

Funny selves: Development of the Humor Efficacy and Identity Short Scales (HEISS)

By: [Paul J. Silvia](#), Rebekah M. Rodriguez, and Maciej Karwowski

Silvia, P.J., Rodriguez, R.M., & Karwowski, M. Funny selves: Development of the Humor Efficacy and Identity Short Scales (HEISS). *Personality and Individual Differences* 182, 111093. <https://doi.org/10.1016/j.paid.2021.111093>



© 2021 Elsevier Ltd. This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#).

Abstract:

Although humor is a universal feature of human communication, people vary widely in how they create and use humor. Guided by a broader model of creative self-beliefs, we developed the Humor Efficacy and Identity Short Scales (HEISS), a pair of 4-item scales measuring humor self-efficacy (“I can” beliefs reflecting confidence about one's ability to be funny) and humor identity (“I am” beliefs reflecting the centrality of humor ability to one's self-concept). Using a large sample of English speakers ($n = 1842$), an item response theory analysis found a suitable range of item difficulty, good item discrimination, and essentially zero gender-based differential item functioning. Three follow-up samples with English ($n = 304$, $n = 400$) and Polish ($n = 385$) speakers found conceptually consistent relationships with humor backgrounds and experiences (e.g., taking classes and holding jobs involving humor), with Big 5 personality traits, and with humor styles and playfulness. Taken together, these scales show promise for research on people's humor self-concepts and for studies of gendered aspects of humor use, creation, and appreciation.

Keywords: humor | self-efficacy | identity | self-concept | creativity

Article:

1. Introduction

Humor is a universal feature of human communication, used for a wide range of social purposes from persuasion to entertainment, aggression to inclusion, and ingratiation to disparagement (Lynch, 2002; Martin & Ford, 2018; Walter et al., 2018; Weinstein et al., 2011). Nevertheless, people vary widely in humor's many facets, such as the kinds of humor they find funny, their typical goals for using humor in interactions, and their ability to generate funny ideas (Martin et al., 2003; Plessen et al., 2020; Ruch & Heintz, 2019). In the present research, we draw upon the broader study of creativity to examine people's *humor self-concepts*—their beliefs about their humor abilities, particularly their *humor self-efficacy* (“I can” beliefs reflecting confidence about the ability to be funny) and *humor identity* (“I am” beliefs reflecting the centrality of being funny to their self-concept). In four samples, brief scales—the Humor Efficacy and Identity Short Scales (HEISS)—that measure these two aspects of humor self-concepts were developed and evaluated.

1.1. Creativity and humor production

The ability to create humor—to come up with ideas and practices that make others laugh—is a growing topic in the psychology of creativity. Humor creation dovetails naturally with popular theories and themes in creativity research (Lu et al., 2019; Perchtold-Stefan et al., 2020; Ruch & Heintz, 2019), which has long been concerned with how people generate ideas that are both new and apt for a particular purpose, audience, or context (Diedrich et al., 2015; Sawyer, 2006). In this view, humor production falls under the broader umbrella of creative cognition. The largest area of research at the intersection of creativity and humor is probably the study of individual differences in humor production: the ability to produce funny ideas (Ruch & Heintz, 2019). Humor production ability has been linked to personality traits (e.g., Openness to Experience, Extraversion, and Right-Wing Authoritarianism; Nusbaum & Silvia, 2017; Silvia et al., 2021a), intelligence (e.g., fluid and crystallized intelligence promote humor; Christensen et al., 2018; Kellner & Benedek, 2017), and gender (men show a small advantage; Greengross et al., 2020).

Applying a creativity lens to humor production suggests that many influential concepts and models from creativity research might be fruitful for the study of humor as well. In particular, the study of creativity distinguishes between individual differences in underlying creative abilities (e.g., the ability to generate creative responses to laboratory tasks measuring divergent thinking) and in people's beliefs about their abilities (Kaufman, 2012; Snyder et al., 2020). This long-standing distinction reflects the simple fact that people with the same ability level—be it intelligence, athleticism, or creativity—can nevertheless have different beliefs about their abilities.

In creativity research, the study of creative self-concepts has focused on two of these beliefs: *creative self-efficacy* and *creative identity* (Beghetto & Karwowski, 2017; Karwowski & Kaufman, 2017). Creative self-efficacy reflects confidence in one's creative abilities; creative identity reflects the centrality of being a creative person to one's self-concept. The large literature that has developed around creative self-concepts shows that people's beliefs about their own creativity play crucial roles in how they select, pursue, and abandon creative goals (Karwowski & Kaufman, 2017; Tierney & Farmer, 2002), consistent with the influential social-cognitive model that undergirds it (Bandura, 1997).

The creative self-concepts framework could be adapted to the specific domain of humor. Individual differences in humor production abilities have been widely studied (Nusbaum & Silvia, 2017; Ruch & Heintz, 2019), but little is known about people's beliefs about their humor abilities. Adapting creative self-concepts for humor yields two parallel concepts. The first is *humor self-efficacy*, people's reasonably stable confidence beliefs about their ability to be funny, to make others laugh, and to come up with ideas that are clever, witty, and humorous. The second is *humor identity*, people's perceptions of how their humor-related activities, abilities, and goals connect to their overall sense of self. Humor self-efficacy primarily captures confidence beliefs about one's humor ability (“I can”); humor identity primarily captures the centrality of humor ability to one's self-concept (“I am”).

As the larger social-cognitive literature shows, people's efficacy and identity beliefs play profound roles in their goals and aspirations, their motivation to achieve them, their reactions to setbacks and barriers, and their eventual success (Bandura, 1997). The concepts of humor efficacy and identity could thus aid in illuminating individual differences in how often and how effectively people use and create humor. For example, as with other abilities, people can be overconfident or underconfident in their humor abilities, and these confidence beliefs could shape if, when, and how people produce humor in their interpersonal interactions. Similarly, people vary in how important it is to them to be a funny person, and the centrality of humor to their identity could shape their preferences for different activities, interaction partners, hobbies, and careers.

1.2. The present research

In the present research, we translate a model of creative self-concepts to the domain of humor and develop brief scales to assess two fundamental components of people's humor self-concepts. Using item response theory methods, we developed scales—the Humor Efficacy and Identity Short Scales (HEISS)—to create tools for researchers interested in individual differences in humor self-concepts. Our goal was to create brief scales that could be efficiently incorporated into survey and experimental research and that could stand alone if researchers wished to focus on only one construct. Because gender is a pervasive variable in humor research (Greengross, 2020; Hofmann et al., 2020; Martin, 2014), a particular focus was to develop items with minimal gender-based differential item functioning (DIF) so that researchers could have more confidence that possible gender differences in the scores reflect real underlying trait differences.

2. Study 1

In Study 1, we describe the development of the HEISS, present an item response theory analysis of the items, and evaluate the items for gender-based DIF.

2.1. Method

2.1.1. *Participants*

All participants in all the studies reported here provided informed consent, and the research was approved by our institutions' respective ethics committees. In Study 1, a final sample of 1842 adults—929 women, 913 men—who ranged in age from 18 to 88 years old ($M = 35.66$ years, $SD = 13.81$, $Mdn = 33$) took part. Participants were recruited from the Prolific.co online survey panel and were eligible if they were at least 18 years old, spoke English as a native language, and identified as female or male. The final sample was refined from a larger sample of 2100 people, for an exclusion rate of around 12%. The details of the sample exclusions are in the online supplementary material (OSM).

2.1.2. *Development process*

An initial item pool of 24 items was created by the three authors, who brought expertise in the assessment of humor, personality, and creative self-efficacy and identity. Item generation was

guided by construct definitions that were anchored in the literature on creative self-efficacy and identity to ensure that humor efficacy and identity were aligned with the broader model. We then collected response data across two waves. The first wave (roughly half of the sample) responded to the full pool of 24 items. Items were then excluded based on descriptive statistics (e.g., items at the floor or ceiling of the response scale), IRT and DIF analyses (e.g., omitting extremely easy or hard items, and omitting items showing gender-based DIF), and wording overlap (e.g. omitting items to avoid highly similar phrasing). The second half of the sample then completed a smaller pool of 12 promising items. The data for these 12 items were combined for the full sample to select and evaluate the final 8 items.

The final items were selected based on the same psychometric criteria—high discrimination parameters, at least moderate item difficulty, and a good spread of category endorsement—as well as practical criteria, such as varied item wording and ease of translation from English into other languages. A requirement for inclusion was that an item had essentially zero gender DIF. Because we anticipated that many researchers may wish to use only one of the subscales, the efficacy and identity scales were developed and evaluated as separate instruments, ensuring that each could stand alone if necessary.

The final 8 items are presented in Table 1. In addition to the HEISS, participants also noted their age (in years) and were asked “How would you describe your gender?” People who selected “Female” (coded 1) or “Male” (coded 0) were retained; because DIF analyses require large subgroup samples, the small handful of people who declined to state their gender or who entered a free-response description were omitted.

Table 1. Humor Efficacy and Identity Short Scales (HEISS).

<p>We're interested in people's ideas and beliefs about <i>being funny</i>: making other people laugh and coming up with witty and humorous ideas. How much do you agree or disagree with the following statements about you?</p> <p>[eff1] I think I can make almost anyone laugh. [eff2] I trust my ability to be funny. [eff3] I feel confident in my humor skills. [eff4] Being funny is something that comes naturally to me. [id1] Being a funny person is a big part of who I am. [id2] It's important to me to be a funny person. [id3] My humor ability is central to who I am. [id4] Making people laugh is important to me.</p>
--

Note. The items are completed on a 1-5 scale using *strongly disagree* and *strongly agree* as anchors. The items should be presented in a random order. We recommend including a directed response item (“For this item, please click “strongly disagree””) and a reverse-coded trap item (e.g., “I’m not a funny person”) to catch long-string responding. The Polish translation is available in the online supplementary material. The Open Science Framework archive has Qualtrics versions for importing directly into surveys.

2.2. Results

2.2.1. Item response theory models

The analyses were conducted in R 4.1 (R Core Team, 2021) using *mirt* (Chalmers, 2012, Chalmers, 2020) and *psych* (Revelle, 2021). Each subscale was unidimensional according

to parallel analyses of polychoric correlations. We conducted graded response models to evaluate and select the items. These models yield estimates of each item's slope/discrimination and boundary location parameters. Analyses of item misfit found no evidence for underfit on infit and outfit metrics. Analyses of residuals found at most small local item dependence (all standardized G^2 values less than Cramer's signed $V = \pm 0.13$; Chen & Thissen, 1997).

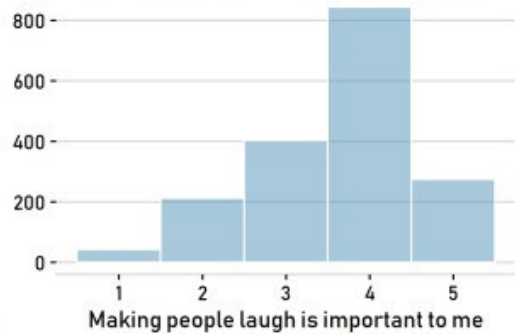
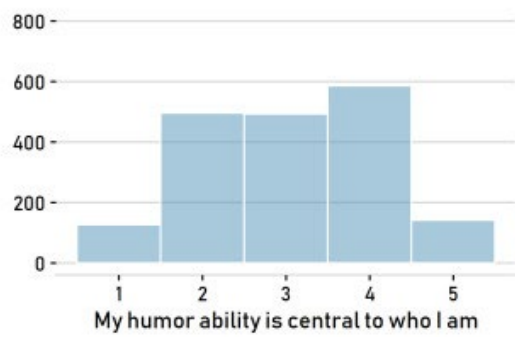
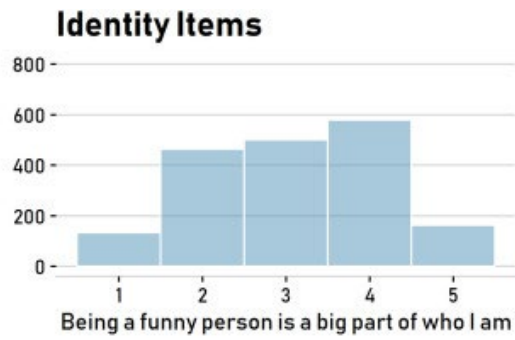
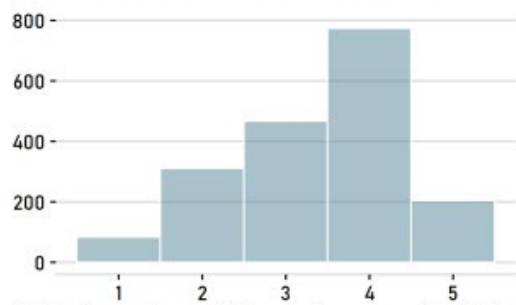
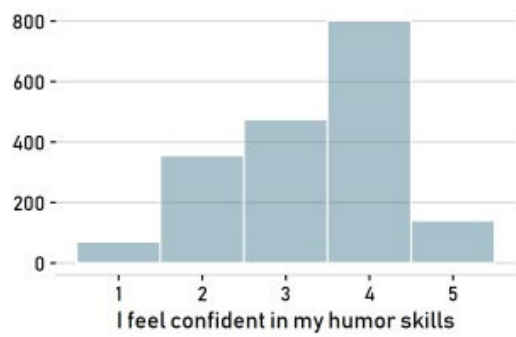
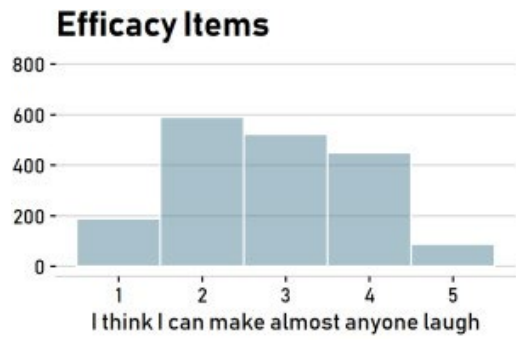
DIF was evaluated using the logistic ordinal regression approach implemented in *lordif* (Choi et al., 2016), which can estimate both uniform and non-uniform DIF (Choi et al., 2011). We used effect size measures (Jodoin & Gierl, 2001; Meade, 2010), particularly McFadden's R^2 (Menard, 2000). Values below $R^2 = 0.02$ (the common cutoff for a small R^2) or 0.01 have been suggested as strict cutoffs (Choi et al., 2011; Crane et al., 2007) for flagging DIF items. Because gender differences are widespread in humor research, we used an especially strict threshold of 0.005 (half of 1%) to identify items for total DIF. None of the efficacy or identity items were flagged, so we can conclude that gender DIF is negligible in this large sample.

Table 2 shows the item parameters, and Fig. 1 shows the distributions of scores for each item. The items had at least good levels of discrimination and boundary locations that captured a reasonably large trait range. The test information functions are shown in the OSM. Trait scores for humor identity and efficacy were computed via expected a posteriori (EAP) scores from the IRT model. The EAP reliability was very high for both efficacy (0.90) and identity (0.89), suggesting strong evidence for the reliability of the scales' scores. Fig. 2 displays the distributions for the humor efficacy and identity trait scores. Shown in Fig. 3, these scores were highly correlated ($r = 0.73$ [0.71, 0.75]), as expected.

Table 2. Item statistics: Study 1.

Item	<i>M</i>	<i>SD</i>	<i>Mdn</i>	<i>Skew</i>	<i>Kurtosis</i>	Discrimination							
						(<i>a</i>)	<i>b</i> ₁	<i>b</i> ₂	<i>b</i> ₃	<i>b</i> ₄	Infit	Outfit	
[eff1] I think I can make almost anyone laugh.	2.81	1.06	3	0.10	-0.80	2.17	-1.64	-0.24	0.69	2.14	0.875	0.852	
[eff2] I trust my ability to be funny.	3.32	0.97	4	-0.51	-0.52	4.56	-1.99	-0.77	-0.05	1.62	0.685	0.570	
[eff3] I feel confident in my humor skills.	3.32	0.99	4	-0.44	-0.53	3.81	-2.01	-0.79	-0.02	1.55	0.757	0.698	
[eff4] Being funny is something that comes naturally to me.	3.38	1.04	4	-0.46	-0.46	3.18	-1.98	-0.87	-0.07	1.37	0.812	0.770	
[id1] Being a funny person is a big part of who I am.	3.09	1.10	3	-0.12	-0.83	3.91	-1.59	-0.50	0.27	1.47	0.662	0.619	
[id2] It's important to me to be a funny person.	3.26	1.01	3	-0.35	-0.51	3.06	-1.91	-0.82	0.11	1.61	0.787	0.761	
[id3] My humor ability is central to who I am.	3.07	1.08	3	-0.09	-0.85	2.79	-1.75	-0.49	0.31	1.68	0.801	0.791	
[id4] Making people laugh is important to me.	3.63	0.95	4	-0.66	0.06	2.52	-2.45	-1.30	-0.43	1.25	0.836	0.805	

Note. $n = 1842$. The items were completed on a 1-5 scale. "Discrimination" refers to the IRT discrimination a parameter; the b_1 - b_4 values are the graded response model boundary locations for the 5 response options.



Being funny is something that comes naturally to me
Fig. 1. Item category responses.

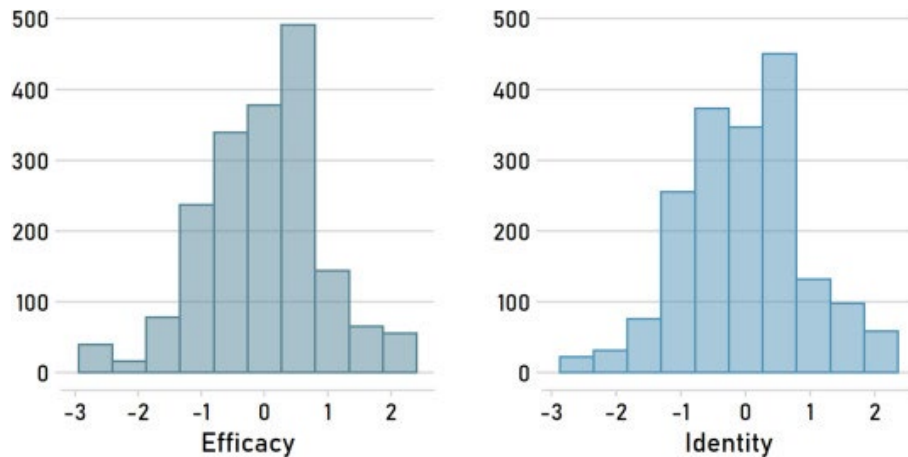


Fig. 2. Distributions of humor efficacy and identity.

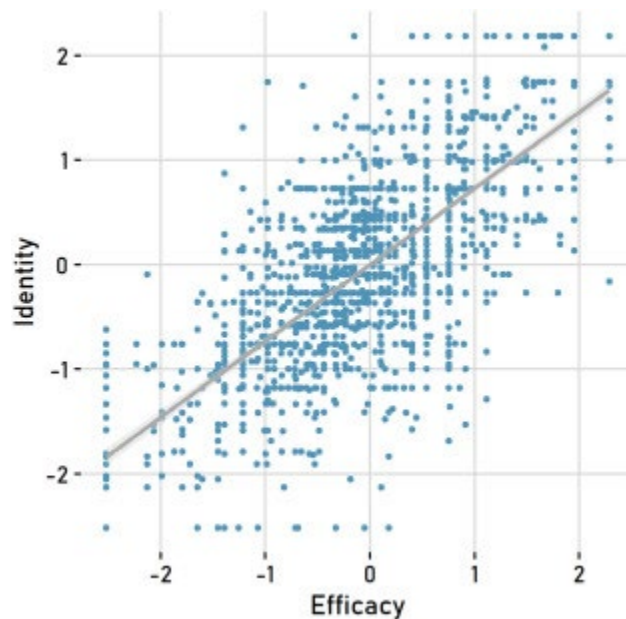


Fig. 3. Correlation between humor efficacy and identity scores.

A confirmatory factor analysis (CFA) of the HEISS supported our view of two factors that are strongly related but nevertheless distinct. Using Mplus 8.1, a CFA modeling humor efficacy and identity as distinct latent variables, each defined by four items, found good model fit: $\chi^2(19) = 206.825$, $p < .001$, RMSEA = 0.073 [90% CI: 0.064, 0.082], SRMR = 0.026, CFI = 0.975. Model fit for the two-factor model was significantly better than fit for a one-factor model that constrained the factor covariance to 1 (Wald (1 df) = 209.498, $p < .001$).¹

2.2.2. Age and gender

Our sample had a wide age range and almost equal numbers of women and men. We had no expectations for age, which had small negative correlations with efficacy ($r = -0.05$ [-0.10,

¹ In all four studies, the proposed two-factor HEISS model fit significantly better than a one-factor model. The OSM reports model fit statistics and Wald comparisons for the one-factor and two-factor models for all the samples.

-0.01]) and identity ($r = -0.13 [-0.17, -0.09]$). For gender, however, significant differences were found. Women had lower humor efficacy and identity than men, as Fig. 4 illustrates. Using the common guidelines of Cohen's $d = 0.20/0.50/0.80$ as small/medium/large effects, gender's effect sizes were in the small-to-medium range for efficacy ($d = 0.35 [0.26, 0.45]$) and for identity ($d = 0.27 [0.18, 0.36]$). Because the scale development process found essentially zero gender DIF, we can be relatively confident that these differences reflect real gender differences in the underlying constructs.

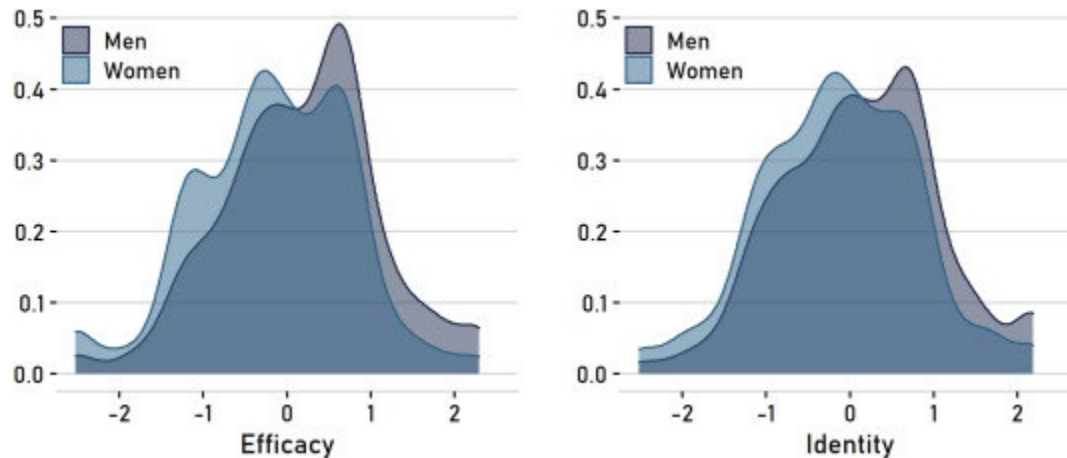


Fig. 4. Gender differences in humor efficacy and identity scores.

2.2.3. Discussion

The scale development process yielded compact scales with strong unidimensionality, good evidence for score reliability, good item discrimination, and an appropriate range of item difficulty for a scale of self-beliefs. Notably, the items showed essentially no gender DIF, making the HEISS a useful tool for studying gender and humor.

3. Study 2

Study 2 sought initial evidence for the validity of the HEISS scores. First, we examined humor backgrounds—past experiences with taking classes, working, creating, or performing in the humor domain—as predictors of humor efficacy and identity. The expectation that people who have a history of training, work, and voluntary engagement in humor behaviors will have higher humor efficacy and identity scores follows naturally from the model of the creative self-concept (Karwowski et al., 2019) and the broader social-cognitive model of efficacy beliefs (Bandura, 1997), which emphasizes the role of prior behavioral experience in the development of self-efficacy. Second, we examined links with the Big Five personality traits, which have rich links to other areas of humor, such as how people create funny ideas, use humor in interactions, and appreciate different kinds of jokes (Nusbaum et al., 2017; Plessen et al., 2020). Drawing on previous research, we expected humor self-efficacy and identity to be positively linked with extraversion and openness to experience (Greengross & Miller, 2009; Karwowski & Lebeda, 2016) and inversely related to conscientiousness and neuroticism (Greengross & Miller, 2009). Third and finally, we expected to replicate Study 1's finding that men had higher humor efficacy and identity scores.

3.1. Method

3.1.1. Participants

A final sample of 304 adults—153 women, 151 men—who ranged in age from 18 to 70 years old ($M = 33.78$ years, $SD = 13.03$, $Mdn = 30$) took part. Participants were recruited from the Prolific.co online survey panel and were eligible if they were at least 18 years old, spoke English as a native language, and identified as female or male. The final sample was refined from a larger sample of 336 people, for an exclusion rate of around 9.5% (see OSM).

3.1.2. Measures

After completing demographic items and the final 8-item HEISS, participants completed a set of six items used to measure their background in humor. On a *Yes/No* scale, people indicated if they had ever *performed comedy in public; created jokes for others to use; published jokes, cartoons, or other works of humor; had a job involving being funny or making people laugh; taken comedy or improv classes; and created a meme, joke, or cartoon to share on social media*. Finally, the Big Five traits were measured with 30-item BFI-2 (Soto & John, 2017). For the BFI-2, Cronbach's α for N, E, O, A, and C was 0.85, 0.72, 0.72, 0.74, and 0.79, respectively.

3.2. Results and discussion

Internal consistency was high for the efficacy ($\alpha = 0.89$, $\omega_h = 0.86$) and identity ($\alpha = 0.87$, $\omega_h = 0.83$) scales. They were modeled as correlated latent variables in Mplus 8.1, using maximum likelihood with robust standard errors and with factor variances set to 1. Model fit was good on most metrics, $\chi^2(19) = 66.10$, $p < .001$, RMSEA = 0.090 [90% CI: 0.067, 0.115], SRMR = 0.038, CFI = 0.964. The factors correlated highly, as in Study 1 ($r = 0.85$ [0.80, 0.90]). In all analyses, the humor efficacy and identity factors were used as correlated outcomes in a multivariate model. Model fit for the BFI-2 Big Five traits was poor, so these traits were modeled as average item scores instead of latent variables. The OSM presents the descriptive statistics and correlation matrix for all variables.

For the humor background items, the effect sizes offered evidence for the validity of the HEISS (see Table 3). For efficacy, the effect sizes (in d) ranged from 0.47 to 0.86; for identity, they ranged from 0.33 to 0.72.

The findings for the Big 5 factors, shown in Table 4, are consistent with past research on personality and humor production. For these models, all 5 factors were included as predictors, and humor efficacy and identity were simultaneous outcomes. People with higher humor self-efficacy were higher in extraversion and openness to experience, with smaller negative effects for neuroticism and conscientiousness. People with higher humor identity were higher in extraversion, with smaller effects for higher openness to experience and lower conscientiousness.

Table 3. Effect sizes for humor background items: studies 2, 3, and 4.

	Humor efficacy			Humor identity			Percent endorsed		
	Study 2	Study 3	Study 4	Study 2	Study 3	Study 4	Study 2	Study 3	Study 4
Performed comedy	0.49 [0.00, 0.97]	0.33 [-0.18, 0.83]	0.63 [0.15, 1.11]	0.33 [-0.18, 0.84]	0.19 [-0.35, 0.73]	1.12 [0.68, 1.55]	5.0%	5.7%	4.0%
Job	0.72 [0.35, 1.09]	0.54 [0.34, 0.74]	0.44 [0.05, 0.83]	0.61 [0.27, 0.96]	0.53 [0.32, 0.73]	0.88 [0.58, 1.17]	5.9%	36.6%	5.3%
Classes	0.86 [0.59, 1.12]	0.42 [0.21, 0.62]	0.56 [0.20, 0.92]	0.42 [0.03, 0.81]	0.42 [0.21, 0.62]	0.75 [0.27, 1.22]	6.6%	44.9%	3.3%
Published	0.77 [0.48, 1.06]	0.74 [0.45, 1.03]	0.42 [0.05, 0.79]	0.63 [0.36, 0.90]	0.46 [0.11, 0.80]	0.57 [0.15, 0.98]	11.5%	5.7%	6.0%
Created	0.67 [0.41, 0.94]	0.36 [-0.04, 0.76]	0.68 [0.47, 0.90]	0.72 [0.48, 0.96]	0.34 [-0.17, 0.85]	0.83 [0.59, 1.07]	20.1%	5.2%	12.8%
Social media	0.47 [0.24, 0.71]	0.47 [0.25, 0.69]	0.39 [0.20, 0.58]	0.46 [0.22, 0.70]	0.49 [0.28, 0.70]	0.47 [0.26, 0.68]	39.9%	34.0%	33.8%

Note. Study 2, $n = 304$ (English), Study 3, $n = 385$ (Polish), Study 4, $n = 400$ (English). The efficacy and identify columns display Cohen's d effect sizes and 95% confidence intervals. These were estimated as Y -standardized regression coefficients in Mplus and will vary slightly from values computed from descriptive statistics due to the MLR estimator.

Table 4. Effect sizes for Big Five traits: Study 2 and Study 3.

Empty Cell	Humor efficacy		Humor identity	
	Study 2	Study 3	Study 2	Study 3
Neuroticism	-0.15 [-0.29, -0.02]	-0.12 [-0.22, -0.02]	0.02 [-0.12, 0.16]	0.02 [-0.10, 0.13]
Extraversion	0.29 [0.16, 0.42]	0.50 [0.39, 0.61]	0.27 [0.14, 0.41]	0.42 [0.31, 0.54]
Openness to experience	0.26 [0.16, 0.37]	0.19 [0.09, 0.30]	0.15 [0.03, 0.27]	0.13 [0.02, 0.24]
Agreeableness	0.04 [-0.09, 0.16]	0.02 [-0.10, 0.15]	0.05 [-0.09, 0.18]	0.15 [0.02, 0.27]
Conscientiousness	-0.16 [-0.28, -0.05]	-0.08 [-0.18, 0.01]	-0.18 [-0.31, -0.05]	-0.16 [-0.26, -0.06]

Note. Study 2, $n = 304$ (English), Study 3, $n = 385$ (Polish). The efficacy and identity columns display standardized regression weights and confidence intervals from a model including both HEISS scores as outcomes and all five traits as predictors.

Finally, as in Study 1, women had lower efficacy ($d = -0.35 [-0.58, -0.12]$) and identity ($d = -0.23 [-0.46, 0.00]$) scores than men, with mostly small effect sizes, and age had small, negative regression effects on efficacy ($\beta = -0.11 [-0.23, 0.01]$) and identity ($\beta = -0.08 [-0.19, 0.03]$).

In sum, the findings provide evidence for the validity of the HEISS scores. Consistent with the social-cognitive model underlying the scales (Karwowski et al., 2019), people with more training, experience, and behavioral engagement in humor reported higher self-efficacy and identity. Furthermore, the Big Five findings are consistent with a great deal of research, such as (1) the prominent roles of openness to experience and extraversion in humor production and use (Nusbaum et al., 2017; Plessen et al., 2020); (2) the salience of high openness and low conscientiousness in the personality profiles of comedians (Greengross & Miller, 2009); and (3) the broader role that openness to experience and extraversion, as facets of broader behavioral flexibility, play in creative domains (Karwowski & Lebuda, 2016).

4. Study 3

Study 3 sought to replicate and extend the findings of Study 2 in a sample from a different language country—Poland—and by including additional validity measures. As in Study 2, we examined humor backgrounds and Big Five personality traits. Additionally, we added two broad constructs that seemed promising as HEISS correlates. The first was creative self-concept: creative self-efficacy and creative personal identity (SSCS; Karwowski et al., 2018); the second was “dark triad” personality factors: Machiavellianism, narcissism, and psychopathy (Paulhus & Williams, 2002). We expected that humor self-efficacy will be predicted by creative self-efficacy rather than creative personal identity, and that humor identity will be more strongly predicted by creative personal identity. One reason for hypothesizing this is that self-efficacy scales (creative and humor related) share the “I can” wording and focus, while identity items share the “I am” emphasis. More substantially, however, humor is an important domain of creativity in everyday life and—as Study 2 demonstrated—people with higher experience in producing humor hold higher humor efficacy (Study 2). That leads us to expect that creative self-efficacy and humor efficacy will be positively related. Likewise, people can certainly define themselves as creative people but not as funny people, but we would expect that being high in humor identity implies a broader view of oneself as creative.

The dark triad was included in an exploratory manner. Although there is evidence that people scoring higher on psychopathy and Machiavellianism tend to use more negative humor (Veselka et al., 2010), to the best of our knowledge, there is no research on dark triad traits and humor self-concepts. What's more, there are reasons to expect that the associations will be virtually null—people higher and lower in dark triad traits can assess their humor efficacy and identity similarly, even if their typical kinds of humor differ (e.g., affiliative vs aggressive humor; Martin et al., 2012).

4.1. Method

4.1.1. *Participants*

A final sample of 385 adults—205 women, 180 men—who ranged in age from 18 to 84 years old ($M = 36.74$ years, $SD = 14.28$, $Mdn = 34$) took part. Participants were recruited from the Syno International online survey panel and were eligible if they were at least 18 years old, spoke Polish as a native language, and identified as female or male. The final sample was refined from a larger sample of 420 people, for an exclusion rate of around 8% (see OSM).

4.1.2. *Measures*

The HEISS was translated from English to Polish and then back-translated into English independently by the third author and an independent researcher fluent in Polish and English. Discrepancies were discussed, and the final Polish version of HEISS was established. After completing demographic items and the final 8-item HEISS, participants completed the same set of six items used to measure their humor background used in Study 2. The Big Five traits were measured with Goldberg's (1999; Goldberg et al., 2006) 50-item BFI-50, which includes ten items per factor (for Polish adaptation, see Strus et al., 2014). Cronbach's α values for N, E, O, A, and C were 0.91, 0.91, 0.74, 0.83, and 0.81, respectively. Additionally, we measured creative self-concept using the Short Scale for Creative Self (SSCS; Karwowski et al., 2018), which

yields scores for creative self-efficacy ($\alpha = 0.90$) and creative identity ($\alpha = 0.90$). Finally, the Dark Triad traits—psychopathy ($\alpha = 0.77$), Machiavellianism ($\alpha = 0.79$), and narcissism ($\alpha = 0.65$)—were assessed using the Dirty Dozen Scale (DDS; Jonason & Webster, 2010; for Polish adaptation, see Czarna et al., 2016). The last three questionnaires were presented in a counter-balanced order.

4.2. Results and discussion

Internal consistency was high for the efficacy ($\alpha = 0.89$, $\omega_h = 0.89$) and identity ($\alpha = 0.88$, $\omega_h = 0.88$) scales. As in Study 2, both scales were modeled as correlated latent variables in Mplus 8.1, using maximum likelihood with robust standard errors and with factor variances set to 1. Model fit was good, $\chi^2(19) = 38.91$, $p = .005$, RMSEA = 0.052 [90% CI: 0.028, 0.076], SRMR = 0.022, CFI = 0.987. The factors correlated highly ($r = 0.83$ [0.78, 0.88]), as in Study 1 and Study 2. Big Five traits, creative self-concept, and dark triad traits were modeled as observed scores formed by their item averages and included as predictors of two latent HEISS scales. The OSM presents the descriptive statistics and correlation matrix for all variables, as well as analyses of measurement invariance between the Polish and English language groups.

For the humor background items, the effect sizes offered evidence for the validity of the HEISS (see Table 3). For efficacy, the effect sizes (in d) ranged from 0.33 to 0.74; for identity, they ranged from 0.19 to 0.63.

The findings for the Big 5 factors, shown in Table 4, fit the patterns observed in Study 2. People with higher humor self-efficacy also scored higher in extraversion and openness to experience, with smaller negative effects for neuroticism and conscientiousness. People with higher identity were higher in extraversion, with smaller effects for higher agreeableness and lower conscientiousness.

As illustrated in Table 5, people higher in humor self-efficacy were also higher in creative self-efficacy and creative personal identity. Humor identity, however, was linked only to creative personal identity and unrelated to creative self-efficacy, indicating greater specificity. For the dark triad traits, no significant links were observed in the case of humor efficacy. Humor identity was positively and significantly related to narcissism, yet the effect size of this relationship was small.

Table 5. Effect sizes for creative self-concept and dark triad: Study 3.

	Humor efficacy	Humor identity
Creative self-concept		
Creative self-efficacy	0.31 [0.10, 0.52]	0.05 [-0.16, 0.26]
Creative personal identity	0.22 [0.03, 0.41]	0.38 [0.17, 0.58]
Dark triad		
Machiavellianism	-0.03 [-0.19, 0.14]	0.03 [-0.13, 0.18]
Psychoticism	-0.10 [-0.26, 0.05]	-0.18 [-0.33, 0.03]
Narcissism	0.13 [-0.02, 0.28]	0.18 [0.04, 0.31]

Note. Study 3, $n = 385$. The efficacy and identify columns display standardized regression weights and confidence intervals from two separate models, one including both outcomes and the two Creative Self-Concept scales, and a second including both outcomes and the three Dark Triad traits.

Finally, contrary to what we observed in the English-speaking samples in Study 1 and Study 2, in Poland there were no gender differences in either efficacy ($d = 0.08$ [$-0.03, 0.18$]) or identity ($d = 0.01$ [$-0.09, 0.12$]). As in Studies 1 and 2, however, age had small, negative regression effects on efficacy ($\beta = -0.12$ [$-0.23, -0.01$]) and identity ($\beta = -0.18$ [$-0.28, -0.07$]).

In sum, Study 3 replicated and expanded the evidence for validity in a Polish-speaking sample. The findings for the Big Five personality traits and humor background broadly replicated, which bolsters evidence for score validity. Relationships with the dark triad traits were generally small, which illustrates some boundaries on the conceptual network of the HEISS. Finally, it was noteworthy that the Polish sample showed no gender differences in either humor efficacy or identity, a finding that is explored later.

5. Study 4

In Study 4, we sought to extend our understanding of humor efficacy and identity by exploring their links to prominent constructs in humor research. Humor styles were assessed with the Humor Styles Questionnaire (HSQ; Martin et al., 2003), perhaps the most widely used self-report scale in modern humor research (Kuiper, 2020; Schermer et al., 2019; Silvia & Rodriguez, 2020). The humor styles model proposes four forms (Kuiper, 2016): two adaptive styles (affiliative and self-enhancing) and two maladaptive styles (aggressive and self-defeating). This dimension is crossed with self-oriented humor (self-enhancing and self-defeating) and other-oriented humor (affiliative and aggressive).

Affiliative humor appears to be the style that would be most closely linked to humor efficacy and identity. According to Martin et al. (2003), people high in the affiliative humor style “tend to say funny things, to tell jokes, and to engage in spontaneous witty banter to amuse others, to facilitate relationships, and to reduce interpersonal tensions” (p. 53). The other styles involve using humor for coping and emotion regulation (self-enhancing), using humor to tease, belittle, and disparage others (aggressive), and using self-disparaging humor to ingratiate oneself, preempt criticism, and avoid social exclusion (self-defeating).

In addition to the HSQ, we measured playfulness, an increasingly prominent construct in humor research (Proyer, 2018). Past theory and research suggest close ties between a playful, non-serious approach to events, experiences, and ideas and the appreciation and production of humor (Chafe, 2007; Proyer, 2012b, Proyer, 2013, Proyer, 2018). Need for uniqueness, the preference for being distinctive and standing out from other people (Lynn & Snyder, 2002), was measured as well. The need to be different is an important motivation for creative behavior (e.g., Dollinger, 2003; Joy, 2004) and thus should be relevant to interpersonal humor. Finally, for further replication, we included the six items about participants' humor background used in Studies 2 and 3 along with gender and age.

5.1. Method

5.1.1. Participants

A final sample of 400 adults—204 women, 196 men—who ranged in age from 18 to 80 years old ($M = 35.08$ years, $SD = 12.79$, $Mdn = 32$) took part. They were recruited from the Prolific.co online survey panel and were eligible if they were at least 18 years old, spoke English as a native language, and identified as female or male. The final sample was refined from a larger sample of 415 people, for an exclusion rate of around 3.6% (see OSM).

5.1.2. Measures

After completing demographic items, people completed the HEISS along with the same 6 items measuring humor backgrounds in Study 2 and 3. Humor styles were measured with the HSQ (Martin et al., 2003), a 32-item scale that measures the four styles—affiliative ($\alpha = 0.86$), self-enhancing ($\alpha = 0.81$), aggressive ($\alpha = 0.72$), and self-defeating ($\alpha = 0.83$)—using a 5-point response scale (see Silvia & Rodriguez, 2020). Playfulness was measured with the Short Measure of Adult Playfulness (SMAP; Proyer, 2012a), which has 5 items measured on a 4-point response scale. The SMAP yields a single score for overall playfulness ($\alpha = 0.86$). Need for uniqueness was measured with the Self-attributed Need for Uniqueness (SANU; Lynn & Harris, 1997) scale, which has 4 items measured on a 5-point scale ($\alpha = 0.87$). It yields a single overall score reflecting people's need to stand out from others (Lynn & Snyder, 2002).

5.2. Results and discussion

Internal consistency was high for the efficacy ($\alpha = 0.89$, $\omega_h = 0.88$) and identity ($\alpha = 0.87$, $\omega_h = 0.83$) scales. They were modeled as correlated latent variables in Mplus 8.1, using maximum likelihood with robust standard errors and with factor variances set to 1. Model fit was good on most metrics, $\chi^2(19) = 113.51$, $p < .001$, RMSEA = 0.112 [90% CI: 0.092, 0.132], SRMR = 0.053, CFI = 0.942. The factors correlated highly, as in the prior samples ($r = 0.76$, [0.68, 0.83]). The four humor styles, playfulness, and need for uniqueness were modeled as item averages. The OSM presents the descriptive statistics and correlation matrix for all variables.

Table 6. Effect sizes for humor styles, playfulness, and need for uniqueness: Study 4.

	Humor efficacy	Humor identity
HSQ: affiliative	0.70 [0.62, 0.77]	0.64 [0.56, 0.72]
HSQ: self-enhancing	0.17 [0.10, 0.25]	0.13 [0.04, 0.22]
HSQ: aggressive	0.09 [0.01, 0.17]	0.03 [-0.05, 0.11]
HSQ: self-defeating	-0.17 [-0.26, -0.09]	0.17 [0.08, 0.25]
Playfulness	0.48 [0.39, 0.57]	0.61 [0.53, 0.69]
Need for uniqueness	0.19 [0.07, 0.31]	0.33 [0.22, 0.44]

Note. Study 4, $n = 400$. The efficacy and identify columns display standardized regression weights and confidence intervals from three separate models: one including all 4 humor styles as predictors, one including only playfulness, and one including only need for uniqueness.

For humor styles, as expected, the strongest relationships were for the affiliative humor style, which assesses a mix of creating humor, valuing being funny, and using humor to amuse and entertain others. The effect sizes were large for affiliative humor (see Table 6), but the correlations with the other three humor styles were much smaller. Neither efficacy nor identity appreciably correlated with the aggressive humor style, and correlations with self-enhancing and self-defeating humor were small as well. This overall pattern—high correlations with affiliative

humor, and smaller correlations with the rest—supports our view of humor efficacy and identity as constructs focused on the interpersonal use of humor.

For playfulness, both humor efficacy and identity had positive correlations, large in effect size (see Table 6), with SMAP scores, consistent with the view that playfulness and humor are closely connected (Proyer, 2018). Likewise, both efficacy and identity correlated positively with need for uniqueness, with effect sizes in the small to medium range (see Table 6), suggesting that people high in humor efficacy and identity prefer to be distinctive and stand out from others.

For the humor background items, the effect sizes resembled the findings from Studies 2 and 3 (see Table 3). Consistent with the prior samples and with the social-cognitive model that underlies the HEISS, people higher in efficacy and identity were more likely to have performed humor and held jobs involving humor, among other activities. For efficacy, the effect sizes (in d) ranged from 0.39 to 0.68; for identity, they ranged from 0.47 to 1.12.

Finally, women had lower efficacy ($d = -0.57 [-0.75, -0.38]$) and identity ($d = -0.43 [-0.63, -0.23]$) scores than men, with mostly medium effect sizes. Age had no effect on efficacy ($\beta = 0.01 [-0.09, 0.11]$) and a small negative effect on identity ($\beta = -0.17 [-0.27, -0.07]$).

In sum, the findings expand and enrich our understanding of the conceptual network of humor efficacy and identity. People with higher self-efficacy for creating humor and who view being funny as central to their self-concepts were likely to be higher in playfulness and have a higher need for uniqueness. The strong links with the affiliative humor style for both humor efficacy and identity, paired with minor links to the aggressive humor style, suggest that the HEISS assesses relatively adaptive humor traits that reflect constructive uses of humor in interpersonal interactions.

6. General discussion

Everyday social interaction is suffused with humor, from off-hand jokes in conversation to memes shared online, so it's important to understand how people create and use humor. In the present research, we developed and evaluated the HEISS, a pair of brief scales that measure humor self-efficacy and humor identity. Research in the broader domain of creativity has shown that studying the creative self-concept can illuminate important motivational features of creativity, from the goals people set to their decisions to persist or quit (Karwowski & Beghetto, 2019; Puente-Diaz & Cavazos-Arroyo, 2018; Royston & Reiter-Palmon, 2019), so extending this framework to humor, a specific domain of creativity, seems likely to be fruitful.

In four samples, the HEISS showed strong psychometric properties in both English-speaking and Polish-speaking samples. The items have good discrimination, an appropriate range of difficulty, and essentially no DIF for gender. Evidence for score reliability, estimated via EAP, alpha, and omega reliability, was substantial in all four samples. Finally, the studies provided initial evidence for score validity by showing that humor efficacy and identity are predicted by prior humor experiences (e.g., performing, studying, and working with humor), by personality traits (openness to experience and extraversion), by playfulness and the affiliative humor style, and by broader creative efficacy and identity beliefs. The two HEISS scales are highly correlated

with each other— $r = 0.73$ [0.71, 0.75] for EAP trait scores in the largest sample—much like the high correlation between creative efficacy and identity in past work (Karwowski et al., 2018)—consistent with the interwoven nature of people's beliefs about their abilities, identities, and capacities proposed by social-cognitive models (Beghetto & Karwowski, 2017).

The psychology of humor has a long tradition of individual differences research (Martin, 2003; Martin & Ford, 2018), and it is helpful to position humor efficacy and identity within the network of other concepts. For the well-known distinction between the appreciation of humor (e.g., what people find humorous as the audience) and the production of humor (what people generate as the creator), we see the HEISS as firmly within the production domain because it assesses self-beliefs about one's ability to create effective humor. Within the humor production domain (Ruch & Heintz, 2019), one finds research on individual differences in the ability to generate funny ideas overall, the kinds of humor people tend to generate (e.g., sexual, sarcastic, dark, or bizarre humor), or the goals and functions that their humor production serves (e.g., reducing conflict, coping with stress, or disparaging others). We see humor efficacy and identity as distinct from these, in that it refers to people's subjective self-beliefs about what they are like and what they can effectively accomplish in the domain of creating humor. These beliefs, as measured by the HEISS, do not imply any particular kind of humor or purpose for being funny. Finally, humor efficacy and identity are self-focused beliefs, in that they refer to people's beliefs about their own traits, abilities, and features, rather than global beliefs about the nature of humor.

The present research represents only the first steps in gaining evidence for the validity of the concepts of humor efficacy and identity and the scales we have developed to measure them. In future work, it would be important to examine the temporal stability of HEISS scores (e.g., test-retest stability) as well as their sensitivity to change as a result of significant experiences (e.g., pre-post change following humor-based interventions). Evidence for incremental validity, both in relation to humor-related constructs and broader efficacy and identity constructs, would further clarify the nature and scope of humor efficacy and identity. Likewise, we should note that aside from the first sample, the present samples were large enough for high statistical power for the kinds of effect sizes common in individual differences research but are more modest with regards to the stability of the observed correlations (Schönbrodt & Perugini, 2013) and latent variable correlations (Kretzschmar & Gignac, 2019), so the estimated effect sizes would benefit from additional large-sample replications.

Regarding implications for future research, we see this scale as a useful tool for researchers interested in studying motivational aspects of humor production and use. For example, people vary in how well they can generate funny ideas on the spot, how often they use humor in everyday interactions, and how effectively humor is used to meet social goals (Heintz, 2017; Nezlek & Derks, 2020). Beliefs captured by the HEISS are likely prominent in the production and self-regulation of humor, the humor-related goals people set for themselves (e.g., whether to pursue activities that involve trying to be funny), and how they evaluate their own attempts to be funny (Silvia et al., 2021b).

In addition, because these scales were crafted to avoid gender DIF, they are well suited to studying the complex intersections of gender and humor (Martin, 2014). In the United States, for example, there are widely held stereotypes about gender and humor, such as the belief that

“women aren't funny” (e.g., Hitchens, 2007; Hooper et al., 2016; Mickes et al., 2012). A recent meta-analysis did show a small advantage for men (Greengross et al., 2020), but the proposed explanations, from evolved differences grounded in sexual selection to cultural norms and disparate social power, have rarely been directly tested. The HEISS could be a useful tool for understanding gender's role in humor use and production. In the present samples, for example, the persistent gender difference in the English-speaking samples was not found in the Polish-speaking sample. Although more work is obviously needed, cross-cultural findings like these imply that cultural differences in stereotypes and norms—captured at the person level by people's humor self-concepts—are likely at work. Whenever differences in interest, motivation, and performance are found, it's worth examining people's beliefs about what's personally important and what they can do well.

Credits

Paul J. Silvia: Conceptualization, Formal analysis, Investigation, Methodology, Writing - original draft. **Rebekah M. Rodriguez:** Conceptualization, Investigation, Methodology, Writing - review & editing. **Maciej Karwowski:** Conceptualization, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing.

References

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.
- Beghetto, R. A., & Karwowski, M. (2017). Toward untangling creative self-beliefs. In M. Karwowski, & J. C. Kaufman (Eds.), *The creative self: Effects of beliefs, self-efficacy, mindset, and identity* (pp. 3-22). Academic Press.
- Chafe, W. L. (2007). *The importance of not being earnest: The feeling behind laughter and humor*. John Benjamins Publishing.
- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48, 1-29.
- Chalmers, R. P. (2020). mirt: Multidimensional item response theory. R package version 1.33.2 <https://cran.r-project.org/package=mirt>.
- Chen, W. H., & Thissen, D. (1997). Local dependence indices for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, 22, 265-289.
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software*, 39(8), 1-30.
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2016). lordif: Logistic ordinal regression differential item functioning using IRT. R package version 0.3.3 <https://cran.r-project.org/package=lordif>.

- Christensen, A. P., Silvia, P. J., Nusbaum, E. C., & Beaty, R. E. (2018). Clever people: Intelligence and humor production ability. *Psychology of Aesthetics, Creativity, and the Arts, 12*(2), 136-143.
- Crane, P. K., Gibbons, L. E., Ocepek-Welikson, K., Cook, K., & Cella, D. (2007). A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research, 16*, 69-84.
- Czarna, A. Z., Jonason, P. K., Dufner, M., & Kossowska, M. (2016). The Dirty Dozen Scale: Validation of a Polish version and extension of the nomological net. *Frontiers in Psychology, 7*, 445.
- Diedrich, J., Benedek, M., Jauk, E., & Neubauer, A. C. (2015). Are creative ideas novel and useful? *Psychology of Aesthetics, Creativity, and the Arts, 9*, 35-40.
- Dollinger, S. J. (2003). Need for uniqueness, need for cognition, and creativity. *Journal of Creative Behavior, 37*, 99-116.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. J. Deary, F. D. Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (pp. 7-28). Tilburg University Press.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & 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.
- Greengross, G. (2020). Sex and gender differences in humor: Introduction and overview. *Humor., 33*, 175-178.
- Greengross, G., & Miller, G. F. (2009). The Big Five personality traits of professional comedians compared to amateur comedians, comedy writers, and college students. *Personality and Individual Differences, 47*, 79-83.
- Greengross, G., Silvia, P. J., & Nusbaum, E. C. (2020). Sex differences in humor production ability: A meta-analysis. *Journal of Research in Personality, 84*, Article 103886.
- Heintz, S. (2017). Putting a spotlight on daily humor behaviors: Dimensionality and relationships with personality, subjective well-being, and humor styles. *Personality and Individual Differences, 104*, 407-412.
- Hitchens, C. (2007). Why women aren't funny. *Vanity Fair, 557*, 54-59.
- Hofmann, J., Platt, T., Lau, C., & Torres-Marín, J. (2020). *Gender differences in humor-related traits, humor appreciation, production, comprehension, (neural) responses, use, and correlates: A systematic review*. *Current Psychology*.
- Hooper, J., Sharpe, D., & Roberts, S. G. B. (2016). Are men funnier than women, or do we just think they are? *Translational Issues in Psychological Science, 2*, 54-62.

- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*, 329-349.
- Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment, 22*, 420-432.
- Joy, S. (2004). Innovation motivation: The need to be different. *Creativity Research Journal, 16*(2-3), 313-330.
- Karwowski, M., & Beghetto, R. A. (2019). Creative behavior as agentic action. *Psychology of Aesthetics, Creativity, and the Arts, 13*, 402-415.
- Karwowski, M., & Kaufman, J. C. (2017). *The creative self: Effects of beliefs, self-efficacy, mindset, and identity*. Academic Press.
- Karwowski, M., & Lebuda, I. (2016). The Big Five, the Huge Two, and creative self-beliefs: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts, 10*, 214-232.
- Karwowski, M., Lebuda, I., & Beghetto, R. A. (2019). Creative self-beliefs. In J. C. Kaufman, & R. J. Sternberg (Eds.), *Cambridge handbook of creativity* (pp. 396-417). Cambridge University Press.
- Karwowski, M., Lebuda, I., & Wisniewska, E. (2018). Measuring creative self-efficacy and creative personal identity. *International Journal of Creativity and Problem Solving, 28*, 45-57.
- Kaufman, J. C. (2012). Counting the muses: Development of the Kaufman Domains of Creativity Scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts, 6*, 298-308.
- Kellner, R., & Benedek, M. (2017). The role of creative potential and intelligence for humor production. *Psychology of Aesthetics, Creativity, and the Arts, 11*, 52-58.
- Kretschmar, A., & Gignac, G. E. (2019). At what sample size do latent variable correlations stabilize? *Journal of Research in Personality, 80*, 17-22.
- Kuiper, N. A. (2020). Humor Styles Questionnaire. In V. Zeigler-Hill, & T. Shackelford (Eds.), *Encyclopedia of personality and individual differences*. Springer.
- Kuiper, N. A. (2016). Model of humor styles. In V. Zeigler-Hill, & T. Shackelford (Eds.), *Encyclopedia of personality and individual differences*. Springer.
- Lu, J. G., Martin, A. E., Usova, A., & Galinsky, A. D. (2019). Creativity and humor across cultures: Where aha meets haha. In S. R. Luria, J. Baer, & J. C. Kaufman (Eds.), *Creativity and humor* (pp. 183-203). Academic Press.
- Lynch, O. H. (2002). Humorous communication: Finding a place for humor in communication research. *Communication Theory, 12*(4), 423-445.
- Lynn, M., & Harris, J. (1997). Individual differences in the pursuit of self-uniqueness through consumption. *Journal of Applied Social Psychology, 27*(21), 1861-1883.

- Lynn, M., & Snyder, C. R. (2002). Uniqueness seeking. In C. R. Snyder, & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 395-410). Oxford University Press.
- Martin, R. (2003). Sense of humor. In S. J. Lopez, & C. R. Snyder (Eds.), *Positive psychological assessment: A handbook of models and measures* (pp. 313-326). American Psychological Association.
- Martin, R. A. (2014). Humor and gender: An overview of psychological research. In D. Chiaro, & R. Baccolini (Eds.), *Gender and humor: Interdisciplinary and international perspectives* (pp. 123-144). Mouton de Gruyter.
- Martin, R. A., & Ford, T. (2018). *The psychology of humor: An integrative approach* (2nd ed.). Academic Press.
- Martin, R. A., Lastuk, J. M., Jeffery, J., Vernon, P. A., & Veselka, L. (2012). Relationships between the Dark Triad and humor styles: A replication and extension. *Personality and Individual Differences, 52*, 178-182.
- Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality, 37*, 48-75.
- Meade, A. W. (2010). A taxonomy of effect size measures for the differential functioning of items and scales. *Journal of Applied Psychology, 95*, 728-743.
- Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician, 54*, 17-24.
- Mickes, L., Walker, D. E., Parris, J. L., Mankoff, R., & Christenfeld, N. J. (2012). Who's funny: Gender stereotypes, humor production, and memory bias. *Psychonomic Bulletin & Review, 19*, 108-112.
- Nezlek, J. B., & Derks, P. L. (2020). Relationships between personality and the everyday use of humor. *Humor, 33*, 361-379.
- Nusbaum, E. C., & Silvia, P. J. (2017). What are funny people like? Exploring the crossroads of humor ability and openness to experience. In G. J. Feist, R. Reiter-Palmon, & J. C. Kaufman (Eds.), *Cambridge handbook of creativity and personality research* (pp. 294-322). Cambridge University Press.
- Nusbaum, E. C., Silvia, P. J., & Beaty, R. E. (2017). Ha ha? Assessing individual differences in humor production ability. *Psychology of Aesthetics, Arts, and Creativity, 11*, 231-241.
- Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556-563.
- Perchtold-Stefan, C. M., Fink, A., Rominger, C., & Papousek, I. (2020). Motivational factors in the typical display of humor and creative potential: The case of malevolent creativity. *Frontiers in Psychology, 11*, 1213.
- Plessen, C. Y., Franken, F. R., Ster, C., Schmid, R. R., Wolfmayr, C., Mayer, A. M., ... Maierwieser, R. J. (2020). Humor styles and personality: A systematic review and meta-

- analysis on the relations between humor styles and the Big Five personality traits. *Personality and Individual Differences*, 154, Article 109676.
- Proyer, R. T. (2012a). Development and initial assessment of a short measure for adult playfulness: The SMAP. *Personality and Individual Differences*, 53(8), 989-994.
- Proyer, R. T. (2012b). Examining playfulness in adults: Testing its correlates with personality, positive psychological functioning, goal aspirations, and multi-methodically assessed ingenuity. *Psychological Test and Assessment Modeling*, 54(2), 103-127.
- Proyer, R. T. (2013). The well-being of playful adults: Adult playfulness, subjective wellbeing, physical well-being, and the pursuit of enjoyable activities. *European Journal of Humour Research*, 1(1), 84-98.
- Proyer, R. T. (2018). Playfulness and humor in psychology: An overview and update. *Humor*, 31(2), 259-271.
- Puente-Diaz, R., & Cavazos-Arroyo, J. (2018). An exploration of some antecedents and consequences of creative self-efficacy among college students. *Journal of Creative Behavior*, 52, 256-266.
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from <https://www.R-project.org>.
- Revelle, W. (2021). psych: Procedures for psychological, psychometric, and personality research. R package version 2.1.3 <https://CRAN.R-project.org/package=psych>.
- Royston, R., & Reiter-Palmon, R. (2019). Creative self-efficacy as mediator between creative mindsets and creative problem-solving. *Journal of Creative Behavior*, 53, 472-481.
- Ruch, W., & Heintz, S. (2019). Humor production and creativity: Overview and recommendations. In S. R. Luria, J. Baer, & J. C. Kaufman (Eds.), *Creativity and humor* (pp. 1-42). Academic Press.
- Sawyer, R. K. (2006). *Explaining creativity: The science of human innovation*. New York, NY: Oxford University Press.
- Schermer, J. A., Rogoza, R., Kwiatkowska, M. M., Kowalski, C. M., Aquino, S., Ardi, R., ... Krammer, G. (2019). Humor styles across 28 countries. *Current Psychology*. Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609-612.
- Silvia, P. J., Christensen, A. P., & Cotter, K. N. (2021a). Right-wing authoritarians aren't very funny: RWA, personality, and creative humor production. *Personality and Individual Differences*, 170, Article 110421.
- Silvia, P. J., Greengross, G., Cotter, K. N., Christensen, A. P., & Gredlein, J. M. (2021b). If you're funny and you know it: Personality, gender, and people's ratings of their attempts at humor. *Journal of Research in Personality*, 92, Article 104089.
- Silvia, P. J., & Rodriguez, R. M. (2020). Time to renovate the Humor Styles Questionnaire? An item response theory analysis of the HSQ. *Behavioral Sciences*, 10 (11), 173.

- Snyder, H. T., Sowden, P. T., Silvia, P. J., & Kaufman, J. C. (2020). The creative self: Do people distinguish creative self-perceptions, efficacy, and personal identity? *Psychology of Aesthetics, Creativity, and the Arts*.
- Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory-2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, 68, 69-81.
- Strus, W., Cieciuch, J., & Rowinski, T. (2014). The Polish adaptation of the IPIP-BFM-50 questionnaire for measuring five personality traits in the lexical approach. *Annals of Psychology*, 17, 347-366.
- Tierney, P., & Farmer, S. M. (2002). Creative self-efficacy: Its potential antecedents and relationship to creative performance. *Academy of Management Journal*, 45, 1137-1148.
- Veselka, L., Schermer, J. A., Martin, R. A., & Vernon, P. A. (2010). Relations between humor styles and the Dark Triad traits of personality. *Personality and Individual Differences*, 48, 772-774.
- Walter, N., Cody, M. J., Xu, L. Z., & Murphy, S. T. (2018). A priest, a rabbi, and a minister walk into a bar: A meta-analysis of humor effects on persuasion. *Human Communication Research*, 44, 343-373.
- Weinstein, N., Hodgins, H. S., & Ostvik-White, E. (2011). Humor as aggression: Effects of motivation on hostility expressed in humor appreciation. *Journal of Personality and Social Psychology*, 100(6), 1043-1055.

Appendix: Online Supplementary Material

1. Polish Translation

[eff1] I think I can make almost anyone laugh.

[PL: Myślę, że jestem w stanie rozśmieszyć niemal każdego]

[eff2] I trust my ability to be funny.

[PL: Wierzę w swoje zdolności do bycia zabawnym]

[eff3] I feel confident in my humor skills.

[PL: Mam zaufanie do swoich umiejętności rozśmieszania innych]

[eff4] Being funny is something that comes naturally to me.

[PL: Bycie zabawnym jest czymś, co przychodzi mi naturalnie]

[id1] Being a funny person is a big part of who I am.

[PL: Bycie zabawnym jest ważną częścią tego, kim jestem]

[id2] It's important to me to be a funny person.

[PL: To dla mnie ważne, aby być zabawną osobą]

[id3] My humor ability is central to who I am.

[PL: Moje poczucie humoru jest kluczowe dla tego, kim jestem]

[id4] Making people laugh is important to me.

[PL: Rozśmieszanie ludzi jest dla mnie ważne]

2. Sampling Details

Study 1

The two samples were collected via the Prolific.co online survey panel. In both samples, we requested an equal number of self-identified women and men who were at least 18 years old, spoke English as their first language, and would complete the survey on a computer or tablet (not a smartphone). Sample 1 requested 510 women and 510 men; Sample 2 requested 500 women and 500 men. We used the *careless* package (Yentes & Wilhelm, 2021) to estimate Mahalanobis's distance (D) and long-string indexes for the full set of humor items. In both samples, people were excluded for several reasons:

- Not completing the survey (partial data)
- Extensive missing data (more than a few items)
- Failing a directed response item (“Please select *strongly disagree*”), by far the largest contributor to exclusion
- Marking a gender other than male or female (only a few cases per sample) due to gender-based DIF testing
- Elevated Mahalanobis's D values

The second sample included a reverse-coded trap item (“I'm not a funny person”). Because all the scale items are positively coded, the reverse-coded trap item enables more decisive exclusions due to long-string indexes because people who respond to all the items with the same value are almost surely responding carelessly. A handful of people were excluded from Sample 2 for this reason.

Study 2

The sampling and screening approach in Study 2 was identical to the second sample in Study 1, except we requested 310 people divided evenly between men and women. Note that the number requested (310) is less than the number of respondents (336) because rejecting a response in Prolific.com due to failing an attention check will reopen a slot for a new participant.

Study 3

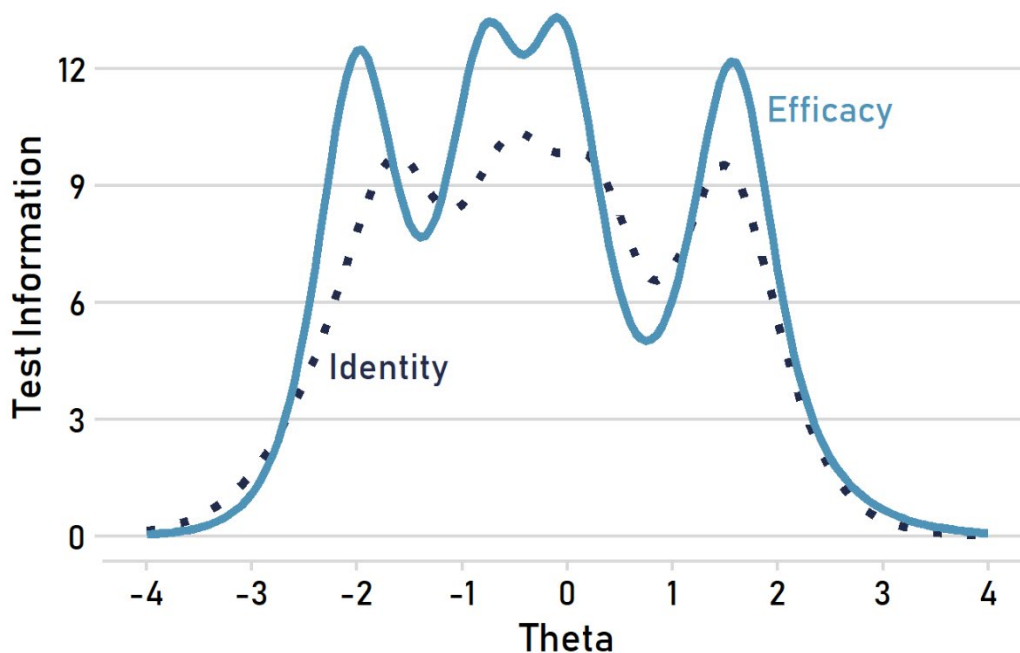
Using Syno International research panel, we requested 210 women and 210 men who were at least 18 years old and spoke Polish as their first language. While excluding the participants we followed the procedure applied for Study 1 and Study 2. Participants were excluded if they failed a directed response item (“Please select strongly disagree”); 21 people (5% of the initial sample) were excluded due to this reason. Additional exclusions were based on elevated Mahalanobis’s D values and careless, long-string responding using a trap item (“I’m not a funny person”); 14 people (3% of the initial sample) were excluded for these reasons.

Study 4

The sampling and screening approach in Study 4 was identical to Study 2, except we requested a final sample of 410 people divided evenly between men and women, and we added an approval rate in past studies of at least 80% as an additional criterion.

3. Information Functions

OSM Figure 1. Test information functions for the humor efficacy and identity scales.



4. Similarity of Polish and English Versions of the HEISS

We explored the similarity of the Polish and English versions of the HEISS via an item response theory framework (IRT) using analyses of differential item functioning (DIF). A virtue of an IRT approach to measurement invariance is the ability to explore uniform and non-uniform forms of item bias. For example, group membership may create a consistent item bias that is the same across all regions of the underlying trait, but it may have a non-uniform effect, such as no bias at low level of the trait but appreciable bias at high levels. Furthermore, DIF catches the core of item bias: whether members of different groups with the same underlying trait level have the same likelihood of giving a particular item response.

Using the same approach used to evaluate gender DIF described in Study 1 of the main article, we combined the data from all four samples and compared the Polish-speaking participants ($n = 400$) to the English-speaking participants ($n = 2546$). We estimated if language groups varied in uniform or non-uniform DIF for the HEISS efficacy and identity scales, using a threshold of 1% of the variance (via McFadden's R^2) as a criterion for flagging items for DIF. No items were flagged for total DIF using the 1% criterion, so the items do not appear to favor either of the language groups.

5. Model Fit for Scales Used in Studies 2-4

Scale	χ^2 (df)	CFI	RMSEA [90% CI]	SRMR
S2: BFI-2	985.81 (395), $p < .001$.761	.070 [.065, .076]	.083
S3: BFI-2	3076.04 (1165), $p < .001$.735	.065 [.062, .068]	.092
S3: Dark Triad	743.12 (51), $p < .001$.754	.188 [.176, .200]	.102
S3: SSCS	133.31 (43), $p < .001$.937	.074 [.060, .088]	.048
S4: HSQ	1019.42 (458), $p < .001$.838	.055 [.051, .060]	.070
S4: SMAP	9.11 (5), $p = .105$.994	.045 [.000, .091]	.017
S4: SANU	7.71 (2), $p = .021$.990	.085 [.028, .151]	.016

Note. The CFA for the dark triad had poor convergence.

6. Tables of Descriptive Statistics and Correlations

OSM Table 1

Study 2 Descriptive Statistics and Correlations

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7
1. Humor Efficacy	2.36	.83	1						
2. Humor Identity	2.37	.83	.75	1					
3. N	2.88	.92	-.20	-.02	1				
4. E	2.95	.75	.35	.25	-.34	1			
5. O	3.63	.72	.32	.20	-.01	.23	1		
6. A	3.77	.67	.10	.07	-.15	.16	.10	1	
7. C	3.56	.77	-.02	-.10	-.32	.25	.01	.20	1

Note. *n* = 304. See Study 2 text for details about the sample and measures used.

OSM Table 2

Study 3 Descriptive Statistics and Correlations

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12
1. Humor Efficacy	3.43	.83	1											
2. Humor Identity	3.50	.83	.73	1										
3. N	2.96	.84	-.36	-.20	1									
4. E	3.32	.78	.59	.46	-.45	1								
5. O	3.67	.49	.37	.31	-.18	.38	1							
6. A	3.93	.58	.28	.30	-.26	.38	.40	1						
7. C	3.75	.59	.08	.01	-.14	.21	.17	.31	1					
8. Creative Efficacy	5.21	.90	.46	.34	-.31	.46	.60	.38	.31	1				
9. Creative Identity	5.32	.99	.44	.40	-.18	.40	.60	.39	.23	.81	1			
10. Narcissism	2.90	1.15	-.03	-.01	.10	-.09	-.02	-.46	-.25	-.07	-.08	1		
11. Machiavellianism	2.82	1.21	.01	.03	.19	-.02	-.04	-.45	-.26	-.05	-.09	.83	1	
12. Psychopathy	2.92	1.18	-.01	.03	.22	-.01	-.07	-.38	-.22	-.08	-.06	.80	.83	1

Note. *n* = 385. See Study 3 text for details about the sample and measures used.

OSM Table 3

Study 4 Descriptive Statistics and Correlations

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8
1. Humor Efficacy	3.34	.86	1							
2. Humor Identity	3.39	.86	.69	1						
3. Playfulness (SMAP)	2.82	.64	.46	.57	1					
4. Need for Uniqueness (SANU)	2.73	.80	.19	.31	.34	1				
5. HSQ Affiliative	3.77	.68	.72	.71	.58	.19	1			
6. HSQ Self-enhancing	3.30	.68	.46	.46	.44	.27	.48	1		
7. HSQ Aggressive	2.65	.62	.19	.24	.22	.13	.23	.14	1	
8. HSQ Self-defeating	2.91	.73	.13	.37	.29	.25	.29	.25	.38	1

Note. $n = 400$. See Study 4 text for details about the sample and measures used.

7. Model Fit and Model Comparisons for the HEISS

	χ^2		CFI		RMSEA [90% CI]		SRMR		Wald
	One Factor	Two Factors	One Factor	Two Factors	One Factor	Two Factors	One Factor	Two Factors	
Study 1	843.811	206.825	.889	.975	.150 [.141, .158]	.073 [.064, .082]	.055	.026	209.50, $p < .001$
Study 2	144.670	66.096	.905	.964	.143 [.122, .166]	.090 [.067, .115]	.053	.038	32.98, $p < .001$
Study 3	167.287	38.91	.907	.987	.138 [.119, .158]	.052 [.028, .076]	.051	.022	45.20, $p < .001$
Study 4	287.111	113.514	.836	.942	.183 [.164, .202]	.112 [.092, .132]	.074	.053	41.43, $p < .001$

Note. Degrees of freedom for χ^2 are 20 and 19 for the one and two factor models, respectively.

8. OSM References

Yentes, R., & Wilhelm, F. (2021). *careless: Procedures for computing indices of careless responding*. R package version 1.2.1. <https://cran.r-project.org/package=careless>