




Social Psychological Bulletin

Intellectual Humility: Validation and Comparison of Four Self-Report Scales in the German Context

Larissa Knöchelmann^{1,2} , Jule Anna Kemmer¹ , J. Christopher Cohrs¹ 

[1] *Department of Psychology, Marburg University, Marburg, Germany.* [2] *Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany.*

Social Psychological Bulletin, 2026, Vol. 21, Article e12465, <https://doi.org/10.32872/spb.12465>

Received: 2023-07-25 • Accepted: 2025-05-06 • Published (VoR): 2026-01-16



Handling Editor: Nikhila Mahadevan, University of Essex, Colchester, United Kingdom

Corresponding Author: Larissa Knöchelmann, Humboldt-Universität zu Berlin, Department of Psychology, Rudower Chaussee 18, 12489 Berlin. E-mail: larissa.knoechelmann@hu-berlin.de

Supplementary Materials: Data, Preregistration [see [Index of Supplementary Materials](#)]



Abstract

Intellectual humility (IH), the recognition of one's intellectual limitations, is a promising characteristic to tackle societal conflicts such as affective political polarization. Despite increasing research on IH in recent years, most research has been conducted in the USA, probably due to a lack of scales that are validated in languages other than English. Our objective was to offer researchers in the German context several IH measures to choose from depending on their specific research question. Therefore, we validated and compared four established IH scales within the German context. Items of three widely-used IH scales (CIHS, SIHS, LIHS) were translated from English into German by two independent researchers, rated by experts ($n = 8$), and pre-tested ($n = 13$). We then assessed the structural, convergent, discriminant, and incremental validity of the scales plus an existing Swiss-German IH scale (IHS) in a preregistered online-survey ($N = 698$), representative for Germany regarding age, gender, and education. Discriminant validity was assessed via social desirability, cognitive closure, need for cognition, HEXACO honesty-humility, and openness, and incremental validity regarding affective polarization. Results showed that the IHS did not meet all of our pre-registered criteria, potentially due to cross-cultural differences between Germany and Switzerland. However, the three translated scales were comparable to the original English scales regarding structural, convergent, and discriminant validity. The SIHS showed the best incremental validity in predicting less affective polarization towards opinion-based outgroups. Limitations and directions for future research regarding IH in German-speaking contexts are discussed.



This is an open access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), CC BY 4.0, which permits unrestricted use, distribution, and reproduction, provided the original work is properly cited.

Keywords

intellectual humility, scale validation, German version, confirmatory factor analysis, affective polarization

Highlights

- Intellectual humility is promising to bridge political divides and affective polarization. However, current research is limited mainly to anglophone contexts and heterogeneous measurement instruments.
- We translated, validated, and compared four established intellectual humility scales in a sample ($N = 698$) that was representative of Germany's population in terms of demographic variables.
- Results showed that our translations of three intellectual humility scales were comparable to the original English scales. A contextualized measure to assess specific intellectual humility achieved the highest incremental validity in predicting less affective polarization beyond many other psychological variables.
- We hope that our research facilitates cross-cultural work on intellectual humility.

Intellectual humility (IH), the recognition of one's intellectual limitations (Porter et al., 2021), is highly relevant to address urgent societal issues studied in psychological research across subdisciplines. For instance, IH might be a promising construct when aiming to tackle biased information processing (Bar-Tal & Halperin, 2011) in the context of societal conflicts and affective political polarization (Porter et al., 2022). Here, IH has been found to be associated with greater openness to opposing political views (Porter & Schumann, 2018), with being more politically tolerant of others (Krumrei-Mancuso & Newman, 2021), and with more knowledge acquisition, such as reflective and open-minded thinking (Krumrei-Mancuso et al., 2020).

Despite increasing research on IH, most research is currently limited to Anglophone contexts (Porter et al., 2022). This may be due to a lack of validated scales other than English. Hence, our aim was to validate and compare four established IH scales with different scopes and contributions in the German context to enable research on IH in this context. We validated the Leary Intellectual Humility Scale (LIHS) by Leary et al. (2017), the Comprehensive Intellectual Humility Scale (CIHS) by Krumrei-Mancuso and Rouse (2016), the Intellectual Humility Scale (IHS) by Alfano et al. (2017), and the Specific Intellectual Humility Scale (SIHS) by Hoyle et al. (2016). Each scale contains unique features that could prove valuable to future research, as we discuss below.

Intellectual Humility Classification Framework

Increasing research on IH in recent years led to a variety of definitions and self-report scales across both philosophy and psychology (Alfano et al., 2020; Du & Cai, 2020; Porter et al., 2021). Porter et al. (2021) conducted a systematic review of definitions and scales,

and developed an Intellectual Humility Classification Framework. This framework helps to understand the scope of a specific IH scale depending on item content, enabling comparisons across scales. According to this analysis, items can be categorized into four quadrants of classification. These quadrants result from the combination of two key dimensions: Whether an IH item is self-focused vs. other-focused as well as whether IH is internal (cognitions) vs. external (observable behaviors) (Porter et al., 2021, Figure 2). Each scale we chose to validate in Germany covers between one and four different quadrants. Moreover, the IHS, CIHS, and LIHS conceptualize IH as a general personality construct, whereas the SIHS measures IH regarding specific attitudes or (political) topics. We decided to validate these scales with their different scopes as our aim is to offer researchers in the German context a variety of IH measurements they can choose from depending on their specific research question. In the following, we present key features of each scale.

Comprehensive Intellectual Humility Scale

The CIHS developed by Krumrei-Mancuso and Rouse (2016) covers inter- and intrapersonal aspects of IH. The 22 items load on four factors: Independence of Intellect and Ego (“When someone contradicts my most important beliefs, it feels like a personal attack”, reverse-coded), Openness to Revising One’s Viewpoint (“I am willing to change my opinions on the basis of compelling reason”), Respect for Others’ Viewpoints (“I am willing to hear others out, even if I disagree with them”), and Lack of Intellectual Overconfidence (“My ideas are usually better than other people’s ideas”, reverse-coded). The measure shows good internal consistency ($\alpha = .82$ to $.87$) for the total scale and a test-retest reliability of $.75$ after one month and $.70$ after three months (Krumrei-Mancuso & Rouse, 2016). In the original study, the fit of the model with an overall factor ($\chi^2(205) = 368.29$, GFI = $.921$, CFI = $.948$, RMSEA = $.046$) was similar to the model without such a factor, $\chi^2(203) = 367.101$, GFI = $.921$, CFI = $.948$, RMSEA = $.046$. The CIHS covers all quadrants in the IH classification framework (Porter et al., 2021) and has frequently been used to examine the relationship between IH and various social and political constructs (e.g., Bowes et al., 2020; Krumrei-Mancuso & Newman, 2020).

Intellectual Humility Scale

The IHS (Alfano et al., 2017) is the only IH scale that has already been translated into (Swiss-)German. The scale consists of 23 items loading on four inter-correlated factors of IH: Open-Mindedness (“I feel no shame learning from someone who knows more than me”), Intellectual Modesty (“I like to be the smartest person in the room”, reverse-coded), Corrigibility (“I appreciate being corrected when I make a mistake”) and Engagement (“I find it boring to discuss things I don’t already understand”, reverse-coded). The authors’ validation study showed a good fit for the 4-factor solution (German version: $\chi^2(224) = 359.89$, CFI = 0.88 , RMSEA = 0.05 , SRMR = 0.06) as well as correlations between self-

and informant-ratings on the four subscales between $r = .28$ and $.47$. The IHS covers all quadrants in the IH classification framework (Porter et al., 2021). The fit indices and factor structure of the Swiss-German version were shown to be similar to the English version.

Leary Intellectual Humility Scale

The LIHS by Leary et al. (2017) consists of 6 items which load on one factor. It exhibits good internal consistency ($\alpha = .73$ to $.82$) and good structural (CFI = $.91$, SRMR = $.05$) and construct validity in the original study. The LIHS covers the two quadrants self-directed cognitions (e.g., “I accept that my beliefs and attitudes may be wrong”) and other directed cognitions (e.g., “I recognize the value in opinions that are different from my own”) in the IH classification framework (Porter et al., 2021). Due to its brevity, the scale holds the potential to be used in a range of questionnaires without excessively burdening participants.

Specific Intellectual Humility Scale

The SIHS (Hoyle et al., 2016) is a unidimensional scale measuring IH regarding specific opinions, beliefs, and positions (e.g., “I recognize that my views about gun control are based on limited evidence”). The SIHS covers the quadrant self-directed cognitions in the IH classification framework (Porter et al., 2021). The nine-item scale (model without equality constraints: $\chi^2(162) = 490.1$, CFI = $.93$, SRMR = $.07$, RMSEA = $.12$) can be abbreviated to a three-item short-scale (model with loadings, uniqueness and topics constrained to equality: $\chi^2(18, N = 804) = 25.0$, CFI = $.99$, SRMR = $.09$, RMSEA = $.04$). Previous research has shown that specific and general IH correlate moderately ($r = .24$ to $.63$ with LIHS) but are distinct aspects of IH (Hoyle et al., 2016). The scale shows good internal consistency (nine-item scale: $\alpha = .88$ to $.93$; three-item scale: $\alpha = .77$ to $.86$; Hoyle et al., 2016). In our validation study, it is the only scale that measures IH in relation to a specific political topic.

Validating IH Scales

Self-report scales can be a valid tool to assess psychological constructs (Haefffel & Howald, 2010). However, it is necessary to validate them thoroughly and be transparent about the process and results (Flake & Fried, 2020) – also when transferring scales to another language or culture (International Test Commission, 2018). After translating the scales to German, we therefore assessed structural, convergent, discriminant and incremental validity of each IH scale.

Structural Validity

To assess structural (or factorial) validity, “the degree to which the scores of the measurement instrument are an adequate reflection of the dimensionality of the construct being

measured” (Aertssen et al., 2016, p. 894), we expected that the original factor solution of each IH scale would be replicated in the German context (H1). We pre-registered to use Schermelleh-Engel et al.’s (2003) criteria to assess model fit in confirmatory factor analyses (good fit: RMSEA \leq .05, NNFI \geq .97, CFI \geq .97, SRMR \leq .05; acceptable fit: RMSEA \leq .08, NNFI \geq .95, CFI \geq .95, SRMR \leq .10). It has to be noted, however, that the pre-registered cut-off values were stricter than those achieved in the original scale validations, which we only recognized after pre-registering the analyses.

Convergent Validity

Convergent validity “is the extent to which a construct measured in different ways yields similar results” (Boateng et al., 2018, p. 5). Based on these theoretical considerations and previous research showing that the CIHS and LIHS correlated at $r = .61$ to $.71$ (Bowes et al., 2020), we expected (H2) that all three general IH scales (CIHS, IHS, LIHS) would be highly positively correlated with each other ($r_s \geq .50$). Additionally, we expected the topic-specific intellectual humility scale (SIHS) to be at least moderately positively correlated ($r_s \geq .20$) with the general intellectual humility scales based on previous findings (Bowes et al., 2020; Hoyle et al., 2016).

Discriminant Validity

To assess discriminant validity, “the extent to which a measure is novel and not simply a reflection of some other construct” (Boateng et al., 2018, p. 14), we measured social desirability, need for cognition, need for cognitive closure, the personality factors honesty-humility and openness, and demographic variables.

Social Desirability – Self-report questionnaires can be confounded with social desirability, the “tendency to present oneself in a positive light” (Paulhus, 2017, p. 2). As this might also apply to IH self-report measures (Meagher et al., 2015), social desirability was included in previous work on IH. The LIHS and SIHS scales showed small, or no correlations with social desirability in the original validation papers (LIHS: $r = .03$, SIHS: $r_s = -.02$ to $-.16$). The CIHS had only a small correlation, or no correlation, with social desirability ($r \leq .18$, Krumrei-Mancuso et al., 2020) whereas some subfacets of the IHS correlated positively with socially desirable responding ($r_s = .09$ - $.40$; Alfano et al., 2017).

Need for Cognition – Intellectually humble people tend to have higher cognitive flexibility and intelligence (Zmigrod et al., 2019) and engage in more information seeking (Gorichanaz, 2022). Thus, validation studies should investigate in how far IH is conceptually different from Need for Cognition (NFC), an “individual’s tendency to engage in and enjoy effortful cognitive endeavors” (Cacioppo et al., 1984, p. 306). Previous research has shown small to moderate correlations of NFC with IH, LIHS: $r = .34$ (Leary et al., 2017), $r = .26$ (Porter & Schumann, 2018). In a regression analysis, the CIHS accounted for 8.2% of

the variance in need for cognition over and above social desirability (Krumrei-Mancuso et al., 2020).

Need for Cognitive Closure — To show that IH is distinct from solely accepting ambiguity, previous research assessed discriminant validity of IH with Need for Cognitive Closure (NCC), “a desire for an answer on a given topic [...] compared to confusion and ambiguity” (Webster & Kruglanski, 1994, p. 1049). There were indeed only zero to small negative correlations between IH and NCC, $r = -.14$ (Leary et al., 2017); $r = -.05$ (Porter & Schumann, 2018).

Honesty-Humility — IH has been conceptualized as an aspect of humility “the virtuous mean between [...] arrogance, [...] and self-deprecation or diffidence” (Church & Barrett, 2016). A related construct is honesty-humility, which includes the notion of being “honest, sincere, fair, and modest versus [being] greedy, conceited, deceitful, and pretentious” (Ashton et al., 2014, p. 140). In previous research, the CIHS was shown to be slightly correlated with the HEXACO Honesty-Humility subscale ($r = .21$) as well as exhibiting small to moderate correlations with other measures of humility (Krumrei-Mancuso & Rouse, 2016). These findings underline the idea of IH being related but not identical to honesty-humility.

Openness to Experience — IH is conceptually different from the personality trait ‘openness to experience’, as the latter includes the notions of “being imaginative, cultured, curious, original, broad-minded, intelligent, and artistically sensitive” (Barrick & Mount, 1991, p. 4) whereas IH describes a meta-cognition of knowing one’s intellectual limitations. Accordingly, previous validation studies showed moderate correlations of IH with this personality trait: The CIHS correlated with IPIP-openness: $r = .40$ (Krumrei-Mancuso & Rouse, 2016, Study 4). The LIHS correlated with BFI-openness: $r = .33$ (Leary et al., 2017, Study 1), and the SIHS with NEO-PI-R openness: $r = .11$ to $.21$ (Hoyle et al., 2016, Study 2). The IHS subscales correlated with Big Six Originality/Talent between $r = -.20$ to $.42$ (Alfano et al., 2017).

Demographic Variables — In previous studies, zero to small correlations were found between IH scores and demographics such as gender, race, education, or age (Krumrei-Mancuso et al., 2020; Krumrei-Mancuso & Rouse, 2016). In the US context, Republicans and Democrats did not differ in their levels of IH (Krumrei-Mancuso & Newman, 2021).

Incremental Validity

When introducing a new measure to a new context, it is important to justify how this measure “provides information that was formerly unavailable or less adequately obtained” (Hunsley & Meyer, 2003, p. 449). This can be achieved by assessing incremental validity, “the improvement obtained by adding a particular procedure or technique to an

existing combination of assessment methods” (American Psychological Association, n.d.). As outlined in the previous section, IH is correlated with other constructs such as need for cognition or other measures of humility. Thus, one may ask how much IH measures might add to the prediction of an important criterion variable, such as seeking to explain affective polarization.

Affective polarization refers to the “extent to which partisans [or political groups more broadly] treat each other as a disliked outgroup” (Wojcieszak & Warner, 2020, p. 1). IH is a promising construct to explain less affective political polarization (Jost et al., 2022; Porter et al., 2022) as it is associated with less biased information processing. For instance, intellectually humble individuals enjoy weighing competing ideas (McElroy et al., 2014), are better in differentiating weak from strong arguments (Leary et al., 2017), and believe less in under-supported political statements (Krumrei-Mancuso & Newman, 2020). Moreover, IH might address affective polarization by making individuals constructively deal with contrary-minded others. For instance, intellectually humbles are less defensive in disagreements (McElroy et al., 2014) and can admit that other viewpoints have positive features, too (Hodge et al., 2021).

Indeed, previous research has shown that IH was associated with less affective polarization regarding political and religious groups (Krumrei-Mancuso & Newman, 2020). Bowes et al. (2020) used three of the IH measures we considered (CIHS, LIHS, SIHS) and found negative correlations ($r_s = -.25$ to $-.44$) between IH scale means and different measures of affective polarization. Regarding assessing incremental validity, IH predicted open-minded thinking and tolerance towards others after accounting for social desirability and general humility (Krumrei-Mancuso & Rouse, 2016). However, there have been no strong tests of incremental validity by testing whether IH can predict affective polarization beyond other established scales, such as need for cognition or need for cognitive closure. Therefore, we assessed incremental validity by measuring affective polarization and determining the extent to which IH can explain variance in reduced affective polarization beyond the discriminant constructs assessed (H4).

Current Study

The aim of the study was to validate and compare four IH scales (Alfano et al., 2017; Hoyle et al., 2016; Krumrei-Mancuso & Rouse, 2016; Leary et al., 2017) in the German context. The scales selected for our study have different scopes according to the IH framework, and can be used to study diverse research questions in future psychological research. Thus, our objective was not to create a new IH scale but to offer researchers in the German context several IH measures they can choose from depending on their specific research question. The availability of the same scales in different language can also facilitate cross-cultural research in IH (Porter et al., 2022). Additionally, we contribute to the IH literature by following a peer-reviewed pre-registration allowing the four IH scales to be systematically compared.

Hypotheses

H1) Structural Validity

The original factor solution of each IH scale is replicated in the German context. The detailed hypotheses can be found in the pre-registration (see Knöchelmann et al., 2021).

H2) Convergent Validity

All three general IH scales (CIHS, IHS, LIHS) are highly positively correlated with each other ($r_s \geq .50$). The topic-specific intellectual humility scale (SIHS) is moderately positively correlated with the general intellectual humility scales ($r_s \geq .20$).

H3) Discriminant Validity

IH is different from other established psychological constructs. We expect zero- to small correlations ($|r_s| \leq .20$) of IH with: social desirability, political orientation, age, gender, and participants' highest degree of education. Moreover, we expect small to medium correlations ($|r_s| \leq .40$) of IH scales with openness (+), honesty-humility (+), cognitive closure (-), and need for cognition (+).

H4) Incremental Validity

Over and above other psychological constructs (social desirability, openness, honesty-humility, cognitive closure, need for cognition), IH predicts less affective polarization.

Method

Transparency and Ethics Approval

The pre-registration is available at <https://doi.org/10.23668/psycharchives.5202>. Data and analyses are available (see Knöchelmann, 2024). The study was approved by the ethics committee of the authors' university on 24th September 2021 (file reference number '2021-39k').

Translation and Revision of IH Items

Based on recommendations of the [International Test Commission \(2018\)](#), the first two authors (both German native speakers) translated the items of the three English IH scales (CIHS, LIHS, SIHS) into German independently. Afterwards, $n = 8$ experts of social psychological research, who were German native speakers and fluent in English, independently rated which translation captured the original item better and made suggestions on how to improve the wording. Then, the two translators jointly chose and partially revised the translations based on the suggestions. Afterwards, $n = 13$ participants pre-tested the whole questionnaire. Based on their comments, we made some minor revisions

to improve the comprehensibility of the items and to correct some small grammatical mistakes in the IHS.

Power Analysis and Data Collection

Based on several a priori power analyses for CFA with the semPower shiny app (Moshagen & Erdfelder, 2016), we needed a sample of at least $N = 697$ to find an RMSEA of .05 of the most complex CFA model for H1 (LIHS with 12 degrees of freedom and six manifest variables) with a desired power of .80 and an alpha error = .05; for details see the pre-registration. Participants were sampled via the panel provider Bilendi&respondi. Participants received 1.10 € for their 20-minute participation in the study. The online study was conducted between December 3rd and 21st 2021 on the platform www.soscisurvey.de. Data collection was funded by PsychLab, a service of Leibniz Institute for Psychology (ZPID), as the pre-registration successfully went through a peer-review process before data collection.

Inclusion and Exclusion Criteria

A total of $N = 744$ participants finished the study. Participants who indicated they were younger than 18 years old or did not have proficient German language skills (accepted answers were “fluent” and “native”) were automatically filtered out at the beginning of the survey. Applying our pre-registered criteria, we excluded participants who indicated they had not answered the questions in a genuine manner (self-report question at the end of the survey, $n = 4$), who did not answer all scales/items relevant for our hypotheses (because of a programming error, $n = 5$), or who needed less than six minutes to fill out the questionnaire ($n = 38$). A sample of $N = 698$ participants remained. No participants had to be excluded because of failing attention checks.

Participants

We used independent quotas to stratify for age, gender, and education (highest school degree) to gain a sample similar to the German population on December 31st 2020 (DESTATIS, 2021b, 2021a). At the end of the sampling process, the panel provider could not find any more participants still going to school or having no school degree resulting in the final quotas for education being slightly different from those preregistered (see supplementary file ‘additional_analyses.pdf’). Participants’ political orientation (measured on scales from 1 to 101) offered a good range (left-right placement: $M = 46.76$, $SD = 20.45$; socialist-market liberal placement: $M = 50.60$, $SD = 22.44$; progressive-conservative placement: $M = 52.39$, $SD = 21.55$).

Procedure

After clicking on the survey link provided on the panel platform, participants were informed that the purpose of the study was to test the suitability of a new questionnaire to support research in German-speaking contexts. Participants who gave consent to the data protection guidelines were asked to indicate their demographics (age, gender, occupation, highest level of education, nationality, current country of residency, and level of German) and political orientation. Underage participants and participants indicating they did not speak German fluently were filtered out automatically. For these individuals, the survey ended and they were thanked for their interest in the study.

The remaining participants filled out the three general IH scales. To fill out the SIHS and assess affective polarization, we then asked participants: “Regarding which policy do you have a clear opinion? Even if this applies to several policies, please choose one option. Participants could choose one out of three options: “Use of gendered language”, “Introduction of rent caps in cities through the federal government”, and “Stop building single family houses because of environmental reasons”. Afterwards, participants filled out the SIHS, followed by the constructs used to assess discriminant and incremental validity. The order of the three general IH scales was randomized, as was the order of the scales assessing discriminant validity, and the order of items within each scale. After asking participants whether they answered the questions in an honest manner, we thanked participants for their participation in the study, provided further information about IH and gave them the opportunity to leave an anonymous comment.

Measures

IH Scales

Comprehensive Intellectual Humility Scale – The CIHS (Krumrei-Mancuso & Rouse, 2016) consists of 22 items (11 reverse-coded), e.g., “I have at times changed opinions that were important to me, when someone showed me I was wrong” which were measured on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*.

Intellectual Humility Scale – The German version of the IHS (Alfano et al., 2017) consists of 23 items (15 reverse-coded), e.g., “I feel no shame learning from someone who knows more than me”, answered on a 5-point scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*.

Leary General Intellectual Humility Scale

The LIHS (Leary et al., 2017) consists of 6 items (none reverse-coded), e.g., “I recognize the value in opinions that are different from my own”, answered on a 5-point scale ranging from 1 = *not at all like me* to 5 = *very much like me*.

Specific Intellectual Humility Scale — The SIHS consists of nine items (none reverse-coded) which can be reduced to three items (Hoyle et al., 2016). It measures IH regarding a specific (political) topic, “My views about [topic] are just as likely to be wrong as other views”. Participants filled out the scale for one of three topics (use of gender-neutral language, introduction of rent-caps in German cities through the federal government, or stop building single-family houses for environmental reasons) on a 5-point scale ranging from 1 = *not at all like me* to 5 = *very much like me*.

Discriminant Validity Scales

Social Desirability — Social desirability was measured with the KSE-G (Kemper et al., 2012) which addresses people’s bias to provide unrealistically positive self-descriptions of personality traits. The scale consists of two factors: Downplaying Negative Qualities (NQ-, three items, none reverse-coded; e.g., “It has happened that I have taken advantage of someone in the past”) and Emphasizing Positive Qualities (PQ+, three items; e.g., “Even if I am feeling stressed, I am always friendly and polite to others”). Participants rated to what degree the statements described themselves on a 5-point scale ranging from 0 = *doesn’t apply at all* to 4 = *applies completely*. The two factors were not aggregated into one scale but used independently (PQ+: $\alpha = .63$, $\Omega = .64$; NQ-: $\alpha = .56$, $\Omega = .59$).

Political Orientation — Political orientation was measured with three items (none reverse-coded) asking participants to rate their political views on 101-point bipolar scales (left-right; socialistic-market liberal; liberal-conservative). We provided explanations for the words “socialistic” and “market liberal”: “socialistic = the state should control the economy strongly”, “market liberal = the state should not intervene in the economy at all”. However, the reliability was unsatisfactory ($\alpha = .57$, $\Omega = .63$), thus we continued the analysis, as pre-registered, with the left-right placement only.

Openness — Openness was measured with the *BFI-K* (Rammstedt & John, 2005) consisting of five items (one reverse-coded, $\alpha = .76$, $\Omega = .78$). Participants rated to what degree the statements held true for them, e.g., “I am interested in a variety of things”, on 5-point scales ranging from 1 = *very incorrect* via 3 = *neither nor* to 5 = *very correct*.

Honesty-Humility — Honesty-Humility was measured with the respective *HEXACO* subscale (Ashton & Lee, 2009; translated to German by Moshagen et al., 2014). Participants rated to what degree ten statements (six reverse-coded) described themselves, e.g., “I would never accept a bribe, even if I were sure I could get away with it”, on a 5-point scale ranging from 1 = *do not agree at all* via 3 = *neither nor* to 5 = *completely agree*; $\alpha = .73$, $\Omega = .73$.

Need for Cognitive Closure — NCC was measured with the *Kognitive Geschlossenheitskala* (Schlink & Walther, 2007) consisting of 16 items (three reverse-coded, $\alpha = .77$, Ω

= .77). Participants rated to what degree the statements described themselves, e.g., “I dislike it when a person’s statement could mean many different things”, on a 6-point scale ranging from 1 = *do not agree at all* to 6 = *completely agree*.

Need for Cognition — NFC was measured with the NFC-K (Beißert et al., 2014) consisting of four items (two reverse-coded, $\alpha = .52$, $\Omega = .55$). Participants rated to what degree the statements described themselves, e.g., “I would prefer complex to simple problems”, on a 7-point scale ranging from 1 = *doesn’t apply at all* via 4 = *neither nor* to 7 = *applies completely*.

Incremental Validity: Affective Polarization

Participants’ Opinion on a Political Topic — We asked participants to indicate their opinion on the political topic chosen on a six-point scale ranging from 1 = *I am completely against it* to 6 = *I am completely in favor of it*.

Affective Polarization Regarding Opinion-Based Groups — To calculate the affective polarization score regarding opinion-based groups, participants were asked to indicate their feelings towards supporters and opponents of the political topic chosen on 101-point feeling thermometers from 1 = *cold (unpleasant)* presented in blue color to 101 = *warm (pleasant)* presented in red color. The affective polarization value (per person) was then calculated as the difference between ingroup and outgroup perceptions on the feeling thermometers, so that higher values [-100; 100] indicate higher affective polarization.

Affective Polarization Regarding Political Parties — Party ingroup and outgroup were assessed by asking for the ingroup “Which party of the current parliament would you most likely vote for?” and for the outgroup “Which party of the current parliament would you never vote for?” to apply the least-liked group paradigm (Gibson, 2013). We explicitly stated that participants should decide on choosing one option – even if the statement could be true for several parties. The parties displayed were *CDU/CSU*, *SPD*, *AfD*, *FDP*, *Bündnis 90/Die Grünen*, *DIE LINKE*. Then, participants rated their ingroup and outgroup on 101-point feeling thermometers from 1 = *cold (unpleasant)* presented in blue color to 101 = *warm (pleasant)* presented in red color. The affective polarization value (per person) was then calculated as the difference between ingroup and outgroup perceptions on the feeling thermometers, so that higher values [-100; 100] indicate higher affective polarization.

Quality Checks

Attention Checks — The IHS as well as the cognitive closure scale each contained an additional attention check item: “We are testing your attention here. Please mark ‘fully agree’ here”.

Honesty in Answering the Questionnaire — At the end of the survey, we asked: “Did you answer the questions of this study honestly so that we can use them?”. Participants had to choose between: “Yes, I answered the questions honestly” and “No, I did not answer the questions honestly and just clicked on anything”.

Data Diagnostics and Analytic Strategy

We conducted all analyses as pre-registered and peer-reviewed before data collection.

Results

Pre-Requisite Checks and Outliers

Bartlett’s tests and Kaiser-Meyer-Olkin tests revealed that our data was suitable for factor analysis (Bartlett test: all $ps < .001$; all KMOs $\geq .82$). Additionally, the assumption of multivariate normality with Mardia’s, Royston’s, Doornik-Hansens’s, Henze-Zirkler’s tests and E-statistic was violated. This is why we ran all CFA analyses with Satorra-Bentler correction (Satorra & Bentler, 1994). Moreover, we found $n = 54$ cases of multivariate outliers based on robust Mahalanobis distances (97.5% adjusted quantile). In cases where the exclusion of these participants changed the results substantially (i.e., changing significance levels or cut-offs), we mention both results with and without outliers. Results without outliers are displayed in supplementary file ‘analyses_and_results_without_outliers.html’.

Reliability

All IH scales and subscales showed adequate reliabilities according to our pre-registered cutoff values, i.e., Cronbach’s $\alpha \geq .65$ and McDonald’s $\Omega \geq .70$, see Table 1. For item statistics (M , SD , item intercorrelations, item selectivity), see supplementary file ‘analyses_and_results.html’.

Table 1*Reliability of the Intellectual Humility Scales and Subscales*

(Sub-)Scale	Cronbach's α		McDonald's Ω	
	α	95% CI	Ω	95% CI
CIHS	.87	[.86, .88]	.86	[.84, .88]
Openness to Revising Viewpoint	.82	[.80, .84]	.82	[.79, .85]
Respect for Others' Viewpoints	.82	[.80, .84]	.82	[.80, .85]
Lack of Intellectual Overconfidence	.77	[.74, .79]	.77	[.73, .80]
Independence of Intellect and Ego	.87	[.85, .88]	.87	[.85, .89]
IHS	.81	[.79, .83]	.81	[.78, .83]
Openness	.63	[.58, .67]	.63	[.58, .68]
Intellectual Modesty	.74	[.71, .77]	.75	[.72, .78]
Corrigibility	.63	[.58, .67]	.66	[.61, .70]
Engagement	.66	[.62, .70]	.67	[.63, .71]
LIHS	.77	[.74, .79]	.77	[.73, .80]
SIHS full scale	.92	[.91, .92]	.92	[.90, .93]
Introduction of Rent Caps	.90	[.88, .92]	.90	[.88, .92]
Use of Gendered Language	.92	[.90, .93]	.92	[.90, .93]
Stop Building Single-Family Houses	.91	[.87, .94]	.91	[.87, .95]
SIHS short scale	.87	[.86, .89]	.88	[.86, .90]
Introduction of Rent Caps	.86	[.83, .88]	.86	[.83, .89]
Use of Gendered Language	.88	[.85, .90]	.88	[.85, .91]
Stop Building Single-Family Houses	.86	[.80, .91]	.86	[.80, .92]

H1) Structural Validity

For testing structural validity, we ran CFA with Satorra-Bentler correction in lavaan (Version 06.14-14).

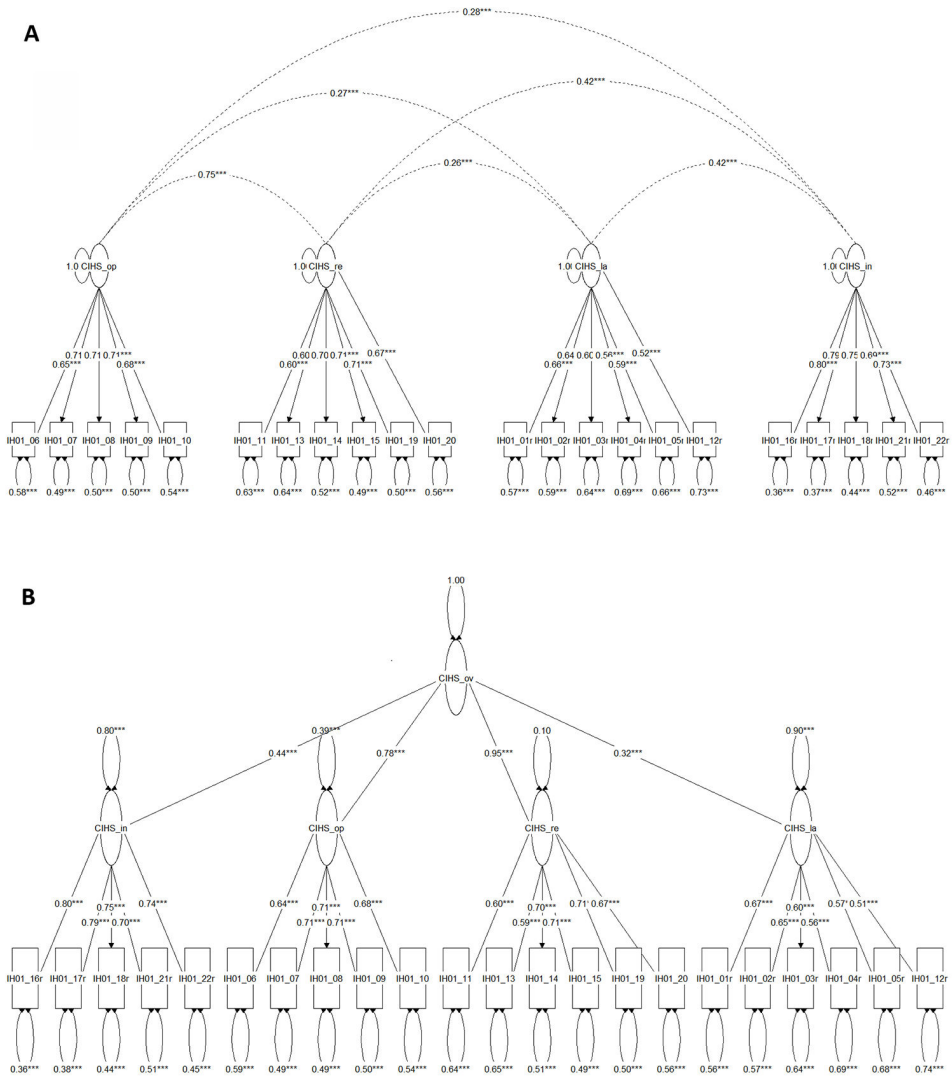
H1a) CIHS

For the CIHS, we tested and compared two competing models as in the original validation study by [Krumrei-Mancuso and Rouse \(2016, p. 215\)](#).

First, we tested a model with four intercorrelated first-order factors. We constrained the variance of the latent variables to be 1, thereby freeing all item loadings. This model met some but not all fit criteria: $\chi^2_m(203) = 473.42$, $p < .001$; $RMSEA = 0.044$, 90% CI [0.039, 0.048], $NNFI = .928$, $CFI = .937$, $SRMR = 0.047$, see [Figure 1A](#). However, model fit was similar to that achieved in the original scale validation.

Figure 1

CIHS Models and Estimates With and Without Higher Order Factor



Note. Standardized parameters. CIHS_op = Openness to Revising one’s Viewpoint, CIHS_re = Respect for Others’ Viewpoints, CIHS_la = Lack of Intellectual Overconfidence, CIHS_in = Independence of Intellect and Ego.

Second, we added an overall second-order IH factor to the four first-order factors (same as above). This model met some but not all fit criteria: $\chi^2_m(205) = 516.47, p < .001$;

RMSEA = 0.055, 90% CI [0.051, 0.060], NNFI = .885, CFI = .897, SRMR = 0.065, see [Figure 1B](#). Comparing the two models with a chi-square difference test showed that introducing an overall second-order factor to the model reduced fit, $\Delta\chi^2(\Delta df = 2) = 34.39$, $p < .001$.

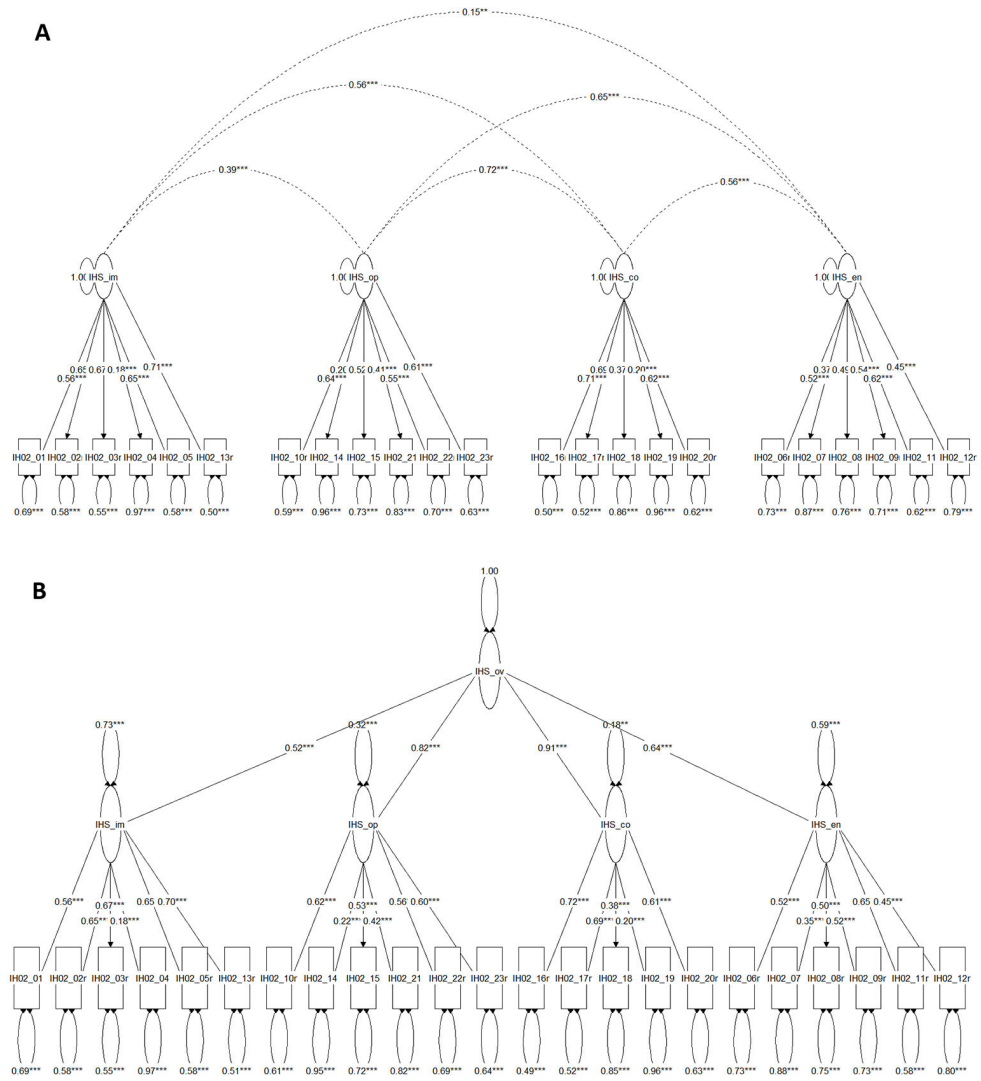
H1b) IHS

For the IHS ([Alfano et al., 2017](#)), we tested and compared two competing models.

First, we tested the authors' original factor solution: A model with four intercorrelated first-order factors. This model met some but not all fit criteria: $\chi^2_m(224) = 552.36$, $p < .001$; RMSEA = 0.046, 90% CI [0.042, 0.050], NNFI = .861, CFI = .877, SRMR = 0.060, see [Figure 2A](#). However, model fit was similar to that achieved in the original scale validation.

Figure 2

IHS Model and Estimates With and Without Higher Order Overall Factor



Note. Standardized parameters. IHS_im = Intellectual Modesty, IHS_op = Open-mindedness, IHS_co = Corrigibility, IHS_en = Engagement.

Second, we added an overall second-order IH factor to the four first-order factors (same as above). This model met some but not all fit criteria: $\chi^2_m(226) = 591.25, p < .001$; RMSEA = 0.047, 90% CI [0.042, 0.051], NNFI = .918, CFI = .927, SRMR = 0.065, see

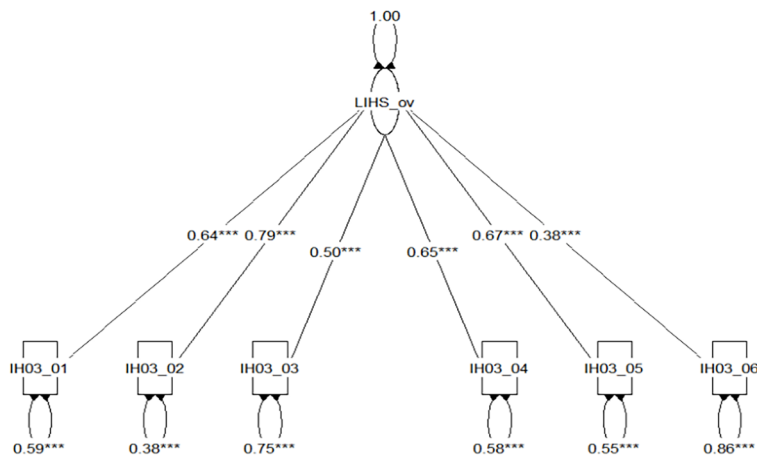
Figure 2B. Comparing the two models with a chi-square difference test showed that adding an overall second-order factor to the model reduced fit, $\Delta\chi^2(\Delta df = 2) = 34.39$, $p < .001$. However, in chi-square tests, even small model deviations can become statistically significant with large sample sizes (Pavlov et al., 2020), and as the model with overall factor showed higher fit regarding NNFI and CFI values, one can also argue that the model with overall factor fits the data better.

H1c) LIHS

For the LIHS (Leary et al., 2017) we tested the authors' original model with all six items loading on one first-order factor. This model revealed acceptable to good fit, except for NNFI: $\chi^2_m(9) = 35.93$, $p < .001$; RMSEA = 0.065, 90% CI [0.048, 0.084], NNFI = .926, CFI = .955, SRMR = 0.043, see Figure 3. The model fit was similar to that achieved in the original scale validation. Excluding multivariate outliers improved model fit even further: NNFI = .953, CFI = .972.

Figure 3

LIHS Model and Estimates



Note. Standardized parameters.

H1d) Specific Intellectual Humility Scale

For the SIHS (Hoyle et al., 2016), we ran multiple group analysis and measurement invariance test for both the full scale and the short scale (as in the original publication).

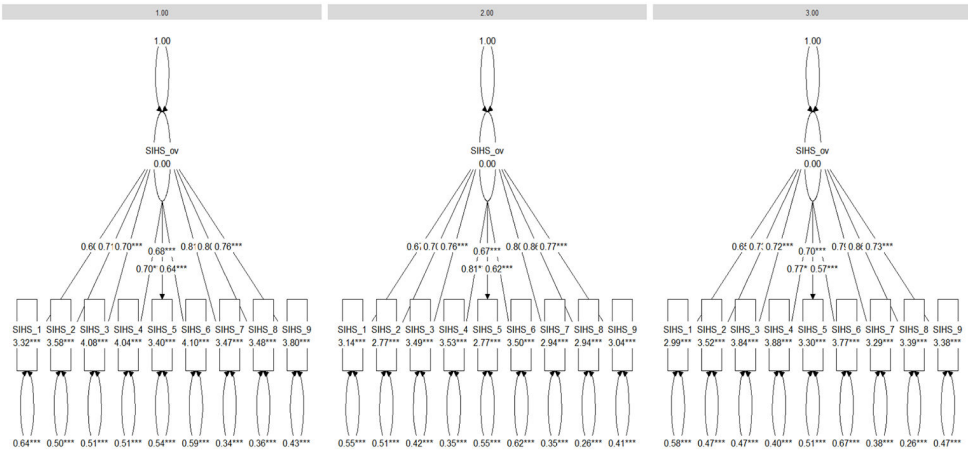
Full Scale — First, we created a model with all nine items loading on one first-order factor and fixing the variance of the latent variable to one, thereby freeing all item loadings. We indicated the three groups (corresponding to the three political topics of

which participants chose one) with a grouping variable without equality constraints. This model did not meet most of our fit criteria: $\chi^2_m(81) = 351.03, p < .001$; RMSEA = 0.120, 90% CI [0.109, 0.130], NNFI = .871, CFI = .903, SRMR = .056, see Figure 4A. However, model fit was similar to that achieved in the original scale validation.

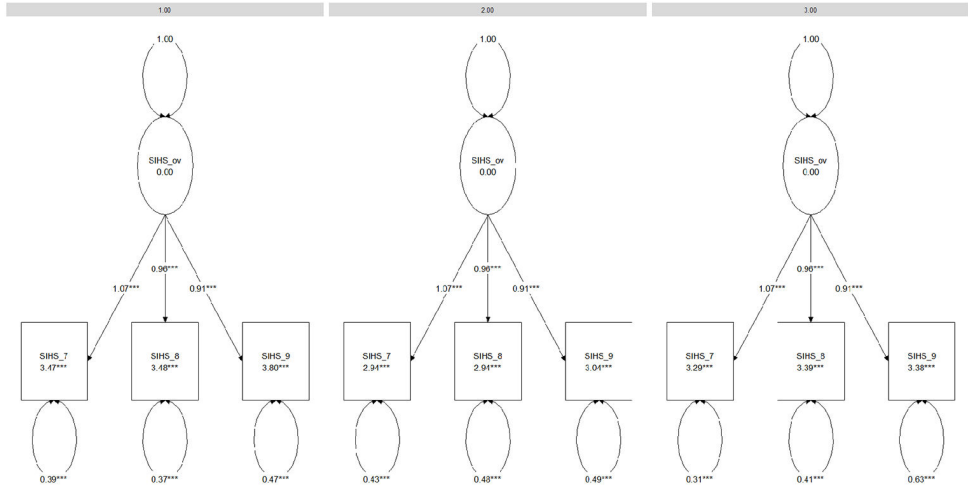
Figure 4

SIHS Model and Estimates

A



B



Note. Standardized parameters. Groups: 1 = Use of gendered language, 2 = Introduction of rent caps, 3 = stop building single family houses. A) Full scale without equality constraints. B) Short scale with equality constraints.

Second, we examined the invariance of all the loadings across topics by testing the same model as above but constraining all item loadings to be equal across groups. This model did not meet our fit criteria: $\chi^2_m(99) = 406.51, p < .001$; RMSEA = 0.116, 90% CI [0.106, 0.126]; NNFI = .880, CFI = .890, SRMR = .113. Comparing the two models with a chi-square difference test showed that the model without equality constraints fitted the data better, $\Delta\chi^2(\Delta df = 18) = 42.25, p = .001$.

Short Scale — First, we created a model with all three short-scale items loading on one first-order factor and fixing the variance of the latent variable to one, thereby freeing all item loadings. We indicated the three groups (corresponding to the three political topics, of which participants chose one) with a grouping variable without equality constraints. We knew that this model would be saturated and that we could not evaluate the model fit. However, we could still estimate and test the parameters as done in the original paper (Hoyle et al., 2016). The standardized parameters were similar across the three different topics: $\lambda = .87, .84$ and $.75$ for group one (topic: use of gendered language), $\lambda = .85, .82$, and $.85$ for group two (topic: introduction of rent caps), and $\lambda = .85, .83$, and $.80$ for group three (topic: stop building single-family houses).

Second, we examined the invariance of all the loadings across topics by testing the same model as above but constraining all item loadings to be equal across groups. This model was testable and met all of our fit criteria with acceptable to good fit: $\chi^2_m(6) = 16.154$, RMSEA = 0.085, 90% CI [0.032, 0.141]; NNFI = .982; CFI = .988; SRMR = 0.079, see Figure 4B. Comparing the item loadings of both models showed that they were quite similar, ranging between .75 and .89.

Except for the 3-item version of the SIHS, the models did not meet our pre-registered cut-off criteria. However, the model fit was generally similar to that achieved in the respective original scale validation. Therefore, we decided to continue with an assessment of convergent, discriminant, and incremental validity.¹

H2) Convergent Validity

To test convergent validity—we calculated two correlation matrices. First, we created a correlation matrix with all IH measures based on scale means and/or subscale means. Additionally, we calculated correlations between the IH scales based on CFA in lavaan as this takes measurement errors into account. Here, we entered all IH scale models (CIHS, IHS, LIHS, SIHS full scale) into one model. As preregistered, we anticipated poor model

1) In addition to the pre-registered analyses, we also explored whether minor modifications (e.g., accounting for reversed item bias or similar wording between items) could improve model fit for each scale. Indeed, when allowing small modifications to our models, model fit of the CIHS, LIHS, and SIHS full scale could be improved substantially, meeting the pre-registered criteria. For the IHS, however, model fit could not be improved enough to meet all cut off values without allowing cross-loadings, thus, structural validity of this scale was not satisfactory in our sample. These exploratory findings can be found in supplementary file 'analyses_and_results.html'.

fit indices when testing multiple translated IH scales simultaneously in one CFA model (with no cross-loadings and correlations of items between scales allowed) as the wording of translated items is highly similar between scales. Therefore, we do not evaluate model fit indices but only provide correlations.

Overall, our hypotheses for convergent validity were mostly confirmed when using Pearson correlations and correlations obtained via SEM (see Table 2), with the following exceptions: The IHS overall scale mean correlated less than expected with the LIHS ($r = .37$; $\varphi = .48$; without multivariate outliers $r = .46$; $\varphi = .65$) and the SIHS ($r = .15$, $\varphi = .19$). The subfacets of the CIHS and IHS were positively correlated with all other general IH scales and subscales, but often smaller than expected. In Pearson correlations, but not SEM, the CIHS subfacet Independence of Intellect and Ego and the IHS subfacet Corrigibility were not correlated with the SIHS ($ps \geq .05$). The other subfacets were positively associated with the SIHS, but often smaller than the expected cut-off value of $r = .20$.²

2) For disattenuated correlation coefficients, see Supplementary Materials “additional analyses.pdf”, Table 3.

Table 2
Convergent Validity Between IH (Sub-) Scales

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. CIHS	–											
2. CIHS_OP	.69** (.79)	–										
	[.65, .73]											
3. CIHS_LA	.65** (.42)	.23** (.33)	–									
	[.61, .70]	[.16, .30]										
4. CIHS_RE	.74** (.85)	.63** (.67)	.21** (.35)	–								
	[.71, .77]	[.59, .68]	[.13, .28]									
5. CIHS_IN	.73** (.53)	.23** (.42)	.34** (.22)	.36** (.45)	–							
	[.69, .76]	[.16, .30]	[.27, .40]	[.29, .42]								
6. IHS	.68** (.87)	.44** (.69)	.38** (.37)	.54** (.74)	.54** (.47)	–						
	[.64, .72]	[.38, .50]	[.32, .45]	[.48, .59]	[.49, .59]							
7. IHS_CO	.49** (.68)	.31** (.54)	.14** (.29)	.40** (.57)	.52** (.36)	.77** (.78)	–					
	[.44, .55]	[.24, .38]	[.07, .22]	[.34, .46]	[.46, .57]	[.74, .80]						
8. IHS_IM	.38** (.39)	.14** (.31)	.35** (.16)	.23** (.33)	.31** (.21)	.67** (.44)	.41** (.34)	–				
	[.31, .44]	[.07, .21]	[.28, .41]	[.16, .30]	[.24, .38]	[.62, .71]	[.34, .47]					
9. IHS_EN	.49** (.61)	.38** (.48)	.28** (.25)	.38** (.51)	.36** (.32)	.66** (.69)	.37** (.54)	.10** (.31)	–			
	[.44, .55]	[.31, .44]	[.21, .35]	[.31, .44]	[.30, .42]	[.62, .70]	[.30, .43]	[.02, .17]				
10. IHS_OP	.55** (.83)	.44** (.66)	.27** (.35)	.53** (.70)	.36** (.44)	.73** (.95)	.48** (.74)	.29** (.42)	.38** (.66)	–		
	[.50, .60]	[.37, .50]	[.20, .34]	[.48, .59]	[.29, .42]	[.70, .76]	[.42, .53]	[.22, .35]	[.31, .44]			
11. LIHS	.47** (.71)	.52** (.56)	.16** (.30)	.50** (.60)	.23** (.38)	.37** (.48)	.30** (.37)	.11** (.21)	.35** (.33)	.30** (.46)	–	
	[.41, .53]	[.46, .57]	[.08, .23]	[.44, .55]	[.16, .30]	[.31, .44]	[.23, .37]	[.04, .19]	[.29, .42]	[.23, .37]		
12. SIHS	.25** (.35)	.29** (.28)	.21** (.15)	.29** (.30)	-.03 (.19)	.15** (.19)	.07 (.15)	.08* (.08)	.10** (.13)	.18** (.18)	.31** (.36)	–
	[.18, .32]	[.22, .35]	[.14, .28]	[.22, .36]	[-.10, .05]	[.08, .22]	[-.01, .14]	[.01, .15]	[.03, .17]	[.11, .25]	[.25, .38]	
13. SIHS_S	.23** (.34)	.29** (.27)	.20** (.14)	.24** (.28)	-.03 (.18)	.12** (.16)	.05 (.12)	.05 (.07)	.11** (.11)	.13** (.15)	.31** (.36)	.92**
	[.16, .30]	[.22, .36]	[.13, .27]	[.17, .31]	[-.10, .05]	[.05, .20]	[-.02, .13]	[.03, .18]	[.06, .20]	[.06, .20]	[.24, .38]	[.91, .93]

Note. Pearson correlations with 95% CI in square brackets. Correlations based on lavaan in round brackets.

* $p < .05$. ** $p < .01$. For exact p -values, see supplementary file analyses_and_results.html.

Table 3
Discriminant Validity

(Sub-)Scale	Age	Gender	Education	PO	CC	NFC	OP	SDNQ	SDPQ	HH
CIHS	.11** (.10) [.03, .18]	-.02 (-.05) [-.10, .05]	.02 (.01) [-.05, .10]	-.06 (-.07) [-.13, .02]	-.40** (-.30) [-.46, -.33]	.23** (.17) [.16, .30]	.11** (.20) [.04, .19]	-.21** (-.35) [-.28, -.14]	.23** (.43) [.16, .30]	.32** (.33) [.25, .39]
CIHS_OP	.08* (.08) [.01, .15]	-.06 (-.04) [-.13, .02]	.07 (.01) [-.01, .14]	-.09* (-.05) [-.16, -.02]	-.20** (-.23) [-.27, -.12]	.19** (.13) [.12, .26]	.18** (.15) [.11, .25]	-.10** (-.27) [-.17, -.03]	.20** (.33) [.13, .27]	.13** (.25) [.06, .20]
CIHS_LA	-.03 (.03) [-.11, .04]	-.12** (-.02) [-.19, -.05]	.02 (.00) [-.06, .09]	-.07 (-.02) [-.14, .01]	-.26** (-.10) [-.33, -.19]	.07 (.06) [-.00, .14]	-.06 (.07) [-.14, .01]	-.09* (-.13) [-.16, -.01]	-.07 (-.14) [-.15, .00]	.26** (.11) [.19, .32]
CIHS_RE	.06 (.10) [-.01, .14]	-.05 (-.04) [-.12, .03]	-.02 (.01) [-.09, .06]	-.06 (-.07) [-.13, .01]	-.22** (-.28) [-.29, -.15]	.20** (.16) [.12, .27]	.22** (.19) [.15, .29]	-.17** (-.33) [-.24, -.09]	.31** (.41) [.24, .37]	.23** (.31) [.16, .30]
CIHS_IN	.17** (.05) [.10, .24]	.13** (-.03) [.06, .20]	-.00 (.00) [-.07, .07]	.04 (-.03) [-.04, .11]	-.40** (-.14) [-.46, -.33]	.21** (.08) [.14, .28]	.03 (.10) [-.05, .10]	-.22** (-.17) [-.29, -.15]	.24** (.21) [.17, .31]	.26** (.16) [.19, .33]
IHS	.18** (.20) [.11, .25]	-.03 (-.09) [-.11, .04]	.05 (.03) [-.03, .12]	-.10** (-.09) [-.18, -.03]	-.43** (-.51) [-.49, -.37]	.25** (.14) [.18, .32]	.14** (.12) [.06, .21]	-.35** (-.62) [-.41, -.28]	.27** (.43) [.20, .33]	.48** (.62) [.42, .54]
IHS_CO	.21** (.18) [.14, .28]	.06 (-.07) [-.02, .13]	-.01 (.03) [-.09, .06]	-.01 (-.08) [-.08, .06]	-.33** (-.45) [-.39, -.26]	.21** (.12) [.14, .28]	.08* (.10) [.00, .15]	-.29** (-.54) [-.36, -.22]	.34** (.38) [.27, .40]	.35** (.54) [.28, .41]
IHS_IM	.26** (.12) [.19, .33]	-.03 (-.04) [-.10, .05]	-.02 (.02) [-.10, .05]	.04 (-.05) [-.03, .11]	-.13** (-.29) [-.20, -.05]	-.03 (.08) [-.10, .04]	-.12** (.07) [-.19, -.05]	-.26** (-.35) [-.33, -.19]	.09* (.25) [.02, .17]	.49** (.36) [.43, .55]
IHS_EN	-.03 (.14) [-.10, .05]	.02 (-.06) [-.06, .09]	.14** (.02) [.07, .21]	-.21** (-.06) [-.28, -.14]	-.55** (-.35) [-.60, -.49]	.42** (.09) [.35, .48]	.29** (.08) [.22, .36]	-.17** (-.43) [-.24, -.10]	.19** (.30) [.11, .26]	.21** (.43) [.13, .28]
IHS_OP	.06 (.15) [-.01, .13]	-.14** (-.08) [-.21, -.06]	.02 (.02) [-.05, .10]	-.11** (-.07) [-.18, -.03]	-.21** (-.40) [-.28, -.14]	.12** (.11) [.05, .19]	.15** (.09) [.08, .22]	-.26** (-.48) [-.33, -.19]	.16** (.34) [.09, .24]	.29** (.48) [.22, .36]
LIHS	.00 (-.02) [-.07, .07]	.03 (.00) [-.05, .10]	.07 (.10) [.00, .15]	-.11** (-.16) [-.19, -.04]	-.23** (-.23) [-.30, -.16]	.21** (.19) [.14, .28]	.18** (.19) [.11, .25]	-.09* (-.14) [-.16, -.01]	.23** (.28) [.16, .30]	.11** (.10) [.04, .19]
SIHS	-.12** (-.12) [-.20, -.05]	-.16** (-.16) [-.23, -.09]	.01 (.02) [-.07, .08]	-.18** (-.18) [-.25, -.11]	-.05 (-.05) [-.13, .02]	.01 (.06) [-.07, .08]	-.03 (-.03) [-.10, .05]	-.00 (.00) [-.07, .07]	.05 (.07) [-.02, .12]	.08* (.17) [.01, .16]
SIHS_S	-.10* (-.11) [-.17, -.02]	-.11** (-.12) [-.19, -.04]	.05 (.06) [-.02, .13]	-.16** (-.16) [-.23, -.08]	-.08** (-.07) [-.15, -.00]	.04 (.10) [-.04, .11]	-.01 (-.01) [-.08, .06]	.02 (.03) [-.05, .09]	.06 (.08) [-.01, .13]	.05 (.05) [-.02, .13]

Note: Pearson correlations with 95% CI in square brackets. Correlations based on CFA in round brackets. Correlations with gender based on $n = 696$ (1 = female, 2 = male), PO = political orientation, CC = cognitive closure, NFC = need for cognition, OP = openness, SDNQ = social desirability, negative qualities, SDPQ = social desirability, positive qualities, HH = honesty-humility.

* $p < .05$. ** $p < .01$. For exact p -values, see supplementary file 'analyses_and_results.html'.

H3) Discriminant Validity

For testing discriminant validity, we first created four correlation matrices with an IH scale and the scales means of the discriminant constructs assessed. Second, we calculated correlations based on CFA in lavaan without evaluating model fit. For each IH scale, we built a model consisting of the IH scale and all discriminant constructs. Each discriminant construct was represented by a first-order factor on which all items of the respective construct loaded, except for the social desirability scale for which two first-order factors were modeled (PQ+ and NQ-).

The two approaches yielded similar results (see Table 3): The SIHS met all of our pre-registered cut-off values. The CIHS met all cut-off values except for both social desirability subscales. The LIHS met all cut-off values except for one social desirability sub-scale. The IHS did not show adequate discriminant validity regarding age, political orientation, social desirability, honesty-humility, need for cognition, and cognitive closure (via Pearson-correlations).²

H4) Incremental Validity

To assess incremental validity, we conducted a blockwise regression for each IH scale with affective polarization as the dependent variable. The first block of independent variables included the psychological constructs already used to assess discriminant validity (social desirability (NQ- and PQ+), openness, honesty-humility, cognitive closure, need for cognition). Then we entered the IH scale to assess the change in explained variance of affective polarization. For the SIHS, we used affective polarization scores regarding opinion-based groups. We found that the SIHS incrementally predicted less affective polarization as expected ($\beta = -.13$, $p < .001$).

For the three general IH scales (CIHS, IHS, LIHS), we used affective polarization scores regarding political parties. Contrary to our hypothesis, the CIHS overall score predicted *more* affective polarization, $\Delta R^2 = .013$, $\beta = .13$, $p = .002$. However, when entering the CIHS subfacets instead of the overall mean, none of the subfacets was associated with affective polarization, $\Delta R^2 = .016$, $|\beta|s < .08$, $ps > .08$. The IHS overall scale predicted *higher* affective polarization, too, $\Delta R^2 = .012$, $\beta = .14$, $p = .004$. When entering the IHS subfacets instead of the overall mean, further variance of affective polarization could be explained, $\Delta R^2 = .032$. Here, Corrigibility ($\beta = .07$, $p = .159$) and Engagement ($\beta = .09$, $p = .062$) did not predict affective polarization significantly. Open-Mindedness predicted *more* affective polarization, $\beta = .12$, $p = .008$, while, as expected, Intellectual Modesty predicted *less* polarization, $\beta = -.10$, $p = .038$. Without outliers, Engagement predicted more polarization, too, $\beta = .10$; $p = .038$. The LIHS did not explain affective polarization beyond the other discriminant constructs, $\Delta R^2 = .003$, $\beta = .06$, $p = .14$. However, when excluding outliers, the LIHS predicted more polarization, too, $\Delta R^2 = .008$, $\beta = .10$, $p = .02$.

Because of these unexpected effects, we ran several exploratory analyses to understand our findings better, see supplementary materials.³

Discussion

With the study presented in this manuscript, we validated and compared four IH scales within the German context. After translating and revising German translations of the three originally English scales, we ran a pre-registered and peer-reviewed online survey to test the validity of the scales. Our sample ($N = 698$) was quota-stratified regarding age, gender, and education for Germany's population. Besides reliability as well as structural and convergent validity, we assessed discriminant validity by including established scales for social desirability, cognitive closure, need for cognition, honesty-humility, and openness. We also assessed incremental validity in relation to affective polarization.

Results showed that the reliability of all scales and subscales was good. In confirmatory factor analyses testing structural validity, only the 3-item SIHS scale met the (rather conservative) pre-registered fit indices by Schermelleh-Engel et al. (2003). However, all scales showed comparable fit to the models published in the original validation articles. Thus, our a priori criteria to assess model fit might have been too strict and should not be overinterpreted according to the authors themselves (Schermelleh-Engel et al., 2003). Additionally, when allowing small modifications to our models, e.g., residual correlations between items with similar wording or accounting for reversed item bias (Weijters et al., 2013), model fit of the CIHS, LIHS, and SIHS full scales could be improved substantially, meeting the pre-registered criteria. Thus, we encourage using these IH scales both as manifest values as well as latent variables in SEM. For the IHS, model fit was satisfactory regarding RMSEA and SRMR, but not regarding NNFI and CFI without allowing cross-loadings. Thus, the factor structure found in Switzerland does not seem to replicate well in the German context.

Convergent validity of the CIHS, LIHS, and SIHS overall scale was met, indicating that these scales measure the same underlying psychological construct (Boateng et al., 2018). However, the IHS overall mean and subscales as well as some subscales of the CIHS and IHS correlated less strongly than expected. These findings underline that the multidimensional scales have a large scope and tap different facets of IH (Porter et al., 2021). The SIHS showed the best discriminant validity, as it was distinct to all other personality constructs assessed. The LIHS and CIHS were distinct to all constructs except slightly higher-than-expected correlations with social desirability. In contrast, the IHS

3) As suggested by an anonymous reviewer, we also tested the predictive validity of the four IH scales by running several multiple regressions with all four IH scales as predictors, and Need for Cognition, Need for Cognitive Closure, Openness to Experience, and Honesty-Humility as dependent variables. These analyses can be found in the Supplementary Materials "additional_analyses.pdf", Table 2.

was not distinct from several psychological constructs (e.g., NCC, NFC, honesty-humility) and demographic variables, thus not meeting our pre-registered cut-off criteria.

Regarding incremental validity, the SIHS predicted less affective polarization towards opinion-based outgroups over and above the other psychological constructs. This effect was driven by warmer feelings towards the political outgroup. Of the general IH scales, only the Intellectual Modesty subscale of the IHS predicted less affective polarization towards political parties. The LIHS, the CIHS subscales as well as the other IH subscales did not predict less affective polarization. In contrast, the CIHS and IHS overall means as well as the IHS subfacet Open-Mindedness predicted slightly more affective polarization.

Taking these results together, we recommend using the CIHS, LIHS or SIHS when studying IH in Germany. The existing scale in German (IHS) did not show adequate validities, potentially because it was originally translated and validated in Switzerland, a different cultural context with differences in the spoken German. We found evidence for this when the item “a disagreement is like a war” was particularly problematic when analyzing structural validity. Because of Germany’s perpetrator role in the second world war, this item might have irritated participants. Therefore, future work is needed to examine whether our German versions of the CIHS, LIHS, and SIHS can be used in different German-speaking contexts, e.g., Switzerland or Austria.

Contributions

Our first aim was to provide high-quality translations in German of three widely-used English IH scales (CIHS, LIHS, SIHS). Thanks to independent item translations by two researchers, a blind evaluation of the translation by $N = 8$ experts and German natives as well as a pre-test ($N = 13$), we can now offer IH items in German to the research community. This is especially important as the only existing IH scale in German (IHS) did not show adequate structural and discriminant validity in our study. Thereby, we want to facilitate studying IH across countries and languages, addressing the call for more cross-cultural research on this topic (Porter et al., 2022).

A second aim was to validate four scales with different scopes and content. Here, we made use of the intellectual humility classification framework (Porter et al., 2021) to evaluate the chosen scales. Two of the scales (CIHS, IHS) are multifaceted scales covering both (meta-)cognitions and behaviors towards the self and others. The LIHS only includes (meta-)cognitive items towards the self and others and the SIHS is the only scale covering (meta-)cognitive and self-directed items regarding a specific (political) topic. With this variety of scope and content, researchers can choose the scale that matches their research aims best (Porter et al., 2022).

Third, we compared four different established IH scales in parallel. This allowed us to systematically compare the scales to other psychological and potentially confounded constructs measured on the same instruments. The original validation studies (Alfano et al., 2017; Hoyle et al., 2016; Krumrei-Mancuso & Rouse, 2016; Leary et al., 2017) used

different scales to assess discriminant validity, making a direct comparison between the scales more difficult. In contrast, we only included established scales validated in the German context, and we were therefore able to directly compare the IH scales regarding their discriminant validity.

Limitations and Future Directions

Our research also has several limitations that can inspire future research. Based on the insights of the study as well as its strengths and limitations, we suggest several avenues for further research on IH assessment.

First, all of the chosen IH measures were self-report questionnaires. As these are a valid tool to assess theoretical constructs and used commonly in psychological research (Haefel & Howald, 2010), our validation of these scales addresses the needs of many researchers. However, it can be beneficial to develop behavioral measures of IH (Van Tongeren et al., 2022a, 2022b), especially when being interested in measuring IH in social interactions (Hanel et al., 2023; Meagher et al., 2020). Here, it would be fascinating to see how different self-report measures, especially those with a rather (meta-)cognitive scope, can predict intellectually humble behaviors. Previous research showed that intellectual humility in conversation (e.g., interrupting others less, more often conveying one's opinion as tentative) coded by linguists was not associated with the CIHS (Hanel et al., 2023), potentially due to the gap between cognitions and behaviors (Minson & Chen, 2021; Van Tongeren et al., 2022a, 2022b). It would be of interest to see if this also applies to other intellectual humility measures, such as the specific intellectual humility scale.

Second, the IH scales chosen for validation conceptualize IH rather as a trait than as a state (Ballantyne, 2023; Du & Cai, 2020). This enables researchers to reliably study between-person variations of IH, and its associations with political constructs (e.g., Bowes et al., 2020; Krumrei-Mancuso & Newman, 2020). However, there might be substantive within-person variation of IH, for instance depending on the norms of a given context (Porter & Cimpian, 2023) or the availability of mental resources in a specific situation (Hanel et al., 2023). Here, it would be beneficial to develop and validate additional state-like IH measures as Zachry et al. (2018) did. This would tap potential within-person variations of IH better than the existing IH measures and enable research investigating in which contexts people show IH.

Third, despite several methodological strengths, there are still some aspects that we could not implement within our study. Future IH validation research could address these open questions. For instance, we only assessed the constructs used for evaluating discriminant and incremental validity after measuring the IH scales. We decided to do so because we did not want the participants' responses on the IH scales to be impacted by having answered these questionnaires before. However, filling out the IH scales might have influenced answers on the discriminant scales, too. Thus, future experimental research could systematically assess order effects of when IH is measured in a survey.

Fourth, the assessment of convergent validity in our study is limited. Ideally, convergent validity of a new measurement instruments is assessed by comparing the new instrument to existing validated scales. However, no IH measurement instrument (or closely related scale such as open-minded cognition) existed in the German context when conducting the study. We therefore assessed convergent validity by correlating the different IH scales with each other. This procedure leaves the possibility that none of our translated scales measures IH very well. However, we evaluate this scenario as relatively unlikely given that all scales used were already validated in an English-speaking context. Additionally, our translated IH scales correlated with a higher need for cognition and a lower need for cognitive closure, which also hint at convergent validity. Nonetheless, future work is needed to assess convergent validity of our IH scales, for instance with behavioral data (Van Tongeren et al., 2022a, 2022b), or text-based information (Stavropoulos et al., 2024).

Lastly, as IH is a promising psychological construct to address societal issues, we assessed incremental in relation to affective polarization. Our findings showed that the SIHS had the highest incremental validity to predict lower affective polarization. This might be due to the fact that for this scale, we measured affective polarization regarding opinion-based groups. Regarding the three general IH scales, we measured affective polarization via political parties where the outgroup was chosen via the least-liked group paradigm (Gibson, 2013). There might be other paradigms, however, which are better suited to assess affective polarization in the context of multiparty-systems (Harteveld, 2021; Wagner, 2021) such as Germany. To sum up, research is just at the starting point to understand how IH is related to affective polarization and how it can bridge political divides. Here, a fascinating avenue might be to study affective polarization in more diverse ways or to use specific behaviors and behavioral intentions such as approaching contrary-minded others (Knöchelmann & Cohrs, 2025).

Conclusion

IH is a psychological construct relevant to addressing societal problems such as bridging political divides. It is associated with many positive outcomes such as higher well-being, or learning outcomes at school. To contribute to more research on IH outside English-speaking contexts, we validated and compared four established IH scales with different content and scope in German. Our sample was representative of age, gender, and education of Germany's population and the study was pre-registered and peer-reviewed before data collection. Results showed that the German versions of the CIHS, LIHS, and SIHS can be compared to the original English scales, whereas the IHS did not meet all of our pre-registered criteria, potentially due to cross-cultural differences between Germany and Switzerland. We hope that our work facilitates more cross-cultural work on IH.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 290878970-GRK 2271, project 13. Data collection of the present study was funded by PsychLab, a service of the Leibniz Institute for Psychology

Acknowledgments: We would like to thank all students and colleagues who supported the research by the preparation of materials, data collection, as well as discussions and feedback on earlier versions of this manuscript: Michaela Bölinger, Christiane Böttcher, Rebecca Cramer, Daniel Kühlewind, Victoria Le, Carmen Lienen, Linus Peitz, Adrian Rothers, Marlene Swietlik.

Competing Interests: The authors have declared that no competing interests exist.

Author Contributions: Larissa Knöchelmann—Conceptualization | Data curation | Formal analysis | Funding acquisition | Investigation | Methodology | Project administration | Supervision | Validation | Visualization | Writing – original draft. Jule Anna Kemmer—Conceptualization | Investigation | Methodology | Validation | Writing – review & editing. J. Christopher Cohrs—Funding acquisition | Resources | Supervision | Writing – review & editing.

Ethics Statement: The study was approved by the University of Marburg ethics committee on 24th September 2021 (file reference number ‘2021-39k’). APA ethical standards as well as recommendations by the German psychological society (DGPs) were followed in the conduct of all studies. All participants gave informed consent before participation.

Preregistration Statement: Our pre-registration is available at <https://doi.org/10.23668/psycharchives.5202>. The pre-registration went through a blind peer-review before data collection thanks to Leibniz Institute for Psychology (ZPID).

Data Availability: Data and analyses are available (see [Knöchelmann et al., 2024](#)).

Supplementary Materials

For this article, the following Supplementary Materials are available:

- Pre-registration (see [Knöchelmann et al., 2021](#))
- Data and analyses (see [Knöchelmann, 2024](#))

Index of Supplementary Materials

Knöchelmann, L., Kemmer, J., & Cohrs, C. (2021). *Intellectual humility: Scale validation in German-speaking contexts* [Pre-registration]. PsychArchives.

<https://doi.org/10.23668/psycharchives.5202>

Knöchelmann, L. (2024). *Intellectual humility - Validation and comparison of four self-report scales in Germany* [Data, analyses]. OSF. <https://doi.org/10.17605/OSF.IO/VF6KJ>

References

- Aertssen, W. F. M., Ferguson, G. D., & Smits-Engelsman, B. C. M. (2016). Reliability and structural and construct validity of the functional strength measurement in children aged 4 to 10 years. *Physical Therapy, 96*(6), 888–897. <https://doi.org/10.2522/ptj.20140018>
- Alfano, M., Iurino, K., Stey, P., Robinson, B., Christen, M., Yu, F., & Lapsley, D. (2017). Development and validation of a multi-dimensional measure of intellectual humility. *PLoS One, 12*(8), Article e0182950. <https://doi.org/10.1371/journal.pone.0182950>
- Alfano, M., Lynch, M. P., & Tanesini, A. (Eds.). (2020). *The Routledge handbook of philosophy of humility*. Routledge.
- American Psychological Association. (n.d.). *Incremental validity*. APA Dictionary of Psychology. Retrieved February 2, 2023, from <https://dictionary.apa.org/incremental-validity>
- Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*(4), 340–345. <https://doi.org/10.1080/00223890902935878>
- Ashton, M. C., Lee, K., & de Vries, R. E. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality Factors: A Review of Research and Theory. *Personality and Social Psychology Review, 18*(2), 139–152. <https://doi.org/10.1177/1088868314523838>
- Ballantyne, N. (2023). Recent work on intellectual humility: A philosopher's perspective. *The Journal of Positive Psychology, 18*(2), 200–220. <https://doi.org/10.1080/17439760.2021.1940252>
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1–26. <https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Bar-Tal, D., & Halperin, E. (2011). Socio-psychological barriers to conflict resolution. In D. Bar-Tal (Ed.), *Intergroup conflicts and their resolution: A social psychological perspective* (pp. 217–240). Psychology Press.
- Beißert, H., Köhler, M., Rempel, M., & Beierlein, C. (2014). *Eine deutschsprachige Kurzsкала zur Messung des Konstrukts Need for Cognition: Die Need for Cognition Kurzsкала (NFC-K)* (GESIS-Working Papers, 2014/32). GESIS - Leibniz-Institut für Sozialwissenschaften. <https://www.ssoar.info/ssoar/handle/document/40315>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health, 6*, Article 149. <https://doi.org/10.3389/fpubh.2018.00149>
- Bowes, S. M., Blanchard, M. C., Costello, T. H., Abramowitz, A. I., & Lilienfeld, S. O. (2020). Intellectual humility and between-party animus: Implications for affective polarization in two community samples. *Journal of Research in Personality, 88*, Article 103992. <https://doi.org/10.1016/j.jrp.2020.103992>
- Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment, 48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

- Church, I. M., & Barrett, J. L. (2016). Intellectual humility. In E. L. Worthington Jr, D. E. Davis, & J. N. Hook (Eds.), *Handbook of humility* (pp. 78–91). Routledge.
<https://doi.org/10.4324/9781315660462>
- DESTATIS. (2021a). *Bevölkerung ab 15 Jahren in Hauptwohnsitzhaushalten: Deutschland, Jahre, Geschlecht, Altersgruppen, Allgemeine Schulausbildung* (12211–0100) [Dataset]. Statistisches Bundesamt Deutschland GENESIS-Online. <https://www-genesis.destatis.de/genesis/online>
- DESTATIS. (2021b). *Bevölkerung: Deutschland, Stichtag, Geschlecht, Altersgruppen, Staatsangehörigkeit* (12411–0009) [Dataset]. Statistisches Bundesamt Deutschland GENESIS-Online. <https://www-genesis.destatis.de/genesis/online>
- Du, J., & Cai, Y. (2020). Owning one's intellectual limitations: A review of intellectual humility. *Psychology (Irvine, Calif.)*, 11(7), 1009–1020. <https://doi.org/10.4236/psych.2020.117066>
- Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, 3(4), 456–465. <https://doi.org/10.1177/2515245920952393>
- Gibson, J. L. (2013). Measuring political tolerance and general support for pro-civil liberties policies: Notes, evidence, and cautions. *Public Opinion Quarterly*, 77(S1), 45–68.
<https://doi.org/10.1093/poq/nfs073>
- Gorichanaz, T. (2022). Relating information seeking and use to intellectual humility. *Journal of the Association for Information Science and Technology*, 73(5), 643–654.
<https://doi.org/10.1002/asi.24567>
- Haefel, G. J., & Howald, G. S. (2010). Self-report: Psychology's four-letter word. *The American Journal of Psychology*, 123(2), 181–188. <https://doi.org/10.5406/amerjpsyc.123.2.0181>
- Hanel, P., Roy, D., Taylor, S., Franjeh, M., Heffer, C., Tanesini, A., & Maio, G. (2023). Using Self-Affirmation to Increase Intellectual Humility in Debate. *Royal Society Open Science*, 10(2), Article 220958. <https://doi.org/10.1098/rsos.220958>
- Harteveld, E. (2021). Fragmented foes: Affective polarization in the multiparty context of the Netherlands. *Electoral Studies*, 71, Article 102332. <https://doi.org/10.1016/j.electstud.2021.102332>
- Hodge, A. S., Hook, J. N., Van Tongeren, D. R., Davis, D. E., & McElroy-Heltzel, S. E. (2021). Political humility: Engaging others with different political perspectives. *The Journal of Positive Psychology*, 16(4), 526–535. <https://doi.org/10.1080/17439760.2020.1752784>
- Hoyle, R. H., Davison, E. K., Diebels, K. J., & Leary, M. R. (2016). Holding specific views with humility: Conceptualization and measurement of specific intellectual humility. *Personality and Individual Differences*, 97, 165–172. <https://doi.org/10.1016/j.paid.2016.03.043>
- Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, 15(4), 446–455.
<https://doi.org/10.1037/1040-3590.15.4.446>
- International Test Commission. (2018). ITC guidelines for translating and adapting tests (second edition). *International Journal of Testing*, 18(2), 101–134.
<https://doi.org/10.1080/15305058.2017.1398166>

- Jost, J. T., Baldassarri, D. S., & Druckman, J. N. (2022). Cognitive–motivational mechanisms of political polarization in social-communicative contexts. *Nature Reviews Psychology*, *1*(10), 560–576. <https://doi.org/10.1038/s44159-022-00093-5>
- Kemper, C. J., Beierlein, C., Bensch, D., Kovaleva, A., & Rammstedt, B. (2012). *Eine Kurzskala zur Erfassung des Gamma-Faktors sozial erwünschten Antwortverhaltens: Die Kurzskala Soziale Erwünschtheit-Gamma (KSE-G)* [(GESIS-Working Papers, 2012/25)]. <https://www.ssoar.info/ssoar/handle/document/33958>
- Knöchelmann, L., & Cohrs, J. C. (2025). Effects of intellectual humility in the context of affective polarization: Approaching and avoiding others in controversial political discussions. *Journal of Personality and Social Psychology*, *129*(1), 91–117. <https://doi.org/10.1037/pspi0000462>
- Krumrei-Mancuso, E. J., Haggard, M. C., LaBouff, J. P., & Rowatt, W. C. (2020). Links between intellectual humility and acquiring knowledge. *The Journal of Positive Psychology*, *15*(2), 155–170. <https://doi.org/10.1080/17439760.2019.1579359>
- Krumrei-Mancuso, E. J., & Newman, B. (2020). Intellectual humility in the sociopolitical domain. *Self and Identity*, *19*(8), 989–1016. <https://doi.org/10.1080/15298868.2020.1714711>
- Krumrei-Mancuso, E. J., & Newman, B. (2021). Sociopolitical intellectual humility as a predictor of political attitudes and behavioral intentions. *Journal of Social and Political Psychology*, *9*(1), 52–68. <https://doi.org/10.5964/jsp.5553>
- Krumrei-Mancuso, E. J., & Rouse, S. V. (2016). The Development and Validation of the Comprehensive Intellectual Humility Scale. *Journal of Personality Assessment*, *98*(2), 209–221. <https://doi.org/10.1080/00223891.2015.1068174>
- Leary, M. R., Diebels, K. J., Davisson, E. K., Jongman-Sereno, K. P., Isherwood, J. C., Raimi, K. T., Deffler, S. A., & Hoyle, R. H. (2017). Cognitive and Interpersonal Features of Intellectual Humility. *Personality & Social Psychology Bulletin*, *43*(6), 793–813. <https://doi.org/10.1177/0146167217697695>
- McElroy, S. E., Rice, K. G., Davis, D. E., Hook, J. N., Hill, P. C., Worthington, E. L., & Van Tongeren, D. R. (2014). Intellectual humility: Scale development and theoretical elaborations in the context of religious leadership. *Journal of Psychology and Theology*, *42*(1), 19–30. <https://doi.org/10.1177/009164711404200103>
- Meagher, B. R., Leman, J. C., Bias, J. P., Latendresse, S. J., & Rowatt, W. C. (2015). Contrasting self-report and consensus ratings of intellectual humility and arrogance. *Journal of Research in Personality*, *58*, 35–45. <https://doi.org/10.1016/j.jrp.2015.07.002>
- Meagher, B. R., Leman, J. C., Heidenga, C. A., Ringquist, M. R., & Rowatt, W. C. (2020). Intellectual humility in conversation: Distinct behavioral indicators of self and peer ratings. *The Journal of Positive Psychology*, *16*(3), 417–429. <https://doi.org/10.1080/17439760.2020.1738536>
- Minson, J. A., & Chen, F. S. (2021). Receptiveness to opposing views: Conceptualization and integrative review. *Personality and Social Psychology Review*, *26*(2), 93–111. <https://doi.org/10.1177/10888683211061037>
- Moshagen, M., & Erdfelder, E. (2016). A new strategy for testing structural equation models. *Structural Equation Modeling*, *23*(1), 54–60. <https://doi.org/10.1080/10705511.2014.950896>

- Moshagen, M., Hilbig, B. E., & Zettler, I. (2014). Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschsprachigen Version des 60-Item HEXACO Persönlichkeitsinventars. *Diagnostica*, *60*(2), 86–97. <https://doi.org/10.1026/0012-1924/a000112>
- Paulhus, D. L. (2017). Socially desirable responding on self-reports. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (Vol. 1, pp. 1–5). Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_1349-1
- Pavlov, G., Shi, D., & Maydeu-Olivares, A. (2020). Chi-square difference tests for comparing nested models: An evaluation with non-normal data. *Structural Equation Modeling*, *27*(6), 908–917. <https://doi.org/10.1080/10705511.2020.1717957>
- Porter, T., Baldwin, C. R., Warren, M. T., Murray, E. D., Cotton Bronk, K., Forgeard, M. J. C., Snow, N. E., & Jayawickreme, E. (2021). Clarifying the content of intellectual humility: A systematic review and integrative framework. *Journal of Personality Assessment*, *104*(5), 573–585. <https://doi.org/10.1080/00223891.2021.1975725>
- Porter, T., & Cimpian, A. (2023). A context's emphasis on intellectual ability discourages expression of intellectual humility. *Motivation Science*, *9*(2), 120–130. <https://doi.org/10.1037/mot0000289>
- Porter, T., Elnakouri, A., Meyers, E. A., Shibayama, T., Jayawickreme, E., & Grossmann, I. (2022). Predictors and consequences of intellectual humility. *Nature Reviews Psychology*, *1*(9), 524–536. <https://doi.org/10.1038/s44159-022-00081-9>
- Porter, T., & Schumann, K. (2018). Intellectual humility and openness to the opposing view. *Self and Identity*, *17*(2), 139–162. <https://doi.org/10.1080/15298868.2017.1361861>
- Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica*, *51*(4), 195–206. <https://doi.org/10.1026/0012-1924.51.4.195>
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Sage Publications. <https://psycnet.apa.org/record/1996-97111-016>
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.
- Schlink, S., & Walther, E. (2007). Kurz und gut: Eine deutsche Kurzskaala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit. *Zeitschrift für Sozialpsychologie*, *38*(3), 153–161. <https://doi.org/10.1024/0044-3514.38.3.153>
- Stavropoulos, A., Crone, D. L., & Grossmann, I. (2024). Shadows of wisdom: Classifying meta-cognitive and morally grounded narrative content via large language models. *Behavior Research Methods*, *56*, 7632–7646. <https://doi.org/10.3758/s13428-024-02441-0>
- Van Tongeren, D. R., Ng, V., Hickman, L., & Tay, L. (2022a). Behavioral measures of humility: Part 1. Theoretical and methodological review. *The Journal of Positive Psychology*, *18*(5), 711–721. <https://doi.org/10.1080/17439760.2022.2109202>

- Van Tongeren, D. R., Ng, V., Hickman, L., & Tay, L. (2022b). Behavioral measures of humility: Part 2. Conceptual mapping and charting ways forward. *The Journal of Positive Psychology, 18*(5), 722–732. <https://doi.org/10.1080/17439760.2022.2109201>
- Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies, 69*, Article 102199. <https://doi.org/10.1016/j.electstud.2020.102199>
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*(6), 1049–1062. <https://doi.org/10.1037/0022-3514.67.6.1049>
- Weijters, B., Baumgartner, H., & Schillewaert, N. (2013). Reversed item bias: An integrative model. *Psychological Methods, 18*(3), 320–334. <https://doi.org/10.1037/a0032121>
- Wojcieszak, M., & Warner, B. R. (2020). Can interparty contact reduce affective polarization? A systematic test of different forms of intergroup contact. *Political Communication, 37*(6), 789–811. <https://doi.org/10.1080/10584609.2020.1760406>
- Zachry, C. E., Phan, L. V., Blackie, L. E. R., & Jayawickreme, E. (2018). Situation-based contingencies underlying wisdom-content manifestations: Examining intellectual humility in daily life. *The Journals of Gerontology: Series B, 73*(8), 1404–1415. <https://doi.org/10.1093/geronb/gby016>
- Zmigrod, L., Zmigrod, S., Rentfrow, P. J., & Robbins, T. W. (2019). The psychological roots of intellectual humility: The role of intelligence and cognitive flexibility. *Personality and Individual Differences, 141*, 200–208. <https://doi.org/10.1016/j.paid.2019.01.016>

PSPS Polish
Social
Psychological
Society

Social Psychological Bulletin (SPB) is an official journal of the Polish Social Psychological Society (PSPS).

zpid ^{LP}

Leibniz-Institut für
Psychologie

PsychOpen GOLD is a publishing service provided by the Leibniz Institute for Psychology (ZPID), Germany.