484 cases—revealed that the eigenvalue for the first factor of the EFA (i.e., 17.648) was substantially larger than the 95<sup>th</sup> percentile (1.778) as well as the average (1.701) of the 50 first eigenvalues generated from the 50 datasets, affirming that a single-factor solution was desirable. After estimating a single-factor model of 54 items (available upon request), we began to remove items with relatively low factor loadings until we had 11 items – three items tapped overarching aspects *a*, *b*, or both *a* & *b* and eight items tapping features 1 to 8 (see the *Accountability construct definition and features* section above for details). The resulting Accountability Scale items all had factor loadings over .500 (see Table 1) and good inter-item reliability (Cronbach’s  $\alpha = .884$ ).

The stringent IRT approach of Rasch modeling supported the unidimensionality of the 11 items, with excellent fit of items to the Rasch rating scale model, indicating that test scores were influenced primarily by the construct of interest rather than construct-irrelevant factors. Furthermore, results showed adequate spread of item difficulties and of person measures on the accountability construct, and no evidence of item bias by gender (see supplemental IRT report pdf <https://osf.io/2audj> and associated output file <https://osf.io/duyp8>).

## Discussion

Study 1 offered evidence of a reliable single-factor scale with 11 items that tapped the construct definition, two overarching aspects, and eight features with an adequate range of item difficulties and little evidence of

**Table 1.** Exploratory and confirmatory factor analysis results: factor loadings.

| Item  | EFA  | CFA  |
|---|------|------|
| (1) I usually welcome being accountable to others. (Overarching aspect <i>a</i> )                     | .589 | .558 |
| (2) I am willing to be held responsible for my contributions on tasks. (Overarching aspect <i>b</i> ) | .638 | .724 |
| (3) I feel responsible for my work with others. (Overarching aspects <i>a</i> & <i>b</i> )            | .662 | .588 |
| (4) I try to understand the perspectives of people who evaluate me. (Feature 1)                       | .667 | .531 |
| (5) I willingly explain my work on a project to people I am responsible to. (Feature 2)               | .621 | .626 |
| (6) I am comfortable showing the details of my work (e.g., for school, job, chores). (Feature 3)      | .632 | .609 |
| (7) I care a lot about whether the people I am accountable to are fair. (Feature 4)                   | .553 | .520 |
| (8) I care about the people affected by what I do. (Feature 5)  | .662 | .586 |
| (9) I welcome corrective feedback from people who evaluate me. (Feature 6)                            | .739 | .600 |
| (10) I take responsibility for my actions even if it costs me. (Feature 7)                            | .627 | .661 |
| (11) I benefit when I am held responsible for my behavior. (Feature 8)                                | .692 | .618 |

Note. The accountability construct aspects and features corresponding to specific scale items are in parentheses.

potential gender bias. This motivated us to confirm the scale's psychometric properties and garner initial evidence of construct and incremental validity.

## Study 2

Using another national sample, we conducted Study 2 to confirm the single-factor structure of the Accountability Scale and assess its construct validity. Specifically, we predicted that accountability would have moderately strong positive correlations with the personality traits of agreeableness and conscientiousness (Petee et al., 2022a) and measures of perceived presence of meaning and flourishing. By contrast, we anticipated a modest correlation between accountability and social desirability, which may inadvertently give higher scores to genuinely virtuous people, such as those who are consistently forgiving, humbly admit mistakes, or do not deliberately say things to hurt others' feelings. Further, we predicted that the Accountability Scale would show incremental validity, going beyond demographics and both agreeableness and conscientiousness to predict the likelihood of relational repair responses after wrongdoing, the presence of meaning in life, and flourishing. Finally, we tested group differences (see below).

## Method

We submitted Open Science Framework materials with predictions before data collection.

## Design

In this study, we first assessed the psychometric quality of the 11-item scale using confirmatory factor analysis (CFA) and two Item Response Theory analyses. Second, to assess initial construct validity, we administered the Big Five personality inventory because the relational and responsible features of accountability commended agreeableness and conscientiousness as key correlates. Third, we assessed incremental validity of the Accountability Scale – beyond the demographic, agreeableness, and conscientiousness variables – to predict the virtue-relevant target variables of engaging in relational repair when one is responsible for wrongdoing against another person, perceived presence of meaning in life, and flourishing. Finally, we assessed known-groups validity by comparing Accountability Scale scores in groups of people who might reasonably be expected to differ on the construct.

## Participants

Participants were 773 adults (389 F, 376 M, 5 'other,' 3 no answer) sampled via Qualtrics Panels to align with US census representation for gender, age, self-identified race and ethnicity, education, and census region. The sample size was based on scale development recommendations for confirmatory-factor analysis studies (DeVellis, 2017), while also allowing for some attrition in a follow-up study. Qualtrics Panels conducted the same quality check as in Study 1, screening out respondents who were likely to fail to provide honest and valid answers.<sup>3</sup>

Participants self-reported ages from 18 to 98 years ( $M = 44.42$ ,  $SD = 16.52$ ). Participants self-selected racial/ethnic identity that 'most describes you,' with 66.8% endorsing White ( $n = 516$ ), and 33.2% endorsing Hispanic, Latino, or Spanish origin ( $n = 99$ ), Black or African American ( $n = 91$ ), Asian or Asian American ( $n = 48$ ), American Indian or Alaska Native ( $n = 10$ ), Middle Eastern or North African ( $n = 3$ ), Native Hawaiian or other Pacific Islander ( $n = 2$ ), and other ( $n = 4$ ). Regarding the self-reported religious and/or spiritual characterization, the sample was 45.1% religious and spiritual, 7.8% religious-but not spiritual, 22.4% spiritual but not religious, and 24.7% neither religious nor spiritual.

## Measures

This study used the following scales, and all items and response options can be found in the OSF materials. We report only the measures analyzed for this study, with Cronbach's alphas based on this sample.

**Accountability Scale.** We used the same instructions and response options as in Study 1, testing the 11-item Accountability Scale.

**Big Five Personality Inventory.** We used the 44-item Big Five Inventory (John et al., 1991), focusing on Agreeableness ( $\alpha = .79$ ) and Conscientiousness ( $\alpha = .83$ ).

**Relational Repair.** This 6-item scale assessed participants' likelihood of reparative responses after actions (or inactions) that hurt others in some way (Witvliet et al., 2019a;  $\alpha = .87$ ).

**Meaning in Life.** Steger et al.'s (2006) scale was administered to assess the Presence of Meaning with 5 items ( $\alpha = .87$ ) and the Search for Meaning in life with 5 items ( $\alpha = .89$ ).

**Flourishing.** Keyes (2002) 14-item flourishing scale assessed feeling good and functioning well with purpose in relationships over the past month ( $\alpha = .94$ ).

**Social Desirability.** We used the 13-item short version of the *true* (1) or *false* (2) social desirability scale (Crowne & Marlowe, 1960) comprised of items 3, 6, 10, 12, 13 (reversed), 15, 16 (reversed), 19, 21 (reversed), 26 (reversed), 28, 30, 33 (reversed);  $\alpha = .72$ .

### Procedure

Institutional Review Board approval was obtained, OSF materials and predictions were uploaded, and Qualtrics Panels solicited national participants who completed informed consent, measures, and debriefing via Qualtrics software.

### Results

To validate the selected accountability items, we conducted a CFA in *Mplus* Version 8 (Muthén & Muthén, 2017) using ML estimation with robust standard errors. The results are presented in Table 1. First, all 11 items had high factor loadings (which ranged from .520 to .724) and good internal reliability ( $\alpha = .859$ ). Second, although the chi-square statistic was significant ( $\chi^2 = 200.815$ ,  $df = 44$ ,  $p < .001$ ) due in part to the large sample size, the common factor model of 11 items had an acceptable fit to data with RMSEA (.068, 90% CI [.059, .078]) in the 'fair fit' range of .050 to .080 (MacCallum et al., 1996), SRMR (.045) smaller than a maximum cutoff of .080 (Hu & Bentler, 1999), and CFI (.906) acceptable

while lower than the desired .950. In sum, the overall CFA results provided support for the 11 items as unidimensional measure of the construct of accountability.

IRT modeling further supported the scale (see IRT report supplement, <https://osf.io/2audj>). Graded response modeling showed an unconstrained model met all assumptions necessary for best practices. Items had high trait discrimination values and homogenous responding patterns. Similar to Study 1, Rasch modeling (<https://osf.io/qpqgt5>) supported the model fit, unidimensionality, and reliability of the items, with no evidence of item bias by gender. In addition, a comparison of item difficulty estimates in Study 1 and Study 2 strongly supported the stability of item structure across the two samples.

### Construct validity correlations

Accountability was positively correlated with agreeableness, conscientiousness, likelihood of relational repair responses, presence of meaning in life, and flourishing, with weak relationships to the search for meaning and social desirability (see Table 2).

### Incremental validity hierarchical regressions

Hierarchical multiple regression analyses showed that the Accountability Scale went beyond demographics (Step 1) and agreeableness and conscientiousness (Step 2) to predict three variables: relational repair after wrongdoing, meaning presence, and flourishing (see Table 3).

### Known-groups validity

Accountability Scale scores differed for people based on follow-up survey preference and recycling tendencies. First, participants who selected 'yes,' indicating they would complete the study a second time four weeks later, had reliably higher accountability scores than the group who said 'no,'  $F(1, 771) = 12.324$ ,  $p < .001$ . A follow-up analysis showed that respondents who predicted they would and actually did complete the study a second time (not exactly 4 weeks later) had marginally higher scores on the Accountability Scale,  $F(1, 771) = 3.08$ ,  $p = .08$ , perhaps reduced by schedule imprecision in survey recontact. Second, participants who reported 'Never' recycling had lower Accountability Scale

**Table 2.** Correlations of accountability with construct validity measures: studies 2 and 3.

| Study 2 <sup>a</sup> |                  | Study 3 <sup>b</sup> |                  |             |                  |                  |                   |
|----------------------|------------------|----------------------|------------------|-------------|------------------|------------------|-------------------|
| Agreeableness        | .42 <sup>†</sup> | Empathy              | .51 <sup>†</sup> | Gratitude   | .37 <sup>†</sup> | Personality Dis. | -.29 <sup>†</sup> |
| Conscientiousness    | .41 <sup>†</sup> | Self-Regulation      | .47 <sup>†</sup> | Forgiveness | .33 <sup>†</sup> | Temper           | -.23 <sup>†</sup> |
| Meaning Presence     | .35 <sup>†</sup> |                      |                  | Humility    | .35 <sup>†</sup> | Anxiety          | -.15*             |
| Meaning Search       | .19 <sup>†</sup> |                      |                  | Autonomy    | .46 <sup>†</sup> | Depression       | -.17*             |
| Flourishing          | .36 <sup>†</sup> |                      |                  |             |                  | Flourishing      | .36 <sup>†</sup>  |
| Relational Repair    | .55 <sup>†</sup> |                      |                  |             |                  |                  |                   |
| Social Desirability  | .23 <sup>†</sup> |                      |                  |             |                  |                  |                   |

Note. <sup>a</sup>Study 2 all variables  $n = 773$ , except Meaning Presence  $n = 770$ . <sup>b</sup>Study 3 Accountability  $n = 234$ ; Empathy and Self-Regulation  $n = 219$ ; Gratitude, Forgiveness, Humility, and Autonomy  $n = 210$ ; Personality Disorders, Temper, Anxiety, Depression, and Flourishing  $n = 211$ . \* $p < .05$ , \*\* $p < .01$ , † $p \leq .001$

**Table 3.** Hierarchical multiple regressions of relational repair after wrongdoing, presence of meaning of life, and flourishing predicted by demographics (model 1), conscientiousness and agreeableness (model 2), and accountability (model 3).

|                       | Relational Repair |                     |                     | Presence of Meaning |                    |                    | Flourishing       |                    |                    |
|-----------------------|-------------------|---------------------|---------------------|---------------------|--------------------|--------------------|-------------------|--------------------|--------------------|
|                       | Model 1           | Model 2             | Model 3             | Model 1             | Model 2            | Model 3            | Model 1           | Model 2            | Model 3            |
| Gender                | .10**             | .03                 | .01                 | .05                 | -.001              | -.01               | -.03              | -.08*              | -.09**             |
| Age                   | .01               | -.15 <sup>†</sup>   | -.11 <sup>†</sup>   | .11**               | -.02               | -.002              | .08*              | -.05               | -.03               |
| Ethnicity             | .00               | -.01                | .01                 | .05                 | .04                | .05                | .03               | .03                | .04                |
| Education             | .01               | .02                 | -.04                | .06                 | .05                | .02                | .14 <sup>†</sup>  | .13 <sup>†</sup>   | .10**              |
| Region                | -.06              | -.05                | -.03                | -.07                | -.06               | -.05               | -.01              | -.01               | -.002              |
| Conscientiousness     | -                 | .12**               | .02                 | -                   | .18 <sup>†</sup>   | .13 <sup>†</sup>   | -                 | .21 <sup>†</sup>   | .16 <sup>†</sup>   |
| Agreeableness         | -                 | .47 <sup>†</sup>    | .35 <sup>†</sup>    | -                   | .29 <sup>†</sup>   | .24 <sup>†</sup>   | -                 | .27 <sup>†</sup>   | .21 <sup>†</sup>   |
| Accountability        | -                 | -                   | .41 <sup>†</sup>    | -                   | -                  | .19 <sup>†</sup>   | -                 | -                  | .20 <sup>†</sup>   |
| Multiple R            | .11               | .54                 | .64                 | .15                 | .43                | .46                | .19               | .45                | .48                |
| R <sup>2</sup>        | .01               | .29                 | .41                 | .02                 | .19                | .22                | .04               | .20                | .23                |
| F                     | 1.95              | 43.60 <sup>†</sup>  | 65.96 <sup>†</sup>  | 3.60**              | 25.08 <sup>†</sup> | 26.06 <sup>†</sup> | 5.58 <sup>†</sup> | 27.00 <sup>†</sup> | 28.32 <sup>†</sup> |
| F change              | -                 | 145.89 <sup>†</sup> | 159.15 <sup>†</sup> | -                   | 76.99 <sup>†</sup> | 26.92 <sup>†</sup> | -                 | 77.72 <sup>†</sup> | 30.29 <sup>†</sup> |
| R <sup>2</sup> change | -                 | .27                 | .12                 | -                   | .16                | .03                | -                 | .16                | .03                |
| n                     | 770               | 770                 | 770                 | 770                 | 770                | 770                | 767               | 767                | 767                |

Note. Standardized coefficients are presented. \* $p < .05$ , \*\* $p < .01$ , <sup>†</sup> $p \leq .001$

scores compared to those who endorsed recycling to some degree, whether 'sometimes,' 'most of the time,' or 'all of the time,'  $F(3, 769) = 5.277$ ,  $p < .001$  (.95 CIs for the 'never' group did not overlap with any other response).

### Discussion

Evidence the psychometric properties of the single-factor 11-item Accountability Scale. Initial construct validity evidence emerged for the Accountability Scale, which showed the predicted positive correlations with the relational trait of agreeableness and the responsible trait of conscientiousness, as well as the likelihood of engaging in relational repair after realizing one is responsible for wrongdoing, perceived presence of meaning in life, and flourishing. The Accountability Scale also was distinct from the search for meaning and social desirability. Initial evidence for group differences showed higher accountability scale scores for those endorsing participation interest in a follow-up study and in recycling. Importantly, the accountability scale also demonstrated incremental validity beyond demographic variables and the theoretically relevant traits of agreeableness and conscientiousness, significantly accounting for additional variance in measures of relational repair, perceived presence of meaning in life, and flourishing consistent with theorizing about accountability as a virtue.

### Study 3

This study was a follow-up to assess additional evidence of construct validity while providing test-retest evidence. Because accountability is conceptualized as a virtue with two overarching aspects of relationality and responsibility, we predicted accountability would show moderately

strong positive correlations with empathy (Davis, 1983) and self-regulation (Carey et al., 2004). Empathic concern and perspective-taking were associated with other relational virtues, including gratitude (McCullough et al., 2002) and forgiveness (Witvliet et al., 2019b). Self-regulation has been theorized to serve as an undergirding mechanism of virtues more broadly (Root Luna et al., 2017) by supporting the capacity to enact appropriate responses in good ways at appropriate times. With respect to accountability, self-regulation is theorized to be particularly important in modulating one's response to guidance and feedback, managing emotions related to evaluation, making corrections and improvements, and fulfilling one's responsibilities to others.

We further assessed the relationship of accountability to the relational virtues of gratitude and forgiveness (McCullough et al., 2002). We also tested the humble capacity to own one's limitations (Haggard et al., 2018) – important because accountability includes a teachable quality with willingness to make corrections where needed, taking responsibility for one's actions even when difficult.

Drawing on mental health theorizing (Peteet et al., 2022a), we assessed the relationship of accountability to psychiatrically relevant variables. Similar to McCullough et al. (2002) findings for the relational virtue of gratitude, we predicted that accountability would have small to moderate inverse relationships with anxiety (Spitzer et al., 2006) and depression (Kroenke et al., 2009). In light of the way that accountability involves the capacities to value others' perspectives and to self-regulate, we theorized that accountability scores would be inversely correlated with disordered personality indicators of antagonism, disinhibition, negative affect, detachment, and psychoticism (Krueger et al., 2012), as well as having a temper (Grasmick et al., 1993). In addition, we predicted that accountability would have a modest direct correlation with a form of

autonomy that includes healthy congruence with one's values – a view that aligns with accountability as a virtue in which fulfilling one's relational responsibilities is vital for a flourishing life (Petee et al., 2022a, 2022b).

### Method

We uploaded Open Science Framework predictions and materials before data collection.

### Design

This study was designed to offer additional evidence of construct validity and initial test-retest reliability evidence of stability a month later as a follow-up to Study 2.

### Participants

Participants included partial completers of the follow-up study to maximize the  $N$  for test-retest and construct validity tests. This yielded a maximum of 234 adults (131 F, 101 M, 2 Other) who completed the Accountability Scale (with 210–219 participants completing the remaining scales). The demographics of the subset who completed at least the Accountability Scale were similar to those in Study 2.

Participants self-reported ages from 18 to 97 ( $M = 46.43$ ,  $SD = 16.32$ ). Participants self-selected racial/ethnic identity that 'most describes you,' with 64% endorsing White ( $n = 149$ ), and 36% endorsing Hispanic, Latino, or Spanish origin ( $n = 28$ ), Black or African American ( $n = 35$ ), Asian or Asian American ( $n = 14$ ), American Indian or Alaska Native ( $n = 3$ ), Middle Eastern or North African ( $n = 1$ ), Native Hawaiian or other Pacific Islander ( $n = 2$ ), and other ( $n = 2$ ). Participants also had a broad range of education and US census regions.

### Measures

The dependent variables for this study included the Accountability and the Flourishing scales from Study 2. Additional measures included the following scales available in OSF, with Cronbach's alphas reported for this sample.

**Accountability Scale.** The Accountability Scale consisted of 11 items (see Appendix) on a single-factor scale assessing the tendency to be accountable to others for carrying out one's responsibilities. The scale showed strong internal reliability,  $\alpha = .89$ .

**Empathic Concern and Perspective Taking.** We administered the brief form of Davis's (1983) Interpersonal Reactivity Index summing 14 items assessing Empathic Concern and Perspective-Taking (not Fantasy items because they do not address real-life relationships),  $\alpha = .82$ .

**Short Self-Regulation Questionnaire.** Using Carey et al.'s (2004) scale, we assessed dispositional self-regulation with 31 items, some of which were reverse-scored,  $\alpha = .80$ .

**Gratitude.** The Gratitude Questionnaire six-item form (GQ-6; McCullough et al., 2002) with two reverse-scored items assessed the disposition to be grateful to others for benefits in one's life,  $\alpha = .80$ .

**Forgiveness.** The Trait Forgiveness Scale (Berry et al., 2005) assessed the disposition to be forgiving toward others for wrongdoing with 10 items,  $\alpha = .81$ .

**Limitations-Owning Humility.** Haggard et al.'s (2018) 4-item subscale assessed respondents' disposition to acknowledge their limitations and capacity to make mistakes,  $\alpha = .63$ .

**Autonomy – Authorship/Self-Congruence.** This 5-item subscale of the Index of Autonomous functioning assessed the capacity to make decisions and enact responses that are in healthy alignment or congruence with one's values and identity (Weinstein et al., 2012),  $\alpha = .90$ .

**Anxiety.** The Generalized Anxiety Disorder 7-item scale (GAD-7; Spitzer et al., 2006) is a clinical screening tool that assesses anxiety symptom levels in the past two weeks,  $\alpha = .94$ .

**Depression.** The Patient Health Questionnaire 8-item scale (PHQ-8; Kroenke et al., 2009) is a brief clinical diagnostic tool assess depressive disorder symptoms in the past two weeks,  $\alpha = .93$ .

**Personality Disorders.** The 25-item Personality Inventory for DSM-5 Brief Form–Adult (Krueger et al., 2012) assesses the personality domains of negative affect, detachment, antagonism, disinhibition, and psychoticism. All items were used in an overall score,  $\alpha = .94$ .

**Temper.** We used four items assessing the tendency to show anger toward others using Grasmick et al.'s (1993) scale of low self-control based on psychometric assessment (Piquero & Rosay, 1998),  $\alpha = .84$ .

**Flourishing.** This 14-item Flourishing Scale (Keyes, 2002) combined three hedonic well-being items and 11 eudaimonic items assessing positive belonging in relationships with a sense of purpose in the most recent month,  $\alpha = .96$ .

## Results

### Construct validity correlations

Correlational evidence supported the construct validity of the Accountability Scale (see Table 2). We found predicted positive correlations for accountability in relation to empathy and self-regulation, the virtues of gratitude, forgiveness, and humility, as well as autonomy to enact decisions congruent with one's values. Anticipated inverse associations were small for accountability in relation to disordered personality features, an

angry temper, anxiety, and depression symptoms. Accountability showed the predicted positive association with flourishing.<sup>4</sup>

### **Test-retest reliability**

The Accountability Scale showed moderate stability across two time points, with a median of 38 days between testing of 234 participants who completed Studies 2 and 3. Intraclass correlation coefficient (ICC) estimates and their 95% confidence intervals were calculated using SPSS based on a single-measure, absolute-agreement, 2-way random-effects model. The single measures ICC of .534 (95% *CI* .436 to .620, *F* [233, 233] 3.304, *p* < .001) suggested moderate test-retest reliability, or stability, of the scale (Koo & Li, 2016).<sup>5</sup>

### **Discussion**

Study 3 found construct validity evidence supporting the scale as a measure of accountability, which emphasizes the interconnection of responsivity and responsibility in relationships by showing convergent validity with empathy and self-regulation, both theorized to play a key role in the virtue (Peteet et al., 2022a). As predicted, accountability showed further convergent validity through positive associations with other relational virtues including gratitude and forgivingness, as well as the humble capacity to own one's limitations – important for receiving corrective feedback well and adapting accordingly. In alignment with Peteet et al. (2022a, 2022b), accountability was directly associated with healthy autonomy in which people make decisions and behave in congruence with their values, an approach that can support fulfilling one's responsibilities to others. As predicted, accountability showed small inverse correlations with anxiety, depression, having a temper, and disordered personality features that do not comport with empathy or responsible follow-through on commitments, especially when doing so is challenging.

Additionally, the test-retest analyses found modest reliability which aligns with the construct of accountability as a disposition that can be shown across a wide range of types of relationship contexts. These test-retest reliability scores are not so high as to assert accountability is a fixed trait; rather, results suggest that interventions to promote accountability may enhance the tendency to welcome accountability.

### **General discussion**

The construct of accountability to others for fulfilling relational responsibilities is an understudied virtue. Based on an interdisciplinary team's definition of the construct and delineation of its features, we developed

a theoretically and psychometrically sound scale to advance psychological research and enhance applications of positive psychology for human flourishing.

Based on classical test theory methodology (exploratory and confirmatory factor analyses) and Item Response Theory modeling (Rasch), Study 1 assessed the dimensionality and quality of items for the Accountability Scale. In a separate national sample (Study 2), confirmatory factor analysis, Rasch modeling, and unconstrained graded response modeling provided additional support for the psychometric properties of the single-factor 11-item scale (<https://osf.io/2audj>).

Two studies garnered evidence of construct validity (also supported by a supplemental study, <https://osf.io/6crfu>). Predicted positive correlations in Study 2 associated Accountability Scale scores with the relational trait of agreeableness and the responsible trait of conscientiousness, the likelihood of relational repair after wrongdoing, perceived presence of meaning in life, and flourishing. Group differences reflected higher accountability scores among participants who volunteered for a follow-up study and participants who reported recycling (compared to those who did not).

Study 3 further supported construct validity of the Accountability Scale through its positive correlations with the relational variable of empathy and the responsibility-oriented variable of self-regulation. Accountability also correlated with other relationally oriented virtues of gratitude and forgivingness, as well as humility. Consistent with Peteet et al. (2022a, 2022b), accountability was positively correlated with healthy autonomy to enact one's agency to behave congruently with one's identity and values. Accountability also showed small inverse correlations with indicators of anxious and depressed symptoms, having a temper, and disordered personality features associated with impaired empathy and self-regulation (see Peteet et al., 2022a). Findings again showed positive associations between accountability and flourishing mental health. Collectively, evidence supported the scale's theoretical and practical value in positive psychology research that intersects with personality, virtues, and mental health.

We also garnered evidence for the predictive value of the Accountability Scale beyond existing measures. In the diverse national sample of Study 2, accountability scores went above and beyond all demographic variables plus agreeableness and conscientiousness to account for additional significant variance in scores for relational repair, the presence of meaning of life, and flourishing. An additional supplemental study (<https://osf.io/6crfu>) which replicated correlations also showed incremental validity of the Accountability Scale beyond humility to predict relational repair and flourishing.

The Accountability Scale showed modest test-retest reliability consistent with virtuous dispositions as stable, not static. Accountability can be shown across an array of relationships in life with changeable circumstances that vary over time in how positive or negative they are. This differs from gratitude shown in response to positive gifts from givers, and forgiveness shown in response to negative wrongs from wrongdoers. Overall, the Accountability Scale demonstrated moderate test-retest reliability, as well as good internal consistency, construct validity, and incremental validity.

### Limitations

This research was conducted with US Census representative samples of adults. So, we do not yet know how the construct or scale will suit respondents in global cultural contexts, younger age ranges, or specialized populations for whom relational responsibility may be especially relevant. Similar to gratitude measurement, we are cautious about what we dub the *Mary Poppins effect* – a tendency for some participants to indicate they are ‘practically perfect in every way.’ Like other self-report measures in positive psychology, accountability will also benefit from dyadic research, observer reports, objective measures, experimental designs, and longitudinal research with multicultural samples.

### Future directions

We view the forward-facing virtue of accountability to others for fulfilling one’s responsibilities as important for healthy relationships and human flourishing. Future research will be needed to explore how accountability develops across the lifespan, in personal and professional domains, physiologically, and across cultures – which vary in individual-collective, power distance, and direct-indirect feedback approaches (e.g., Hofstede et al., 2010). Thus, we view accountability as a ripe area for research on human flourishing globally.

One key area to develop is the psychophysiology of accountability. For example, genetic and gender predictors of empathy have indirectly predicted relational virtues of gratitude and forgiveness (Witvliet et al., 2019b, 2018), commending parallel research for accountability. Self-regulation is a focus of virtue theory (Root Luna et al., 2017), suggesting a role for cardiac regulation in facing the stressors of evaluation and feedback in accountability relationships.

Echoing the influence of positive psychology, the emerging field of positive criminology is ripe for work on developing accountability to aid offender rehabilitation, crime and recidivism reduction, and prosocial behavior

through mentoring programs in prisons (Jang et al., 2021; Johnson et al., 2021). Research suggests that offenders who have grown in accountability themselves may be assets in engendering accountable change in other offenders (Hallett et al., 2016). A common feature of transforming programs in prisons is that they reduce social isolation and shame among prisoners through prosocial and responsible behavior change consistent with accountability as a virtue – beyond accountability as merely backward-looking punishment (see Deuchar, 2020).

Positive psychology intersections with mental health may also be enriched by explicit work on accountability given its positive associations with flourishing and its inverse relationships with dysfunctionality. Recent clinical work has conceptualized the developmental and neurobiological underpinnings of empathy and self-regulation for healthy accountability as a complement to values-congruent autonomy. Further perspectives show the relevance of accountability for assessment and treatment, the character of the therapist, clinical professionalism, and training models (Petee et al., 2022a; Witvliet & Petee, 2022).

### Conclusion

We offer accountability as a positive psychology construct and have provided a reliable and valid scale to measure it. We hope this conceptual and measurement work catalyzes positive psychology research and applications across relational domains wherever people – regardless of status or context – will benefit from being responsive and responsible in relationships to promote flourishing.

### Notes

1. Please see the Open Science Framework (OSF) time-stamped, public registration of the accountability scale development plan with hypotheses for analyses reported (<https://doi.org/10.17605/OSF.IO/J2DES>) and the OSF time-stamped public registration of scale development studies with components containing files of materials and deidentified data files (<https://doi.org/10.17605/OSF.IO/YZHCJ>). There, readers will find two registered components with files of supplemental reports to accompany this publication: (a) ‘Accountability Supplement Study 1 & 2 IRT Rasch and Graded Response Modeling.pdf’ (<https://osf.io/2audj>) and (b) ‘Accountability Supplemental Study construct validity & hierarchical regressions with humility.pdf’ (<https://osf.io/6crfu>) with accompanying deidentified data (including variable view labels, item content and response options, as well as reverse-scored items and scale totals). Finally, we note that Transcendent Accountability and Relational Repair scale

disseminations will address predictions relating to those measures as well as transcendent, religious/spiritual, and forgiveness-seeking topics.

2. We decided not to include a twelfth item described in the registration because it did not perform well, likely because the wording 'Being accountable helps me do my best' (EFA study variable name 'HACCT\_25') could be misunderstood as indicating that one needs to be held accountable externally to do one's best rather than embracing accountability as a virtue. Accordingly, analyses of OSF data should use only the 11 items in the Accountability Scale.
3. Upon collecting a sample ( $N = 613$ ), we identified that Qualtrics Panels programming unintentionally excluded participants who did not endorse having a transcendent guide for living. This was because an awareness check item had been embedded in a 'transcendent' question these participants did not receive. Rectifying this and ensuring our sample had religious-spiritual diversity, Qualtrics Panels provided an additional sample of 160 participants who reported no transcendent guide to match the percentage of Study 1 participants indicating they had no transcendent guide; these participants were selected to have US census demographic representation. Thus, the CFA study had a total of 773 participants (we acknowledge a typo in one registered filename that said 713 despite containing data for the 773 participants).
4. To further assess incremental validity, we tested whether accountability (Step 3) would show incremental validity beyond Step 1 demographics, Step 2 empathy and self-regulation in predicting virtues, relational repair, and flourishing. Results showed incremental validity of the Accountability Scale in predicting only humility ( $R^2$  change = .030; total  $R^2 = .422$ ) and likelihood of relational repair ( $R^2$  change = .018; total  $R^2 = .176$ ).
5. Because we had predicted a higher test-retest reliability value, we conducted two additional test-retest studies with new national samples, each selected for an equal gender split. First, we tested 187 participants at a median interval of 14 days, with  $ICC = .541$  (.95  $CI$  .431 to .635,  $F(186, 186) = 3.346$ ,  $p < .001$ ). Second, we modified the scale instructions to add parenthetical examples: Think about *how you usually respond to people who hold you accountable (e.g., parent, spouse, partner, close friend, teacher, supervisor, 'accountability partner,' etc.)*. A sample of 274 participants, with a median retest interval of 13 days, yielded  $ICC = .580$  (.95  $CI$  .496 to .653,  $F(273, 273) = 3.754$ ,  $p < .001$ ). Deidentified data for these additional test-retest studies are available in the associated project (<https://doi.org/10.17605/OSF.IO/DZ3ST>) for the OSF registration.

## Acknowledgments

This work was supported by a grant from the Templeton Religion Trust (TRT 0171). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Templeton Religion Trust. We appreciate the

seven positive psychology researchers who provided feedback and ratings of item quality and fit with the accountability construct, as well as feedback from interdisciplinary colleagues in the Institute for Research in Psychology and Spirituality, the Accountability Summer Seminar, and the Living Accountably conference.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by a grant from the Templeton Religion Trust (TRT 0171). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Templeton Religion Trust.

## Data availability statement

The data described in this article are openly available in the Open Science Framework at <https://doi.org/10.17605/OSF.IO/YZHCJ>

## Open Scholarship



This article has earned the Center for Open Science badges for Open Data and Open Materials through Open Practices Disclosure. Please see the Open Science Framework (OSF) public registration (<https://doi.org/10.17605/OSF.IO/YZHCJ>) of the scale development studies with downloadable deidentified raw data files (see 'variable view' for scale items and response options) as well as reverse-scored items, and scale totals; to locate the registered data and materials by study component, select 'Components,' and for each component, select 'Files.' The registration's 'Associated project' (<https://doi.org/10.17605/OSF.IO/DZ3ST>) also contains public raw data organized in folders for parallel analysis (supporting Study 1) and test-retest additional studies (see Endnote 5).

## ORCID

Charlotte V.O. Witvliet  <http://orcid.org/0000-0003-4364-6816>

## References

- Bergsteiner, H., & Avery, G. C. (2003). Responsibility and accountability: towards an integrative process model. *International Business & Economics Research Journal*, 2(2), 31–40. <https://doi.org/10.19030/iber.v2i2.3763>
- Berry, J. W., Worthington, E. L., Jr., O'Connor, L. E., Parrott, L., III, & Wade, N. G. (2005). Forgiveness, vengeful rumination, and affective traits. *Journal of Personality*, 73(1), 183–225. <https://doi.org/10.1111/j.1467-6494.2004.00308.x>

- Carey, K. B., Neal, D. J., & Collins, S. E. (2004). A psychometric analysis of the self-regulation questionnaire. *Addictive Behaviors, 29*(2), 253–260. <https://doi.org/10.1016/j.addbeh.2003.08.001>
- Carey, K. B., Neal, D. J., & Collins, S. E. (2004). A psychometric analysis of the self-regulation questionnaire. *Addictive Behaviors, 29*(2), 253–260.
- Crowne, D. P., & Marlowe, D. A. (1960). A new scale of social desirability independent of pathology. *Journal of Consulting Psychology, 24*(4), 349–354. <https://doi.org/10.1037/h0047358>
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113–126. <http://dx.doi.org/10.1037/0022-3514.44.1.113>
- Deuchar, R. (2020). 'I Get More in contact with My Soul': Gang disengagement, desistance, and the role of spirituality. *Youth Justice, 20*(1–2), 113–127. <https://doi.org/10.1177/1473225419889195>
- DeVellis, R. F. (2017). *Scale development: theory and applications* (4th ed.). Sage.
- Evans, C. S. (2021). Accountability as part of the human moral condition and as a virtue. In A. C. Pelsler & W. S. Cleveland (Eds.), *Faith and virtue formation* (pp. 275–291). Oxford University Press.
- Frink, D. D., & Klimoski, R. J. (1998). Toward a theory of accountability in organizations and human resource management. In G. R. Ferris (Ed.), *Research in personnel and human resources management* (Vol. 16, pp. 1–51). Elsevier Science/JAI Press.
- Grasmick, H. G., Tittle, C. R., Bursik, R. J., Jr, & Arneklev, B. J. (1993). Testing the core empirical implications of Gottfredson and Hirschi's general theory of crime. *Journal of Research in Crime and Delinquency, 30*(1), 5–29. <https://doi.org/10.1177/0022427893030001002>
- Haggard, M., Rowatt, W. C., Leman, J. C., Meagher, B., Moore, C., Fergus, T., Whitcomb, D., Battaly, H., Baehr, J., & Howard Snyder, D. (2018). Finding middle ground between intellectual arrogance and intellectual servility: Development and assessment of the limitations-owning intellectual humility scale. *Personality and Individual Differences, 124*, 184–193. <https://doi.org/10.1016/j.paid.2017.12.014>
- Hall, A. T., Frink, D. D., & Buckley, M. R. (2017). An accountability account: A review and synthesis of the theoretical and empirical research on felt accountability. *Journal of Organizational Behavior, 38*(2), 204–224. <https://doi.org/10.1002/job.2052>
- Hallett, M., Hays, J., Jang, S. J., Johnson, B. R., & Duwe, G. (2016). *The Angola prison seminary: Effects of faith-based ministry on identity transformation, desistance, and rehabilitation*. Routledge.
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7*(2), 191–205. <https://doi.org/10.1177/1094428104263675>
- Hofstede, G., Hofstede, G., & Minkov, M. (2010). *Cultures and organizations: software of the mind* (3rd ed.). McGraw Hill.
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jang, S. J., Witvliet, C. V. O., Bradshaw, M., Leman, J., & Johnson, B. R. (October 27, 2021). *Religion, spirituality, and accountability in America: Exploring their relationships with psychological well-being and human behavior. Panel for living accountably conference*, Baylor University, Waco, TX.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory—versions 4a and 54*. Berkeley, Institute of Personality and Social Research.
- Johnson, B. R., Hallett, M., & Jang, S. J. (2021). *The restorative prison: Essays on inmate peer ministry and prosocial corrections*. Routledge.
- Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43*(2), 207–222. <https://doi.org/10.2307/3090197>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders, 114*(1–3), 163–173. <https://doi.org/10.1016/j.jad.2008.06.026>
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*(9), 1879–1890. <https://doi.org/10.1017/S0033291711002674>
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*(2), 255–275. <https://doi.org/10.1037/0033-2909.125.2.255>
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149. <https://doi.org/10.1037/1082-989X.1.2.130>
- McCullough, M. E., Emmons, R. A., & Tsang, J. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology, 82*(1), 112–127. <https://doi.org/10.1037/0022-3514.82.1.112>
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus: statistical analysis with latent variables: user's guide (version 8)*.
- Peteet, J. R., Witvliet, C. V. O., & Evans, C. S. (2022a). Accountability as a key virtue in mental health and human flourishing. *Philosophy, Psychiatry, & Psychology, 29*(1), 49–60. <https://doi.org/10.1353/ppp.2022.0008>
- Peteet, J. R., Witvliet, C. V. O., & Evans, C. S. (2022b). Accountability in relation to autonomy, motivation, and psychiatric treatment. *Philosophy, Psychiatry, & Psychology, 29*(1), 69–71. <https://doi.org/10.1353/ppp.2022.0011>

- Piquero, A. R., & Rosay, A. B. (1998). The reliability and validity of Grasmick et al.'s self-control scale: A comment on Longshore et al. *Criminology*, 36(1), 157–173. <https://doi.org/10.1111/j.1745-9125.1998.tb01244.x>
- Root Luna, L., Van Tongeren, D. R., & Witvliet, C. V. O. (2017). Virtue, positive psychology, and religion: Consideration of an overarching virtue and an underpinning mechanism. *Psychology of Religion and Spirituality*, 9(3), 299–302. <https://doi.org/10.1037/rel0000127>
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, 53(1), 80–93. <https://doi.org/10.1037/0022-0167.53.1.80>
- Torrance, A. (2021). The virtue of accountability. *Studies in Christian Ethics*, 34(3), 307–315. <https://doi.org/10.1177/09539468211009755>
- Van Tongeren, D. R., Davis, D. E., Hook, J. N., & Witvliet, C. V. O. (2019). Humility. *Current Directions in Psychological Science*, 28(5), 463–468. <https://doi.org/10.1177/0963721419850153>
- Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2012). The index of autonomous functioning: development of a scale of human autonomy. *Journal of Research in Personality*, 46(4), 397–413. <https://doi.org/10.1016/j.jrp.2012.03.007>
- Witvliet, C. V. O., Root Luna, L. M., VanderStoep, J. V., Vlisides-Henry, R. D., Gonzalez, T., & Griffin, G. D. (2018). OXTR rs53576 and gender predict trait gratitude. *The Journal of Positive Psychology*, 14(4), 417–426. <https://doi.org/10.1080/17439760.2018.1460689>
- Witvliet, C. V. O., Jang, S. J., Berry, J., Evans, C. S., Johnson, B. R., Roberts, R. C., Peteet, J. R., & Torrance, A. (2019a). August. Accountability to people: A new frontier in positive psychology. Presented at the American Psychological Association (APA) annual convention.
- Witvliet, C. V. O., Root Luna, L. M., VanderStoep, J. L., Gonzalez, T., & Griffin, G. D. (2019b). Granting forgiveness: State and trait evidence for genetic and gender indirect effects through empathy. *The Journal of Positive Psychology*, 15(3), 390–399. <https://doi.org/10.1080/17439760.2019.1615108>
- Witvliet, C. V. O. (2020). Forgiveness, embodiment, and relational accountability: victim and transgressor psychophysiology research. In E. L. Worthington Jr. & N. Wade (Eds.), *Handbook of forgiveness (second edition ed., pp. 167–177)*. Brunner-Routledge.
- Witvliet, C. V. O., & Peteet, J. R. (2022). Accountability. In J. R. Peteet (Ed.), *the virtues in psychiatric practice* (pp. 17–32). Oxford University Press.

## Appendix

### Accountability Scale

Think about **how you usually respond to people who hold you accountable**. Think about people to whom you owe a response for your actions or inaction (lack of action) – whether they are peers, those who supervise you, or those you supervise.

Please select a response to indicate how much you honestly disagree or agree with each statement based on **how you typically are in real life**.

| 1                    | 2                    | 3                             | 4                 | 5                 |
|----------------------|----------------------|-------------------------------|-------------------|-------------------|
| Disagree<br>strongly | Disagree<br>somewhat | Neither agree<br>nor disagree | Agree<br>somewhat | Agree<br>strongly |

- (1) \_\_\_ I usually welcome being accountable to others.
- (2) \_\_\_ I am willing to be held responsible for my contributions on tasks.
- (3) \_\_\_ I feel responsible for my work with others.
- (4) \_\_\_ I try to understand the perspectives of people who evaluate me.
- (5) \_\_\_ I willingly explain my work on a project to people I am responsible to.
- (6) \_\_\_ I am comfortable showing the details of my work (e.g., for school, job, chores).
- (7) \_\_\_ I care a lot about whether the people I am accountable to are fair.
- (8) \_\_\_ I care about the people affected by what I do.
- (9) \_\_\_ I welcome corrective feedback from people who evaluate me.
- (10) \_\_\_ I take responsibility for my actions even if it costs me.
- (11) \_\_\_ I benefit when I am held responsible for my behavior.