JOURNAL OF PERSONALITY ASSESSMENT
2016, VOL. 98, NO. 2, 209–221
http://dx.doi.org/10.1080/00223891.2015.1068174

Routledge
Taylor & Francis Group

The Development and Validation of the Comprehensive Intellectual Humility Scale

Elizabeth J. Krumrei-Mancuso and Steven V. Rouse

Department of Psychology, Pepperdine University

ABSTRACT

A series of studies was conducted to create the 22-item Comprehensive Intellectual Humility Scale on the basis of theoretical descriptions of intellectual humility, expert reviews, pilot studies, and exploratory and confirmatory factor analyses. The scale measures 4 distinct but intercorrelated aspects of intellectual humility, including independence of intellect and ego, openness to revising one's viewpoint, respect for others' viewpoints, and lack of intellectual overconfidence. Internal consistency and test-retest analyses provided reliable scale and subscale scores within numerous independent samples. Validation data were obtained from multiple, independent samples, supporting appropriate levels of convergent, discriminant, and predictive validity. The analyses suggest that the scale has utility as a self-report measure for future research.

ARTICLE HISTORY
Received 21 July 2014
Revised 11 May 2015

To a certain extent, scientific progress in understanding humility has been hampered by the challenges involved in defining and measuring this virtue. Definitions of humility are diverse and often multifaceted. Some definitions emphasize intrapersonal qualities such as an accurate view of oneself, one's abilities, and one's accomplishments, acknowledgment of one's limitations, and low self-focus (Tangney, 2000). Other definitions emphasize interpersonal qualities such as being other-oriented rather than selfish (e.g., Davis, Worthington, & Hook, 2010). Most humility researchers agree that humility involves having an accurate view of self, including an accurate perspective of one's place relative to other people and circumstances (Bollinger & Hill, 2012; Davis et al., 2010; Emmons, 1999; Exline, Campbell, Baumeister, Joiner, & Krueger, 2004; Tangney, 2000).

The research reported here focuses on the distinct but parallel construct of intellectual humility (IH). Whereas humility is based in perceptions of self, IH relates to one's perceptions of knowledge, beliefs, opinions, and ideas. That is, IH involves being humble with regard to the way one acquires and applies knowledge (Stafford, 2010). Just as humility involves being able to accurately see oneself as an imperfect being who is not at the center of the universe, IH involves accepting that one's knowledge and cognitive faculties are limited and imperfect. Thereby, IH presumably holds unique potential to promote human thriving through tolerance of other's ideas, collaboration, and civil discourse. IH has received the most attention within the disciplines of philosophy and theology, but has become of recent interest in the field of psychology. Most published work on IH is theoretical in nature (e.g., Baehr, 2011; Roberts & Wood, 2007; Zagzebski, 1996), but empirical studies are beginning to appear. To encourage further research on IH, the goal of this research was to offer a way to quantify IH through the development and validation of a comprehensive, theoretically grounded self-report measure.

Conceptualizations of intellectual humility

Humility is often classified as a virtue (e.g., Exline et al., 2004). The term *virtue* can bring morality to mind; however, virtues can have an intellectual rather than moral dimension (Baehr, 2011). Thus, whereas humility could be considered a moral virtue that promotes being a person of good character, IH can be classified as an epistemic virtue that promotes being a good knower (e.g., Brady & Pritchard, 2003; Paul & Elder, 2001; Stafford, 2010).

Diverse theoretical foundations have resulted in a variety of definitions of IH. For example, Roberts and Wood (2007) defined intellectual virtues as the absence of their vice counterparts, such as intellectual arrogance, and added that IH becomes a virtue rather than merely the absence of a vice when a lack of concern for one's intellectual status is taken over by an overriding concern for knowledge and truth.

Another conceptualization of IH involves the Aristotelian mean between two vices, such as an accurate view of one's cognitive capacities as a mean between the two extremes of intellectual grandiosity and intellectual diffidence (e.g., Zagzebski, 1996). Thus, although IH calls for openness to new ideas (Gruppen, 2014; McElroy et al., 2014) and openness to changing one's viewpoint when warranted (Hopkin, Hoyle, & Toner, 2014), it does not involve being self-effacing or yielding to others at all times (Vorobej, 2011). An intellectually humble person is able to find the right balance between dogmatically rejecting the dissenting viewpoints of others and yielding too quickly in the face of intellectual conflict (Vorobej, 2011). Thus, IH involves being able to embrace one's beliefs with

CONTACT Elizabeth J. Krumrei-Mancuso Elizabeth.krumrei@pepperdine.edu Department of Psychology, Social Science Division, Pepperdine University, 24255 Pacific Coast Highway, Malibu, CA 90263.

© 2016 Taylor & Francis

confidence while being open to alternative evidence. This ability is derived from an awareness of one's epistemic limits and the fallibility of one's knowledge (e.g., Baltes & Smith, 2008; Hopkin et al., 2014; Jones, 2012; McElroy et al., 2014; Ryan, 2012) and the ability to distinguish what one knows from what one does not know (Elder & Paul, 2012).

Other theories emphasize that knowledge and strength of belief should be derived from one's epistemic position rather than other sources (Jones, 2012). For example, IH involves the disinclination to regard a belief as true simply because it is one's own (Gregg & Mahadevan, 2014); rather, IH involves a lack of overinvolvement of one's ego with one's viewpoints (Sherif & Cantril, 1947; Wayment & Bauer, 2008; L. T. Zagzebski, personal communication, June 20, 2013). As a result, an intellectually humble person is able to exchange differing viewpoints without causing or taking offense (McElroy et al., 2014) and is able to respect those with different views (Gruppen, 2014).

In our work, we consolidated existing theories of IH into our definition of IH as a nonthreatening awareness of one's intellectual fallibility. Consistent with the literature on IH, such awareness should result in openness to revising one's viewpoints, lack of overconfidence about one's knowledge, respect for the viewpoints of others, and lack of threat in the face of intellectual disagreements. As such, we conceptualized IH as both an intrapersonal and interpersonal construct that would be associated with outcomes such as open-mindedness and tolerance for others.

Existing measures of intellectual humility

The earliest measure of IH we could locate consists of a coding scheme for structured interviews used to evaluate wisdom in reasoning about social situations, including recognition of uncertainty and the limits of knowledge (Grossmann et al., 2010). Trained coders rate interview transcripts for IH on a scale from 1 (*not at all*) to 3 (*a lot*). The initial study did not report interrater reliability of the IH domain separate from the overall wisdom ratings; however, a subsequent study reported interrater reliability for IH as greater than .90 (Kross & Grossmann, 2012). Although this context-rich assessment of IH poses clear strengths, our goal was to create a more comprehensive measure of IH in a format that would be more versatile and efficient to employ.

We were able to locate only two IH scales: a self-report measure specific to religious IH and an informant-report measure of general IH. Hopkin et al. (2014) wrote items about religious beliefs that were face valid in expressing IH that resulted in the following four factors: awareness of fallibility of beliefs (e.g., "When it comes to religious or spiritual beliefs, mine are more accurate than others," reverse scored), discretion in asserting beliefs (e.g., "Even when I have a strong religious or spiritual belief I don't need everyone to know it"), comfort in keeping beliefs private (e.g., "It's important to share my religious or spiritual views with others regardless of whether they agree with me," reverse scored), and respect for others' belief (e.g., "I listen to others' religious or spiritual beliefs without disagreeing even when I think I am right"), with internal consistency of the scores ranging from .71 to .89. Thus, this measure offers a relatively broad conceptualization of religious IH. The two domains of this scale that seem most central to common themes in the literature on general IH are awareness of the fallibility of one's beliefs and respect for others' beliefs. However, given that this scale was developed to assess religious IH with items referring specifically to religious and spiritual beliefs, it was not designed for use as a measure of general IH.

The only scale of general IH we could locate was an informant-report measure developed by McElroy et al. (2014). The items assess perceptions of a target's IH based on the theory that IH is fundamentally relational in nature, involving regulating interactions with others on topics related to one's beliefs and worldview. McElroy et al. (2014) proposed that IH is especially pertinent any time there is a competition or negotiation of ideas in a relationship or group. As such, the items primarily describe observable behaviors thought to be reflective of IH or arrogance. The scale has two factors, intellectual openness (7 items; e.g., "Is good at mediating controversial topics") and intellectual arrogance (9 reverse-scored items; e.g., "Always has to have the last word in an argument"). A confirmatory factor analysis indicated good fit (Comparative Fit Index [CFI] = .98, standardized root mean residual [SRMR] = .04, root mean square error of approximation [RMSEA] = .03), and the scale scores displayed high internal consistency (.92–.94 for intellectual openness and .93 for intellectual arrogance). No test-retest reliability was reported.

Given that IH is a growing area of interest in the field of psychology, we hope to contribute to the research literature by providing a broad measure of IH that is self-report in nature. Self-report allows researchers to gather data from target individuals themselves, without requiring informant reporters or observational periods. This would supplement informant-report methodology by offering more accurate assessment of the intrapsychic feelings, attitudes, and beliefs underlying IH, in addition to assessing behavioral components of IH. Of course, there are concerns regarding the use of self-report measures to assess humility constructs, as individuals might lack awareness or be biased in responding (Bollinger & Hill, 2012). Nevertheless, there are also empirical indications that self-report can be a useful way to measure IH. For example, Rowatt et al. (2006) found that self-reported humility offered similar information to an implicit assessment of humility, and Landrum (2011) found that self-report measures of humility were not correlated with social desirability.

Present research and hypotheses

Although current empirical knowledge of IH is very limited, the construct carries potential implications for the thriving of individuals, relationships, and society. Presumably, IH is integral to interpersonal interactions on a microlevel as well as peaceful interactions within pluralistic societies and sociopolitical interactions throughout the world. Clearly, there is a need for empirical research to validate these assumptions and examine specifically how IH can contribute to individuals, relationships, communities, and society. A prerequisite of such empirical research is the availability of scales that provide reliable and valid measurement of IH. This research was conducted to develop and validate a self-report measure of IH. Studies 1 and

2 aimed to create a comprehensive, theoretically grounded measure of IH. Studies 3 through 5 were designed to examine hypotheses related to convergent, discriminant, construct, and incremental validity and test-retest reliability.

Convergent validity

We expected that the IH measure would correlate positively with an existing measure of the same construct. In addition, we anticipated that humility specific to one's thought processes (IH) would also correlate positively with the related, but broader construct of general humility. This resulted in Hypothesis 1, that the IH scale would correlate with measures of intellectual and general humility, thereby supporting the scale's convergent validity (Study 3).

Construct validity

Furthermore, we sought to demonstrate that the new scale would be predictive of the primary outcomes of IH, being that humility in the way one acquires and maintains knowledge would result in greater openness and respect for other people, ideas, and experiences (e.g., McElroy et al., 2014). This resulted in Hypothesis 2, that the IH scale would correlate with measures of open-mindedness, tolerance, and openness to experience (Studies 3 and 4), thereby supporting the scale's construct validity.

Discriminant validity

Although there is controversy in the field regarding the use of self-report methodology to assess humility constructs, there are some indications in previous research that self-reported humility is not correlated with social desirability (Landrum, 2011) and offers similar information to implicit assessments of humility (Rowatt et al., 2006). A goal of this research was to examine whether IH could successfully be measured through self-report, while not being overly influenced by social desirability tendencies.

In addition, although there is a tendency to confuse humility with low self-regard, self-abasement, or low self-esteem (Tangney, 2000), research has shown that humble individuals can have positive opinions of themselves (Exline et al., 2004). Building on this, theories of IH propose that IH is distinct from intellectual diffidence and does not equate to being self-effacing, lacking self-confidence, or yielding too quickly to others (Vorobej, 2011; Zagzebski, 1996). On these bases, our Hypothesis 3 is that the IH scale would demonstrate no to small correlations with social desirability, low self-regard, social conformity, and confidence, thereby supporting the scale's discriminant validity (Studies 3 and 4).

Incremental validity

A valuable method of validating a new scale involves demonstrating that it accounts for incremental variance in predicting salient criterion scores beyond the variance attributable to existing measures. Given that IH is a fairly new construct in the social sciences, there are no clearly identified benchmarks against which to evaluate it. We started by comparing the new IH scale to a recently developed measure of IH. McElroy et al.'s (2014) informant-report Intellectual Humility Scale (IHS) was under review at the time we collected our data. We rewrote these items to reflect a self-report format, expecting that our scale would positively correlate with their Intellectual Openness subscale and negatively correlate with their Intellectual Arrogance subscale (see Hypothesis 1), but that the broader content coverage of the new IH scale would offer greater predictive validity compared to the existing measure of IH in predicting open-mindedness, a salient outcome of IH that is also central to the construct of intellectual openness measured by the IHS (Study 3). This is our Hypothesis 4.

Next, we wanted to demonstrate how IH is distinct from humility, employing the two most frequently used self-report measures of humility (Davis et al., 2010), the Honesty-Humility subscale of the HEXACO Personality Inventory (using the two humility scales to increase specificity) and the Modesty-Humility subscale of the Values in Action Strengths Inventory. We theorized that IH and humility would differentially predict certain outcomes. Specifically, given that IH involves having insight about the limits of one's knowledge, it is marked by openness to new ideas (McElroy et al., 2014), whereas general humility involves a broader perception of self, and low scores have often been shown to predict qualities such as narcissism and psychological entitlement (Exline et al, 2004). Therefore, we offer Hypothesis 5, that the IH scale would predict open-mindedness above measures of humility, and Hypothesis 6, that low scores on measures of humility would predict narcissism and psychological entitlement more so than low scores on the IH scale (Study 3).

We also desired to demonstrate how IH relates to well-known personality traits. Given that IH is marked by openness to new ideas (McElroy et al., 2014), we proposed Hypothesis 7: The IH scale would predict the personality trait of openness to experience, even beyond the personality trait of individualism, which assesses the construct of unpretentiousness (Study 4).

Further, because IH involves reflection on and awareness about one's thought processes, it is likely to be associated with greater intellectual complexity and curiosity (Samuelson, Church, Jarvinen, & Paulus, 2012). This is a potential confound that has not frequently been addressed in the literature on IH. We wanted to demonstrate that outcomes of IH would not be attributable simply to tendencies toward intellectual engagement. This led to Hypothesis 8: The IH scale would predict key outcomes of IH, specifically open-mindedness and tolerance, above the tendency to desire understanding and engage in critical thinking (Study 4).

Test-Retest reliability

Finally, we conceptualized IH as a dispositional quality and therefore expected individuals to fluctuate around dispositional levels of IH on the basis of situational factors. Therefore Hypothesis 9 was that test-retest analyses would reveal moderately stable IH scores with correlations around .70 or higher (Study 5).

General procedures

All studies were conducted after receiving institutional review board approval and all participants provided informed

consent. Data were collected via online surveys. Items of the Comprehensive Intellectual Humility Scale (CIHS) developed in this study were presented to participants in randomized order. In Studies 3 and 4, participants completed the CIHS first and then completed a battery of other surveys in randomized order.

Data were deleted listwise for participants who completed the survey unreasonably fast (less than an average of 2.5 sec per question), were multivariate outliers with a detectible, unreasonable pattern of responding, or responded incorrectly to factual questions embedded in the survey, indicating a lack of attention. Prior to analyses, items were reverse coded where necessary, so that higher values indicated greater levels of each construct assessed. Confirmatory factor analysis was conducted using Amos 22.0; all other analyses were conducted using SPSS Statistics 22.0.

Study 1: Item selection and factor structure

Item development and content validity

On the basis of extensive literature review, we defined IH as a nonthreatening awareness of one's intellectual fallibility. Such an awareness brings with it a healthy independence between one's intellect and ego, meaning that a person will not feel threatened by intellectual disagreements, will not be overconfident about his or her knowledge, will respect the viewpoints of others, and will be open to revising his or her viewpoints. As such, we conceptualized IH as both an intrapersonal and interpersonal construct.

We conducted two focus groups to examine how individuals naturally speak about IH. The first consisted of six individuals in academia (two faculty members, two graduate students, and two undergraduate students); the second consisted of eight individuals in the general population (ranging in age from 21–62 years). The individuals were asked to offer an explanation of the meaning of IH in their own words on the basis of a definition of IH provided to them. Subsequently, they were asked to describe individuals they considered high and low in IH. Their responses were used to generate words, phrases, and potential scale items consistent with the language used by those in the general population.

The literature review and focus groups resulted in 187 potential items for measuring IH. Eighteen expert reviewers provided feedback on our description of IH and item pool. These individuals had expertise in IH, general humility, relational humility, spiritual humility, cultural humility, humility as a personality dimension, humility as a virtue, and virtue and character education. All experts held terminal degrees (16 PhD and 2 EdD) in psychology ($n = 11$), philosophy ($n = 5$), theology ($n = 1$), or education ($n = 1$), and most were professors and researchers. Seven of the experts had published one or more humility scales. The experts rated each item on a 4-point Likert scale for face validity, construct fidelity, and clarity of wording. They also provided open-ended feedback about our definition of IH and individual scale items. We revised items based on this feedback and retained 118 items on the basis of analyses for item retention (V statistic; Aiken, 1985). Only items consistently rated highly by reviewers were retained, corresponding to an average rating of at least a 3 (*good fit*) or 4 (*excellent fit*) on a 4-point scale for construct fidelity.

Pilot testing

We conducted three pilot tests among individuals who had no known previous exposure to information about IH, who completed the measure and marked items that were confusing in any way, with the option of explaining what was confusing. Items were reworded or deleted based on the first pilot test of 139 college students, resulting in the retention of 76 items. A second pilot test among 167 adults in the general population indicated that three additional items were confusing to multiple participants, resulting in the deletion of these items. A third pilot test among 129 college students confirmed that the remaining 73 items were clear and understandable to each participant.

Participants

Participants throughout the United States were recruited through Amazon Mechanical Turk (Mturk). The data set was divided into two random halves ($n = 380$ each) to be used in Study 1 and Study 2. Study 1 participants consisted of 56% men and 44% women between 18 and 71 years old ($M = 33.41$, $SD = 12.01$). The sample identified as 77% White, 9% Asian, 8% Hispanic, 4% Black or African American, and 2% other. In terms of highest level of education completed, 36% had completed high school, 56% had completed college, and 8% had completed graduate school.

Measure

The 73 items retained from pilot testing were administered. The items assessed both intrapersonal and interpersonal forms of IH, including thoughts and behaviors, and were inclusive of positively and negatively worded phrases. All items were rated on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*.

Procedure

Participants were paid through Mturk for their participation, on the basis of the length of the survey. We conducted a maximum likelihood principal-axis factor analysis (PAF) of the 73 scale items retained from pilot testing.

Results

Predicting that aspects of IH would be intercorrelated, we made use of oblique rotation (Promax, kappa = 4). The Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) was .93 and Bartlett's test of sphericity was significant, $\chi^2(2,628) = 13532.36$, $p < .001$.

Factor extraction was guided by both empirical methods and theory. We examined the Kaiser criterion, the Cattell scree test plot, and parallel analysis. We compared the resulting options on the basis of a theoretical interpretation of the extent to which the factors were consistent with our definition of IH. On

the basis of the combined empirical-theoretical approach, the clearest solution resulted from four factors. This model provided the richest content coverage of the construct of IH while avoiding redundancy. The eigenvalue of the fifth factor (unretained) was 1.93. The four factors accounted for 41.05% of the variance.

We implemented a number of a priori decision criteria in selecting items. To achieve a relatively brief measure, we selected up to six items per factor with the highest factor pattern and structure coefficients (minimum .40) and deleted items with factor pattern cross-loadings greater than .30. Factor structure cross-loadings exceeded .30 in some instances; however, we selected items with the greatest distance between the two largest factor structure coefficients.

A principal axis factoring of the 22 selected items had a KMO of .89 and a significant test of sphericity, $\chi^2(231) = 3458.65$, $p < .001$. All items continued to load on their respective factors. The four factors accounted for 57.24% of the total variance and 48.25% of the common variance. Table 1 displays the associated factor pattern matrix, the variance explained by each factor, and the correlations between factors (the factor structure matrix is available on request). As expected, correlations were present among the factors, ranging from .31 to .67.

A readability analysis of the selected 22 items resulted in a Flesch-Kincaid Reading Ease score of 76.1 and a Flesch-Kincaid grade level of 3.9 (Kincaid, Fishburne, Rogers, & Chissom, 1975). Table 2 summarizes the psychometric properties of the final scale for all studies. In Study 1, coefficient alpha for the full scale was .88, and for the subscales it ranged from .73 to .89.

Discussion

The goal of Study 1 was to create a self-report scale of IH. This resulted in a 22-item, four-factor model of IH. Four factors provided a robust assessment of IH that minimized ambiguity between factor content. The exploratory factor analytic solution was well above our a priori established KMO measure of sampling adequacy statistic of .80. The findings indicated that the full scale and four factors had satisfactory internal consistency, each meeting our a priori goal of a coefficient alpha of .70 or higher. As expected, the factors were intercorrelated with one another.

Study 2: Confirmatory factor analysis

Participants

Study 2 participants ($N = 380$) consisted of 59% men and 41% women between 18 and 72 years old ($M = 32.87$, $SD = 11.54$). The sample identified as 77% White, 8% Asian, 5% Hispanic, 5% Black or African American, 3% multiracial, and 2% other. In terms of highest level of education completed, 33% had completed high school, 55% had completed college, and 12% had completed graduate school.

Measure

Study 2 used the same items described in Study 1. Although all 73 items retained from the pilot testing were administered,

Table 1. Factor pattern of principal-axis factor analysis of retained items with Promax rotation.

Items	Factor 1	Factor 2	Factor 3	Factor 4	h^2
73. I feel small when others disagree with me on topics that are close to my heart.[a]	**0.82**	0.12	−0.09	−0.09	0.65
50. When someone contradicts my most important beliefs, it feels like a personal attack.[a]	**0.80**	−0.07	0.01	0.12	0.69
49. When someone disagrees with ideas that are important to me, it feels as though I'm being attacked.[a]	**0.78**	0.01	0.04	0.04	0.67
53. I tend to feel threatened when others disagree with me on topics that are close to my heart.[a]	**0.77**	−0.12	0.07	0.08	0.61
68. When someone disagrees with ideas that are important to me, it makes me feel insignificant.[a]	**0.75**	0.10	−0.03	−0.12	0.56
28. I am open to revising my important beliefs in the face of new information.	0.03	**0.77**	−0.04	−0.01	0.56
26. I am willing to change my position on an important issue in the face of good reasons.	0.02	**0.73**	−0.01	0.00	0.53
29. I am willing to change my opinions on the basis of compelling reason.	0.08	**0.69**	−0.02	−0.04	0.49
25. I have at times changed opinions that were important to me, when someone showed me I was wrong.	−0.07	**0.61**	0.06	0.04	0.41
33. I'm willing to change my mind once it's made up about an important topic.	−0.01	**0.49**	0.05	0.14	0.33
61. I can respect others, even if I disagree with them in important ways.	0.02	−0.19	**0.84**	0.01	0.56
47. I can have great respect for someone, even when we don't see eye-to-eye on important topics.	0.02	−0.02	**0.82**	−0.08	0.62
39. Even when I disagree with others, I can recognize that they have sound points.	−0.05	0.18	**0.64**	−0.05	0.54
65. I am willing to hear others out, even if I disagree with them.	0.06	0.21	**0.56**	−0.09	0.51
45. I welcome different ways of thinking about important topics.	−0.02	0.25	**0.47**	0.10	0.50
34. I respect that there are ways of making important decisions that are different from the way I make decisions.	−0.06	0.19	**0.46**	0.10	0.39
1. My ideas are usually better than other people's ideas.[a]	0.02	−0.11	−0.02	**0.68**	0.42
3. For the most part, others have more to learn from me than I have to learn from them.[a]	−0.09	0.08	−0.06	**0.66**	0.41
9. When I am really confident in a belief, there is very little chance that belief is wrong.[a]	0.01	−0.05	0.15	**0.59**	0.43
24. On important topics, I am not likely to be swayed by the viewpoints of others.[a]	0.06	0.06	−0.06	**0.50**	0.26
21. I'd rather rely on my own knowledge about most topics than turn to others for expertise.[a]	−0.04	0.14	−0.06	**0.46**	0.24
41. Listening to perspectives of others seldom changes my important opinions.[a]	0.09	0.02	0.00	**0.46**	0.29
Initial eigenvalue	6.81	2.45	2.11	1.23	
Initial % variance	30.94	11.15	9.58	5.58	
Extraction % variance	28.76	9.27	6.83	3.38	
Rotation sum of squared loadings	4.28	4.70	4.95	3.11	
Correlation with Factor 1	1.00				
Correlation with Factor 2	0.38	1.00			
Correlation with Factor 3	0.41	0.67	1.00		
Correlation with Factor 4	0.33	0.31	0.40	1.00	

Note. $N = 308$. h^2 = communality coefficient. Factor labels are as follows: Factor 1, independence of intellect and ego; Factor 2, openness to revising one's viewpoint; Factor 3, respect for others' viewpoints; Factor 4, lack of intellectual overconfidence. Bold formatting represents factor loadings of items.
[a]Reverse-coded items.

Table 2. Psychometric properties of the Comprehensive Intellectual Humility Scale and subscales.

	Full scale (22 items; possible range: 22–110)			Subscale 1: Independence ofintellect and ego (5 items; possible range: 5–25)			Subscale 2: Openness to revising one's viewpoint (5 items;possible range: 5–25)			Subscale 3:Respect for others' viewpoints (6 items; possible range: 6–30)			Subscale 4:Lack of intellectual overconfidence (6 items; possible range: 6 –30)		
Sample	*M*(*SD*)	Range	α	*M*(*SD*)	Range	α	*M*(*SD*)	Range	α	*M*(*SD*)	Range	α	*M*(*SD*)	Range	α
Sample 1($N = 380$)	80.20 (9.75)	52–110	0.88	16.85 (4.64)	5–25	0.89	20.00 (2.32)	11–25	0.80	24.42 (2.84)	13–30	0.85	18.92 (3.58)	11–30	0.73
Sample 2 ($N = 380$)	81.89 (9.57)	59–108	0.87	17.78 (4.46)	5–25	0.89	20.31 (2.38)	12–25	0.81	24.78 (2.83)	14–30	0.84	19.03 (3.74)	7–29	0.74
Sample 3 ($N = 509$)	82.01 (9.21)	47–107	0.85	17.57 (4.39)	5–25	0.87	20.13 (2.47)	10–25	0.80	24.97 (2.77)	11–30	0.83	19.34 (3.65)	6–30	0.72
Sample 4($N = 179$)	79.84 (8.64)	56–103	0.82	16.27 (3.99)	5–25	0.82	19.20 (2.60)	11–25	0.76	24.77 (2.80)	14–30	0.78	19.61 (3.57)	6–28	0.70

only the 22 items selected in Study 1 were entered into each confirmatory factor analysis (CFA) in Study 2.

Procedure

We used structural equation modeling based on maximum-likelihood estimation to evaluate three separate CFA models. The first was the four-factor model identified in Study 1. Because positive correlations were observed among the four factors in Study 1, our second model added a higher order factor, representing the common variance among the four lower level facets. Third, we tested a one-factor model by fixing the covariances between the latent factors to a value of one; doing so nested this model within the four-factor model, allowing for a direct statistical comparison across models.

Results

The analysis of the four-factor model resulted in model-fit indexes within ranges considered indicative of strong fit (Byrne, 2010; Kline, 2011): $\chi^2(203) = 367.101$, $p < .001$; goodness-of-fit index (GFI) = .921, CFI = .948, and RMSEA = .046, 90% CI [.039, .054], $p = .793$.

The four-factor model with a higher order factor yielded strong fit indexes that were extremely similar to those from the four-factor model, $\chi^2(205) = 368.291$, $p < .001$; GFI = .921, CFI = .948, and RMSEA = .046; 90% CI [.038, .053], $p = .793$. The difference between the χ^2 values was not statistically significant, $\chi^2\mathbf{\Delta}(2) = 1.19$, $p > .05$.

However, the one-factor model did not yield strong indexes: $\chi^2(209) = 1496.591$, $p < .001$, GFI = .657, CFI = .594, and RMSEA = .127; 90% CI [.121, .134], $p < .001$. The difference between this model's χ^2 value and that of the four-factor model was statistically significant, $\chi^2\mathbf{\Delta}(6) = 1129.49$, $p < .001$.

Although the addition of a higher order factor only resulted in a negligible, nonsignificant improvement in model fit over the four-factor model, the higher order model was selected for all subsequent analyses because this model was consistent with the anticipated use of the measure as a general scale with subscales. This model and its parameter estimates are provided in Figure 1. Among Study 2 participants, coefficient alpha was .87 for the full scale and ranged from .74 to .89 for the subscales.

Discussion

The second study was designed to determine whether the factor structure obtained in the first study could be confirmed in an independent sample. Because the first study used an oblique rotation method and obtained factors that were intercorrelated, a higher order factor structure was explored, and the data showed strong fit to this model. Furthermore, reliability estimates obtained for this sample were comparable to those in the first sample. These analyses bolster confidence in the psychometric strength and factor structure of the 22-item, four-factor CIHS, supporting the use of both the full scale and the subscales for additional validation analyses.

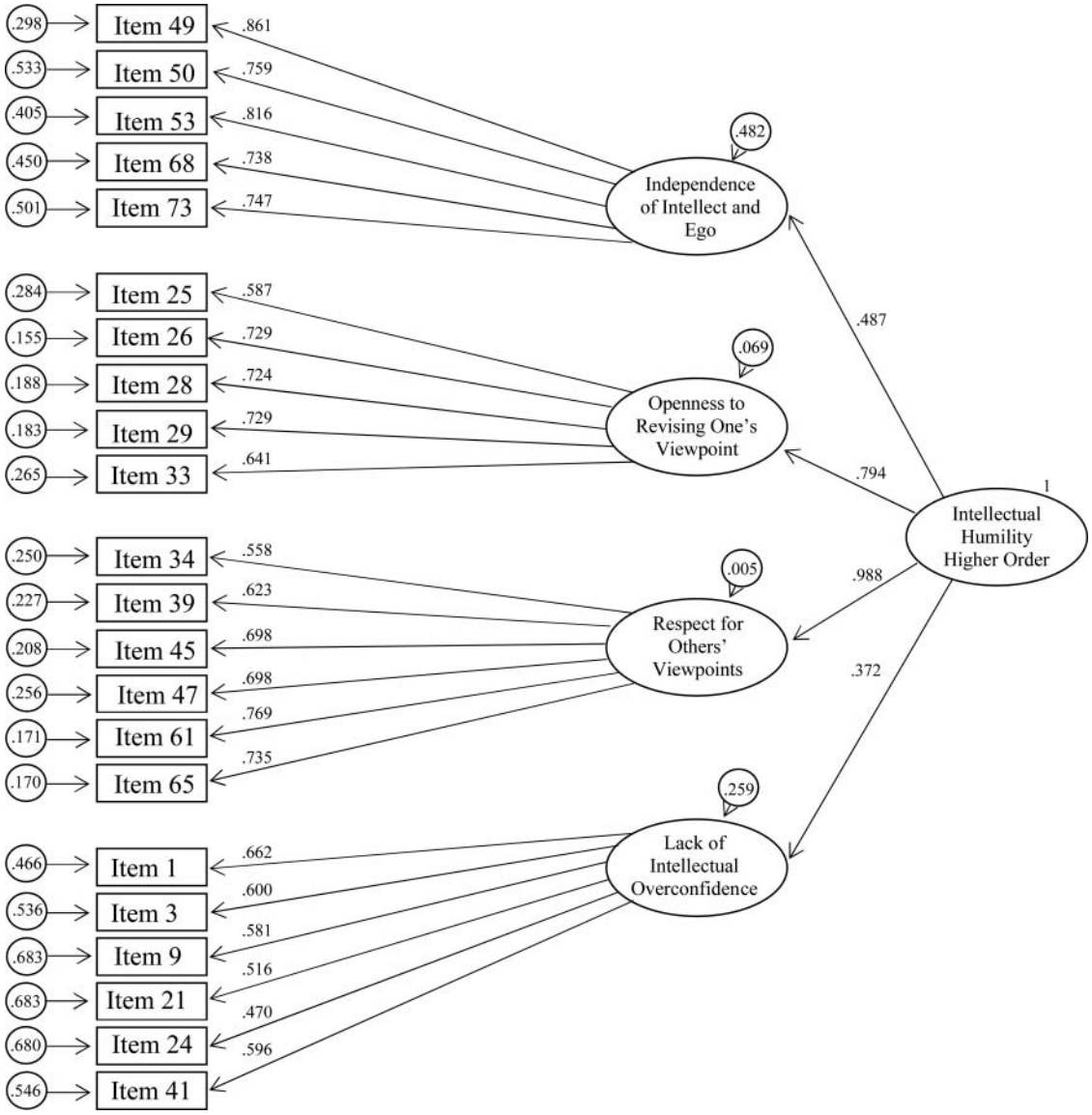


Figure 1. Model and parameter estimates for a four-factor model with higher order factor.

only the 22 items selected in Study 1 were entered into each confirmatory factor analysis (CFA) in Study 2.

Procedure

We used structural equation modeling based on maximum-likelihood estimation to evaluate three separate CFA models. The first was the four-factor model identified in Study 1. Because positive correlations were observed among the four factors in Study 1, our second model added a higher order factor, representing the common variance among the four lower level facets. Third, we tested a one-factor model by fixing the covariances between the latent factors to a value of one; doing so nested this model within the four-factor model, allowing for a direct statistical comparison across models.

Results

The analysis of the four-factor model resulted in model-fit indexes within ranges considered indicative of strong fit (Byrne, 2010; Kline, 2011): $\chi^2(203) = 367.101$, $p < .001$; goodness-of-fit index (GFI) = .921, CFI = .948, and RMSEA = .046, 90% CI [.039, .054], $p = .793$.

The four-factor model with a higher order factor yielded strong fit indexes that were extremely similar to those from the four-factor model, $\chi^2(205) = 368.291$, $p < .001$; GFI = .921, CFI = .948, and RMSEA = .046; 90% CI [.038, .053], $p = .793$. The difference between the χ^2 values was not statistically significant, $\chi^2\mathbf{\Delta}(2) = 1.19$, $p > .05$.

However, the one-factor model did not yield strong indexes: $\chi^2(209) = 1496.591$, $p < .001$, GFI = .657, CFI = .594, and RMSEA = .127; 90% CI [.121, .134], $p < .001$. The difference between this model's χ^2 value and that of the four-factor model was statistically significant, $\chi^2\mathbf{\Delta}(6) = 1129.49$, $p < .001$.

Although the addition of a higher order factor only resulted in a negligible, nonsignificant improvement in model fit over the four-factor model, the higher order model was selected for all subsequent analyses because this model was consistent with the anticipated use of the measure as a general scale with subscales. This model and its parameter estimates are provided in

Figure 1. Among Study 2 participants, coefficient alpha was .87 for the full scale and ranged from .74 to .89 for the subscales.

Discussion

The second study was designed to determine whether the factor structure obtained in the first study could be confirmed in an independent sample. Because the first study used an oblique rotation method and obtained factors that were intercorrelated, a higher order factor structure was explored, and the data showed strong fit to this model. Furthermore, reliability estimates obtained for this sample were comparable to those in the first sample. These analyses bolster confidence in the psychometric strength and factor structure of the 22-item, four-factor CIHS, supporting the use of both the full scale and the subscales for additional validation analyses.

Study 3: Convergent, discriminant, and incremental validity

Participants

Study 3 made use of a new sample of U.S. participants recruited through Mturk ($N = 509$), 52% women and 47% men between 18 and 81 years old ($M = 34.43$, $SD = 11.41$). The sample identified as 76% White, 7% multiracial, 6% Asian, 6% African American or Black, 4% Hispanic, and less than 2% other. In terms of highest level of education completed, 31% had completed high school, 57% had completed college, and 12% had completed graduate school.

Measures

Comprehensive Intellectual Humility Scale

The 22-item CIHS developed in Study 1 was used to assess IH. The items, displayed in Table 1, were rated on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Items were summed for each of the four subscales and the full scale, with higher scores indicating greater IH.

HEXACO, Humility subscale

The Humility subscale of the HEXACO (Ashton & Lee, 2008) was used to assess modesty with 4 items ($\alpha = .70$) and lack of greed with 4 items ($\alpha = .80$). Internal consistency of the full 8-item subscale was .80. Items were rated on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*, and summed into scale scores with higher scores indicating greater humility.

Humility

The 9-item Humility subscale of the International Personality Item Pool's (IPIP) version of the Values in Action Inventory of Strengths (Goldberg et al., 2006) was used to assess participants' report of humble behavior and being viewed by others as humble. Items were rated on a 5-point Likert scale ranging from 1 = *very much unlike me* to 5 = *very much like me*, and summed such that higher scores indicated greater amounts of humility. The internal consistency was acceptable ($\alpha = .78$).

Dispositional Humility Scale

The Dispositional Humility Scale (Landrum, 2011) offers an alternative to self-report assessment of humility by which participants are asked to rate the extent to which they like people with various characteristics on the assumption that humble people like other humble people. The scale is composed of filler items along with two scales measuring humility and four scales measuring nonhumility qualities. The humility subscales assess being self-correcting (i.e., admit their mistakes and weaknesses and are open-minded; $\alpha = .90$) and accurate perception of self (i.e., are aware of their limitations and keeping their accomplishments in perspective; $\alpha = .83$). In addition, we made use of scales assessing low self-regard (i.e., being socially weak, shy, and lacking in self-esteem; $\alpha = .87$) and self-confidence ($\alpha = .74$). Responses were rated on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*, and summed into scale scores with higher scores indicating greater levels of each quality.

Intellectual Humility Scale

The Intellectual Humility Scale (IHS; McElroy et al., 2014) is an informant-report measure of IH consisting of two subscales: Intellectual Openness (7 items; $\alpha = .87$) and Intellectual Arrogance (9 items; $\alpha = .88$). We adapted this measure to be administered in a self-report format. The items were rated on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*, and summed into scale scores with higher scores indicating greater intellectual openness and intellectual arrogance, respectively.

Composite Actively Open-Minded Thinking Scale

The Composite Actively Open-Minded Thinking Scale (Stanovich & West, 2007) is a comprehensive, 41-item measure used to assess open-minded and flexible thinking ($\alpha = .92$). Items were rated on a 6-point Likert scale ranging from 1 = *disagree strongly* to 6 = *agree strongly* and summed such that higher scores indicated greater open-mindedness.

Narcissistic Personality Inventory–13

We used the 13-item version of the Narcissistic Personality Inventory (NPI–13; Gentile et al., 2013) to assess narcissism. It includes three subscales that assess sense of entitlement and tendency to manipulate others (Entitlement/Exploitativeness, $\alpha = .55$), arrogance and exhibitionism (Grandiose Exhibitionism; $\alpha = .76$), and desire to have power over others (Leadership/Authority; $\alpha = .74$). Internal consistency for the full 13-item scale was .81. The format consists of paired choice responses, and higher scale scores indicate greater narcissistic qualities.

Psychological Entitlement Scale

The 9-item Psychological Entitlement Scale (Campbell, Bonacci, Shelton, Exline, & Bushman, 2004) was used to assess the pervasive belief that one deserves and is entitled to more than others ($\alpha = .90$). The items were rated on a 7-point Likert scale ranging from 1 = *strongly disagree* to 7 = *strongly agree*, and summed such that higher scores indicate that respondents view themselves as more worthy than others.

Social Desirability Scale

The Social Desirability Scale, Form C (Crowne & Marlowe, 1960), was used to assess the tendency to try to appear to behave in ways viewed favorably by others ($\alpha = .81$). The 13 items had a true-false format, with higher scores indicating a greater desire to appear to behave in socially favorable ways.

Procedure

Participants were recruited through Mturk and paid for participating commensurate with the length of the survey. We used two-tailed Pearson correlations to examine the convergent and discriminant validity of the CIHS. We conducted hierarchical regressions to examine incremental validity of the CIHS.

Results

Internal consistency

Table 2 displays the psychometric properties of the CIHS within Sample 3. The internal consistency of the full scale was strong (.85) and was acceptable to strong for each subscale (ranging from .72–.87).

Correlations

The CIHS displayed a small positive correlation with age ($r = .09$, $p < .05$). Therefore, age was controlled in Study 3 analyses. The full scale was not correlated with any other demographic factors, including gender, race, education, or religious affiliation.

Consistent with Hypothesis 1 the CIHS demonstrated moderate correlations with the IHS ($r = .52$, $p < .001$, for Intellectual Openness; $r = -.53$, $p < .001$, for Intellectual Arrogance), and small to moderate correlations with measures of humility ($r = .23$, $p < .001$, with the HEXACO Humility subscale; $r = .21$, $p < .001$, with the Humility subscale of the IPIP Values in Action Scales; $r = .42$, $p < .001$, with Landrum's Self-Correction Humility subscale; and $r = .30$, $p < .001$, for Landrum's Accurate Self-Perspective Humility subscale). Consistent with Hypothesis 2, the CIHS was moderately correlated with open-minded thinking ($r = .56$, $p < .001$). Consistent with Hypothesis 3, the CIHS was not redundant with social desirability; nevertheless, there was a small, positive correlation between the two ($r = .22$, $p < .001$), which was driven by the first and second factors of the CIHS (independence of intellect and ego and openness to revising one's viewpoints). Also consistent with Hypothesis 3, the CIHS was unrelated to Landrum's Low Self-Regard subscale ($r = -.04$, $p = .32$), and even showed a small, positive correlation with Landrum's Self-Confidence subscale ($r = .13$, $p < .01$).

Hierarchical regressions

Table 3 displays the results of four hierarchical regressions. Consistent with Hypothesis 4, Panel A displays that the CIHS (entered in Step 3) predicted variance in open-minded thinking beyond the variability attributable to age and social desirability (entered in Step 1), and a self-report assessment of an existing informant report IH scale (IHS; entered in Step 2). The CIHS accounted for 12.4% of the variance in open-minded thinking beyond age, social desirability, and the IHS.

Consistent with Hypothesis 5, Panel B displays that the CIHS (entered in Step 3) predicted variance in open-minded thinking beyond the variability attributable to age and social desirability (entered in Step 1), and three measures of humility (entered in Step 2). The CIHS accounted for 26.2% of the variance in open-minded thinking beyond age, social desirability, and general humility.

Consistent with Hypothesis 6, Panel C displays that measures of humility (entered in Step 3) predicted variance in narcissism and psychological entitlement beyond the variability attributable

Table 3. Hierarchical regression models demonstrating incremental validity.

CIHS predicting open-mindedness beyond demographic factors, social desirability, and the intellectual humility scale

	Actively Open-Minded Thinking		
	B (*SE*)	β	ΔR^2
Step 1			.004
Age	.15 (.10)	.06	
Social desirability	−.01 (.34)	−.00	
Step 2			.220***
Intellectual openness	1.95 (.25)	.33***	
Intellectual arrogance	−1.21 (.20)	−.29***	
Step 3			.124***
CIHS	1.29 (.13)	0.46***	
Total R^2			.348

CIHS predicting open-mindedness beyond demographic factors, social desirability, and commonly used humility measures

	Actively Open-Minded Thinking		
	B (*SE*)	β	ΔR^2
Step 1			.004
Age	.15 (.10)	.06	
Social desirability	−.01 (.34)	−.00	
Step 2			.087***
HEXACO Greed avoidance	.39 (.34)	.06	
HEXACO Modesty	.28 (.51)	.07***	
IPIP VIA Humility scale	−.46 (.25)	−.10	
Step 3			.262***
CIHS	1.53 (.11)	0.54***	
Total R^2			.353

Commonly used humility measures predicting narcissism and psychological entitlement beyond demographic factors, social desirability, and the CIHS

	Narcissism			Psychological entitlement		
	B (*SE*)	β	ΔR^2	*B* (*SE*)	β	ΔR^2
Step 1			.059***			.002
Age	−.05 (.01)	−.20***		−.03 (.04)	−.03	
Social desirability	−.10 (.04)	−.12**		−.10 (.14)	−.03	
Step 2			.020**			.048***
CIHS	−.04 (.01)	−.14**		−.26 (.05)	−.22***	
Step 3			.346***			.308***
HEXACO Greed avoidance	−.07 (.03)	−.09*		−.34 (.12)	−.12**	
HEXACO Modesty	−.30 (.04)	−.31***		−1.76 (.18)	−.48***	
IPIP VIA Humility scale	−.17 (.21)	−.34***		−.14 (.08)	−.08	
Total R^2			.425			.358

Note. $N = 509$. CIHS = Comprehensive Intellectual Humility Scale; IPIP VIA = International Personality Item Pool's version of the Values in Action Inventory of Strengths.
*$p < .05$. **$p < .01$. ***$p < .001$.

to age and social desirability (entered in Step 1), and the CIHS (entered in Step 2). The measures of humility accounted for 34.6% of the variance in narcissism and 30.8% of the variability in psychological entitlement beyond age, social desirability, and IH.

Discussion

The purpose of Study 3 was to continue to assess the internal consistency of the CIHS and to evaluate the validity of the CIHS. The internal consistency of the scale supports the use of the CIHS in future research. In addition, the analyses provided evidence of appropriate levels of construct, convergent, and discriminant validity, with anticipated positive links between the CIHS and an existing measure of IH as well as measures of similar constructs such as general humility and open-minded thinking. Importantly, IH was not equivalent to a lack of self-confidence or low self-regard. This refutes the popular misconception that humility somehow represents a negative view of self and supports the idea that one can be humble about one's beliefs and values without lacking confidence or taking a position of diffidence.

The CIHS displayed only small associations with social desirability and was able to predict outcome measures when controlling levels of social desirability. The CIHS displayed incremental predictive value in open-minded thinking beyond an existing measure of IH (adapted from informant report to self-report format) and commonly used measures of humility. This indicates that the CIHS seems to be a valuable addition to the IHS, perhaps due to the broader content coverage. Further, IH seems to be a unique construct from general humility, which in turn is better at predicting outcomes such as narcissism and psychological entitlement, which are more closely aligned with definitions of general humility.

Study 4: Convergent, discriminant, and incremental validity

Participants

The sample consisted of college students ($N = 179$) who were between 18 and 28 years old ($M = 19.02$, $SD = 1.50$). The majority were female (68%), and they were racially diverse: 48% White, 21% Asian, 14% multiracial, 10% Hispanic, 4% African American or Black, and 3% declined to answer.

Measures

In addition to the CIHS, Study 4 made use of two measures previously described in Study 3: the Composite Actively Open-Minded Thinking Scale ($\alpha = 0.88$) and the Social Desirability Scale ($\alpha = 0.64$).

Select scales of the International Personality Item Pool Six Factor Personality Questionnaire

One factor scale and two facet scales of the IPIP version of the Six Factor Personality Questionnaire (Goldberg et al., 2006) were used. The intellectual openness/openness to experience factor was used to assess a preference for new experiences and change, intellectual curiosity, and variety of interests (10 items; $\alpha = .73$). The comprehension/understanding facet was used to assess participants' desire to understand many areas of knowledge, synthesize ideas, and engage in logical thinking (10 items; $\alpha = .81$). The unpretentiousness/individualism facet was used to assess unconcern about reputation, social standing, or praise and approval from others (9 items; $\alpha = .74$). All items were rated on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*, and summed such that higher scores indicated higher levels of each construct.

Select scales of the International Personality Item Pool Jackson Personality Inventory

Three subscales of the IPIP version of the Jackson Personality Inventory (Goldberg et al., 2006) were used. The conformity/cooperativeness scale was used to assess susceptibility to social influence and group pressure, the tendency to modify behavior to standards set by others, and the desire to fit in (9 items; $\alpha = .63$). The Social Confidence scale was used to assess confidence in social interactions and self-confidence (7 items; $\alpha = .76$). The Tolerance scale was used to assess acceptance of diverse people and ideas and being free of prejudice (9 items; $\alpha = .59$). The items consisted of a true–false response format, with higher scale scores indicating higher levels of each construct.

Procedure

Participants were recruited at the affiliated university and received assignment credit in a psychology class for participating. We used two-tailed Pearson correlations to examine the convergent and discriminant validity of the CIHS. We conducted hierarchical regressions to examine whether the CIHS would predict outcome measures beyond the variance attributable to existing measures.

Results

Internal consistency

Table 2 displays the psychometric properties of the CIHS within Sample 4. The internal consistency was strong for the full scale (.82) and was acceptable to strong for each subscale (ranging from .70–.82).

Correlations

None of the assessed demographic factors were correlated with the full-scale CIHS, including age, gender, race, education, or religious affiliation. Consistent with Hypothesis 2, the CIHS was positively correlated with open-minded thinking ($r = .57$, $p < .001$) and tolerance toward other people and ideas ($r = .28$, $p < .001$). Consistent with Hypothesis 3, the scale displayed only a small correlation with social desirability ($r = .15$, $p < .05$) and was unrelated to measures of conformity ($r = -.14$, $p = .07$) and social confidence ($r = .04$, $p = .60$). Finally, consistent with Hypothesis 7, the CIHS was positively correlated with openness to experience ($r = .40$, $p < .001$).

Hierarchical regressions

Table 4 displays the results of three hierarchical regressions conducted to examine whether the CIHS (entered in Step 3) would explain additional variance in outcome measures beyond social desirability (entered in Step 1) and measures of individualism or comprehension (entered in Step 2).

Consistent with Hypothesis 7, the CIHS accounted for 15.8% of openness to experience after accounting for social desirability and individualism. Consistent with Hypothesis 8, the CIHS accounted for 28.6% of open-minded thinking and 5.1% of tolerance, after accounting for social desirability and comprehension.

Discussion

The purpose of Study 4 was to continue to assess the internal consistency and construct validity of the CIHS. The study provided additional support for the use of the CIHS on the basis of the internal consistency. In addition, the study supported appropriate levels of convergent and discriminant validity, with positive links between the CIHS and measures of open-mindedness and tolerance of other people and ideas. Building on the findings of Study 3, IH was not equivalent with the tendency to conform to others or a lack of social confidence. Thus, IH is not inconsistent with being an independent thinker or social leader.

Similar to Study 3, the CIHS displayed only small associations with social desirability and was able to predict outcomes above social desirability. Study 4 provided evidence of convergent validity in the correlation between IH and openness to experience, which remained even after accounting for the personality facet of individualism, which represents unpretentiousness. Finally, Study 4 was able to distinguish the predictive power of IH from any underlying cognitive complexity tendencies, as the CIHS was able to predict open-minded thinking and tolerance, after controlling the degree to which participants desire to understand many areas of knowledge, synthesize ideas, and engage in logical thinking.

Study 5: Test–Retest reliability

Participants

A sample of college students ($N = 137$) and a smaller sample of adults in the general population recruited through Mturk ($N = 40$) participated. The college students were between 18 and 27 years old ($M = 18.99$, $SD = 1.21$), predominately female (73%), and racially diverse: 47% identified as White, 23% as Asian, 14% as multiracial, 9% as Hispanic, 4% as African American or Black, 1% as Pacific Islander, and 2% declined to respond. The community sample ranged in age from 21 to 66 years old ($M = 34.60$, $SD = 11.44$). There were slightly more men (58%) than women (43%). The sample identified as 85% White, 5% Asian, 3% Hispanic, 3% African American or Black, and 3% Pacific Islander. In terms of highest level of education completed, 35% had completed high school, 53% had completed college, and 13% had completed graduate school.

Measure and procedure

The CIHS was used to assess IH. Student participants completed the CIHS twice, 1 month apart. They received assignment credit in a psychology class for their participation. Mturk participants completed the CIHS twice, 3 months apart. They were paid for participating. Pearson correlations were used to examine the temporal reliability of scores on the CIHS for each sample.

Results

The test-retest correlation for the full scale was .75 after 1 month and .70 after 3 months. One-month test-retest for

Table 4. Hierarchical regression models demonstrating incremental validity.

CIHS predicting openness to experience beyond social desirability and individualism			
	IPIP intellectual openness – 6FPQ openness to experience factor		
	B (*SE*)	*β*	ΔR^2
Step 1			.003
Social desirability	.11 (.16)	.05	
Step 2			.001
IPIP unpretentiousness –6FPQ individualism facet	.04 (.10)	.03	
Step 3			.158***
CIHS	.24 (.04)	.41***	
Total R^2			.161

CIHS predicting open-mindedness and tolerance beyond social desirability and understanding						
	Composite Actively Open-Minded Thinking			Tolerance		
	B (*SE*)	*β*	ΔR^2	*B* (*SE*)	*β*	ΔR^2
Step 1			.004			.056**
Social desirability	.50 (.63)	.06		.18 (.06)	.24**	
Step 2			.022			.010
IPIP comprehension – 6FPQ understanding facet	.47 (.24)	.15		.03 (.02)	.10	
Step 3			.286***			.051**
CIHS	1.29 (.16)	.55***		.05 (.02)	.23**	
Total R^2			.311			.117

Note. $N = 179$. CIHS = Comprehensive Intellectual Humility Scale; IPIP = International Personality Item Pool; 6FPQ = Six Factor Personality Questionnaire.
*$p < .05$. **$p < .01$. ***$p < .001$.

Factor 1 was .74, Factor 2 was .59, Factor 3 was .60, and Factor 4 was .46. Three-month test-retest for Factor 1 was .59, Factor 2 was .50, Factor 3 was .76, and Factor 4 was .69. All test-retest coefficients were significant at $p < .001$.

Discussion

Scores for the CIHS were stable over 1- and 3-month periods, which represent relatively long follow-up times for test-retest reliability (Weiner & Greene, 2008). The full scale met Weiner and Greene's (2008) recommended correlation coefficient of .75 as a general standard for short-term test-retest reliability (ranging from 1 day to a few weeks). The longer term test-retest fell below this level, at .70. Test-retest reliability for the subscales fell below this standard for both test-retest periods, supporting the use of the full scale CIHS score.

General discussion

There are many fruitful avenues of exploration regarding IH that have direct psychological, philosophical, and theological implications, yet they can only be pursued with a psychometrically strong measure of IH. The purpose of this research was to develop a self-report measure of IH with strong construct validity that that would be efficient to employ.

Exploratory and confirmatory factor analyses resulted in a 22-item, four-factor model with higher order factor. Therefore, the full-scale CIHS can be used to provide an overall score for IH that is based on an assessment of independence of intellect and ego, openness to revising one's viewpoint, respect for others' viewpoints, and lack of intellectual overconfidence. Where desirable, individual subscale scores can be taken into consideration as well. In terms of complexity, the scale requires only 4 years of education to be readable.

The CIHS and four factors were found to be psychometrically robust and internally consistent across numerous samples drawn from the general population and academic settings, presenting diversity in age, ethnicity, and gender. One- and 3-month test-retest reliabilities for the full scale were moderate to strong, making the scale useful for longitudinal research.

The analyses supported appropriate levels of convergent validity, showing links between the CIHS and measures of IH, general humility, open-mindedness, tolerance, and openness to experience. The scale provided evidence of discriminant validity in a lack of correlations with conceptually distinct constructs, including self-regard, self-confidence, and social conformity. This indicates that being intellectually humble does not mean that one blindly adopts the views of others or lacks confidence in one's own beliefs and values. Thus, there seems to be no conflict in being both confident in and humble about one's viewpoints. This is consistent with the theories of others, who have articulated that IH does not imply spinelessness or submissiveness (Elder & Paul, 2012), and that IH can be conceptualized as a balance between the two vices of intellectual arrogance and intellectual cowardice (Jones, 2012). Future research is needed to examine how the CIHS relates to other constructs, such as agreeableness, which has been linked to general humility.

The CIHS displayed modest, positive correlations with social desirability tendencies, with between 2.25% and 4.80% shared variance between the CIHS and social desirability. Nevertheless, the CIHS was able to predict expected outcomes of IH beyond the variance attributable to social desirability. The scale also provided incremental validity in salient outcome measures beyond measures of IH, general humility, and cognition; thus, the CIHS seems to assess a unique construct and is not redundant with existing measures. This initial research is promising, and future research is needed to continue to validate the CIHS. For example, confirmatory factor analysis can be used to examine how the CIHS relates to existing measures of humility, including measures of general humility, cultural humility, IH, and religious IH.

To our awareness, the CIHS is currently the only self-report measure of general IH. Two scales of IH have previously been published, one specific to religious humility (Hopkin et al., 2014) and one that was validated as an informant-report measure (McElroy et al., 2014). Hopkin et al.'s (2014) measure assesses respondents' humility with regard to religious and spiritual beliefs and viewpoints, whereas the CIHS assesses general IH regarding any kind of beliefs, opinions, and values that are important to the individual. In comparison to McElroy et al.'s IHS, the CIHS might have stronger construct validity by assessing four rather than two domains of IH. The CIHS also takes a complementary approach to the IHS, which is an informant-report measure and, therefore, might be most accurate in assessing observable behaviors. The CIHS is completed by the target individual rather than an observer, and thereby might be more effective at assessing intrapsychic feelings, attitudes, and beliefs that are not necessarily observable by others.

Self-report is a common and advantageous mode of measurement because it elicits firsthand accounts without requiring other reporters or specialized materials. Nevertheless, some have raised questions about the use of self-report methods to assess humility variables, because those low in humility might self-inflate the virtue, whereas those high in humility might display a modesty effect, contributing to decreased variability (Bollinger & Hill, 2012; Davis et al., 2010). The CIHS makes use of neutral language as much as possible and contains items keyed for both socially desirable and undesirable responses. The data displayed sufficient variance and the incremental validity of the scale seems to indicate that individuals were able to provide information about their levels of IH without being inhibited by social desirability tendencies. We believe future research will benefit from evaluating IH through a combination of self-report, other report, behavioral, and observational methods.

The CIHS promotes the goals of positive psychology by providing a tool for assessing IH, a variable that has the potential to foster positive intrapersonal and interpersonal characteristics. Hopefully, the scale will be useful for future research to discover more about how IH impacts human flourishing and how IH can be promoted.

Acknowledgments

Portions of this article were presented at the annual convention of the Association for Psychological Science in May 2014.

Funding

This article was made possible through the support of a grant from the Fuller Theological Seminary/Thrive Center in concert with the John Templeton Foundation. The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the Fuller Thrive Center or the John Templeton Foundation.

References

Aiken, L. A. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, *45*, 131–142.

Ashton, M. C., & Lee, K. (2008). The HEXACO model of personality structure and the importance of the H factor. *Social and Personality Psychology Compass*, *2*, 1952–1962. doi:10.1111/j.1751-9004.2008.00134.x

Baehr, J. (2011). *The inquiring mind: On intellectual virtues and virtue epistemology*. New York, NY: Oxford University Press.

Baltes, P. B., & Smith, J. (2008). The fascination of wisdom: Its nature, ontogeny, and function. *Perspectives on Psychological Science*, *3*, 56–64. doi:10.1111/j.1745-6916.2008.00062.x

Bollinger, R. A., & Hill, P. C. (2012). Humility. In T. G Plante (Ed.), *Religion, spirituality, and positive psychology: Understanding the psychological fruits of faith* (pp. 31–48). Santa Barbara, CA: Praeger.

Brady, M. S., & Pritchard, D. (2003). Moral and epistemic virtues. *Metaphilosophy*, *34*, 1–11.

Byrne, B. A. (2010). Structural equation modeling with AMOS: Basic concepts, applications, and programming (2nd ed.). New York, NY: Routledge.

Campbell, W. K., Bonacci, A. M., Shelton, J., Exline, J. J., & Bushman, B. J. (2004). Psychological entitlement: Interpersonal consequences and validation of a self-report measure. *Journal of Personality Assessment*, *83*, 29–45.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349–354. doi:10.1037/h0047358

Davis, D. E., Worthington, E. J., & Hook, J. N. (2010). Humility: Review of measurement strategies and conceptualization as personality judgment. *The Journal of Positive Psychology*, *5*, 243–252. doi:10.1080/17439761003791672

Elder, L., & Paul, R. (2012). Critical thinking: Competency standards essential to the cultivation of intellectual skills, part 4. *Journal of Developmental Education*, *35*, 30–31.

Emmons, R. A. (1999). *The psychology of ultimate concerns*. New York, NY: Guilford.

Exline, J. J., Campbell, W. K., Baumeister, R. F., Joiner, T., & Krueger, J. (2004). Humility and modesty. In C. Peterson, & M. Seligman (Eds.), *The Values in Action (VIA) classification of strengths* (pp. 461–475). Cincinnati, OH: Values in Action Institute.

Gentile, B., Miller, J. D., Hoffman, B. J., Reidy, D. E., Zeichner, A., & Campbell, W. (2013). A test of two brief measures of grandiose narcissism: The Narcissistic Personality Inventory–13 and the Narcissistic Personality Inventory–16. *Psychological Assessment*, *25*, 1120–1136. doi:10.1037/a0033192

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C., & Gough, H. G. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*, 84–96. doi:10.1016/j.jrp.2005.08.007

Gregg, A. P., & Mahadevan, N. (2014). Intellectual arrogance and intellectual humility: An evolutionary-epistemological account. *Journal of Psychology & Theology*, *42*, 7–18.

Grossmann, I., Na, J., Varnuma, M. W., Park, D. C., Kitayama, S., & Nisbett, R. E. (2010). Reasoning about social conflicts improves into old age. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 7246–7250. doi:10.1073/pnas.1001715107

Gruppen, L. D. (2014). Humility and respect: Core values in medical education. *Medical Education*, *48*, 53–58. doi:10.1111/medu.12269

Hopkin, C. R., Hoyle, R. H., & Toner, K. (2014). Intellectual humility and reactions to opinions about religious beliefs. *Journal of Psychology & Theology*, *42*, 50–61.

Jones, W. E. (2012). Higher education, academic communities, and the intellectual virtues. *Educational Theory*, *62*, 695–711. doi:10.1111/edth.12005

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy enlisted personnel* (Research Branch Rep. No. 8-75). Millington, TN: Naval Technical Training, U.S. Naval Air Station.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford.

Kross, E., & Grossmann, I. (2012). Boosting wisdom: Distance from the self enhances wise reasoning, attitudes, and behavior. *Journal of Experimental Psychology: General*, *141*, 43–48. doi:10.1037/a0024158

Landrum, R. E. (2011). Measuring dispositional humility: A first approximation. *Psychological Reports*, *108*, 217–228. doi:10.2466/02.07.09.pr0.108.1.217-228

McElroy, S. E., Rice, K. G., Davis, D. E., Hook, J. N., Hill, P. C., Worthington, E. L., Jr., & Van Tongeren, D. R. (2014). Intellectual humility: Scale development and theoretical elaborations in the context of religious leadership. *Journal of Psychology & Theology*, *42*, 19–30.

Paul, R., & Elder, L. (2001). *Miniature guide to critical thinking: Concepts & tools*. Dillon Beach, CA: Foundation for Critical Thinking.

Roberts, R. C., & Wood, W. J. (2007). *Intellectual virtues: An essay in regulative epistemology*. Oxford, UK: Clarendon.

Rowatt, W. C., Powers, C., Targhetta, V., Comer, J., Kennedy, S., & Labouff, J. (2006). Development and initial validation of an implicit measure of humility relative to arrogance. *Journal of Positive Psychology*, *1*, 198–211. doi:10.1080/17439760600885671

Ryan, S. (2012). Wisdom, knowledge and rationality. *Acta Analytica*, *27*, 99–112. doi:10.1007/s12136-012-0160-6

Samuelson, P. L., Church, I. M., Jarvinen, M., & Paulus, T. (2012). *The science of intellectual humility white paper*. Unpublished manuscript, Fuller Theological Seminary School of Psychology, Pasadena, CA.

Sherif, M., & Cantril, H. (1947). *The psychology of ego-involvements*. New York, NY: Wiley.

Stafford, S. P. (2010). Intellectual virtue in environmental virtue ethics. *Environmental Ethics*, *32*, 339–352.

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, *13*, 225–247.

Tangney, J. P. (2000). Humility: Theoretical perspectives, empirical findings and directions for future research. *Journal of Social and Clinical Psychology*, *19*, 70–82. doi:10.1521/jscp.2000.19.1.70

Vorobej, M. (2011). Distant peers. *Metaphilosophy*, *42*, 708–722. doi:10.1111/j.1467-9973.2011.01714.x

Wayment, H. A., & Bauer, J. J. (2008). *Transcending self-interest: Psychological explorations of the quiet ego*. Washington, DC: American Psychological Association.

Weiner, I. B., & Greene, R. L. (2008). *Handbook of personality assessment*. Hoboken, NJ: Wiley.

Zagzebski, L. T. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge, UK: Cambridge University Press.

Copyright of Journal of Personality Assessment is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.