I Should Have Known Better: Development of a Self-Report Measure of Gullibility

Supplementary Materials

Study 1: Exploratory Factor Analysis of Gullibility Items

Data Preparation

After reverse scoring relevant items, the distribution of each gullibility item was examined, and one item removed due to extreme skew. Based on the correlation matrix of the remaining 65 items, eight items were removed as they had either zero or only one correlation greater than r=.3 with the other items.

Exploratory Factor Analysis

After running an initial principal axis factor (PAF) analysis, with oblimin rotation on the data, sixteen items with factor loadings <.32 and low communalities were removed, as low communality scores can potentially distort data interpretation (Tabachnick & Fidell, 2007), leaving a total of 41 items. Eight factors had eigenvalues over one, meeting Kaiser's criterion. However, three of those factors had less than three items loading above .32 (and most of those items cross-loaded onto another factor). A parallel analysis with Principal Components analysis using the rawpar.sps program (O'Connor, 2000) with permutations on the original dataset suggested extracting three factors. However, the point of inflection on the scree plot suggested retaining 4-factors. Therefore, the PAF analysis was conducted extracting both 3- and 4-factor solutions. The 3-factor solution had a high proportion of cross-loading items and explained 40.49% of the variance, while the 4-factor solution explained 42.89% of the variance. Although the factors were interpretable, two items had factor loadings of less than .32, and four items had communalities below .25. Hence, these items were removed. The PAF was run again on the 35 items. The point of inflection on the scree plot suggested retaining 4-factors. Only four factors had eigenvalues over one, meeting Kaiser's criterion, explaining 46.37% of the variance. Table 1 shows factor loadings after rotation (suppressing values below .32) and Table 2 shows

descriptives for each item. All items have reasonable variability and there are no ceiling or floor effects.

Table 1

Study 1: Rotated Factor Matrix of the 4-Factor Gullibility Scale

	1	2	3	4
People think I'm a little naïve	.745			
My family thinks I am easily led	.704			
If anyone is likely to fall for a scam, it's me	.642			
I am probably a little too quick to believe others	.632			
I guess I am more gullible than the average person	.607			
My family think I am an easy target for scammers	.587			
I often fall for things when I should know better	.586			
My friends think I'm easily fooled	.583			
Overall, I'm pretty easily manipulated	.543			
My friends think I'm too trusting	.506			
When debating an idea, I am easily convinced of another person's point of view	.499			
People say I will agree to anything	.465			
I believe most people are honest		.722		
I believe most people can be relied upon to keep their word		.692		
I trust what people say		.654		
Usually people don't try to take advantage of others		.603		
If you are not careful, people will try to take advantage of you*		.585		
Completely trusting someone is asking for trouble*		.540		
I believe people are sincere when they flatter me		.535		
When people compliment me, it is because they want something from me*		.517		
Most people only look out for themselves*		.517		
People are usually honest in all aspects of their lives		.509		
I am often surprised when people are untrustworthy		.450		
I am often put in a situation where I have to pay for others			.760	
People often take advantage of my generosity			.646	
I usually offer to pay for others, even when I don't have much money			.580	
People often use me to get what they want			.522	
I have been persuaded to make donations to charities when I couldn't really afford it			.519	
I often end up doing other people's work			.501	
I'm pretty good at working out when someone is trying to fool me*				.827
I'm usually quick to notice when someone is trying to cheat me*				.793
It usually takes me a while to 'catch on' when someone is deceiving me				.674
I'm pretty poor at working out if someone is tricking me				.672
I'm not that good at reading the signs that someone is trying to manipulate me				.647
I quickly realize when someone is pulling my leg*				.551

Note. * denotes a reverse-scored item

Table 2

Study 1: Item-level descriptive statistics of the 4-Factor Gullibility Scale

	1	2	3	4
People think I'm a little naïve	3.51 (1.60)			
My family thinks I am easily led	3.01 (1.53)			
If anyone is likely to fall for a scam, it's me	2.29 (1.30)			
I am probably a little too quick to believe others	3.89 (1.47)			
I guess I am more gullible than the average person	3.18 (1.51)			
My family think I am an easy target for scammers	2.32 (1.38)			
I often fall for things when I should know better	3.47 (1.50)			
My friends think I'm easily fooled	2.88 (1.42)			
Overall, I'm pretty easily manipulated	2.79 (1.40)			
My friends think I'm too trusting	3.66 (1.58)			
When debating an idea, I am easily convinced of another person's point of view	3.41 (1.49)			
People say I will agree to anything	2.80 (1.45)			
I believe most people are honest		4.32 (1.31)		
I believe most people can be relied upon to keep their word		4.33 (1.92)		
I trust what people say		4.55 (1.23)		
Usually people don't try to take advantage of others		3.67 (1.39)		
If you are not careful, people will try to take advantage of you*		3.08 (1.31)		
Completely trusting someone is asking for trouble*		3.85 (1.60)		
I believe people are sincere when they flatter me		4.08 (1.43)		
When people compliment me, it is because they want something from me*		4.44 (1.43)		
Most people only look out for themselves*		3.39 (1.42)		
People are usually honest in all aspects of their lives		3.54 (1.36)		
I am often surprised when people are untrustworthy		3.76 (1.48)		
I am often put in a situation where I have to pay for others			3.10 (1.55)	
People often take advantage of my generosity			4.04 (1.59)	
I usually offer to pay for others, even when I don't have much money			4.06 (1.65)	
People often use me to get what they want			3.24 (1.39)	
I have been persuaded to make donations to charities when I couldn't really afford it			2.81 (1.72)	
I often end up doing other people's work			3.72 (1.58)	
I'm pretty good at working out when someone is trying to fool me*				3.12 (1.16)
I'm usually quick to notice when someone is trying to cheat me*				3.14 (1.26)
It usually takes me a while to 'catch on' when someone is deceiving me				3.39 (1.40)
I'm pretty poor at working out if someone is tricking me				3.15 (1.39)
I'm not that good at reading the signs that someone is trying to manipulate me				3.16 (1.49)
I quickly realize when someone is pulling my leg*				3.27 (1.28)

Note. Means for each item are presented with the standard deviations in the parentheses. * denotes a reverse-scored item

Study 2: Confirmatory Factor Analysis of the Gullibility Scale

Confirmatory Factor Analysis

A correlation matrix of all 35 items showed that none of the correlations exceeded .7, indicating no evidence of multicollinearity or singularity (Tabachnick & Fidell, 2007). Several goodness-of-fit indices were used. The ratio of the χ^2 to degrees of freedom was used to minimize the effect of sample size. Wheaton, Muthen, Alwin, and Summers (1977) recommend a figure of five or less for this ratio, whereas Tabachnick and Fidell (2007) recommend a more stringent figure of less than two. Other goodness-of-fit measures included the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Akaike Information Criterion (AIC), and the Root Mean Square Error of Approximation (RMSEA). Both CFI and TLI range from zero to one and Hu and Bentler (1999) recommend that values above .95 indicate a well-fitting model. The RMSEA also ranges from zero to one, but lower values represent a better model fit, with values below .06 deemed acceptable (Hu & Bentler, 1999). The AIC includes a parsimony adjustment; smaller values suggest a better fitting and more parsimonious model (Akaike, 1974; Tabachnick & Fidell, 2007).

See Table 3 for all the models tested. First, the common factor model (Model 1) was tested, wherein all the 35 items were constrained to a single latent factor. This model did not fit the data well. Although the model's ratio of χ^2 to degrees of freedom was less than five, not all the items had significant loadings onto the latent variable. Overall, the common factor model was not a good fit for the data and provided further evidence to suggest that gullibility comprises more than one latent factor.

Table 3

Study 2: Fit Indices for Various Models

No.	Model	χ^2	*df*	χ^2/df	RMSEA	CFI	TLI	AIC
1	Common factor model	2512.32*	560	4.49	.10	.60	.58	2652.32
2	Four-factor model	1170.22*	554	2.11	.06	.87	.86	1322.22
3	Three-factor model (minus Trustability)	495.56*	249	1.99	.06	.93	.92	597.56
4	Common factor model (minus Trustability items)	797.66*	252	3.17	.08	.84	.83	893.66
5	Common factor model (minus Trustability and Unassertiveness)	376.99*	135	2.79	.07	.92	.91	448.99
6	Two-factor model (minus Trustability and Unassertiveness items)	272.05*	134	2.03	.06	.95	.95	346.05
7	Common factor model (12 item scale)	202.94*	54	3.76	.09	.93	.91	250.94
8	Two-factor model (2 factors with 6 items per factor)	129.46*	53	2.44	.07	.96	.95	179.46

Note. * *p*<.0005

For Model 2, the four-factor model, the ratio of χ^2 to degrees of freedom was less than five, all the items loaded onto the latent variables significantly, and the AIC was lower than for Model 1. Although this model fit the data well, the standardized covariance between Persuadability and Trustability was not significant (.04, *p*=.346). Similarly, the covariance between Insensitivity and Trustability was not significant (-.05, *p*=.363). After removing the Trustability factor and its associated items (see Model 3, the three-factor model), the AIC reduced and the ratio of χ^2 to degrees of freedom was less than two. The remaining goodness-of-

fit indices were closer the cut-off of .95 (CFI=.93, TLI=.92), with RMSEA=.06. The common factor model, without the items associated with the Trustability factor was also tested (i.e., Model 4), however the three-factor model was a better fit. Consistent with Study 1, the Trustability factor did not relate strongly to the remaining three factors. As the Trustability factor items did not improve the model fit, while adding unnecessary complexity, the Trustability items were removed from the Gullibility Scale.

Although this three-factor model fit the data well, four further models were tested based on the correlations between the factors and variables, as well as theoretical grounding of the concept. The Unassertiveness factor had only a weak to moderate correlation with the other factors (see Table 2 in the paper). Furthermore, this factor and its associated items did not fit the theoretical understanding of gullibility. Therefore, a common factor model without the Trustability and Unassertiveness factor items (Model 5) and a two-factor model (without the Trustability and Unassertiveness factors; Model 6) were tested. Again, the two-factor model fit better than the common factor model, and the two-factor model fit the theoretical understanding.

Reducing the scale to 18 items made both theoretical sense (definitions of unassertiveness and trust did not relate to gullibility) and was supported by the data. However, this led to an unbalanced scale. There was unnecessary conceptual duplication in the Persuadability factor. Therefore, based on the factor loadings, the six lowest scoring items from that factor were removed to balance the number of items in each factor. The common factor model (Model 7) with 12 items fit moderately well, but the 2-factor model (Model 8) fit well and made theoretical sense. After removing items that loaded onto the Trustability factor and the Unassertiveness factor as well as six of the Persuadability items, the Gullibility Scale now comprised 12 items. Cronbach's alpha increased to .92 for the overall scale. Table 4 presents the factor loadings for each item onto their respective factors.

Table 4

Study 2: Factor Loadings for the Final 12-item Gullibility Scale

	P	I
I guess I am more gullible than the average person	.789	
If anyone is likely to fall for a scam, it's me	.758	
My friends think I'm easily fooled	.847	
My family thinks I am an easy target for scammers	.740	
People think I'm a little naïve	.737	
Overall, I'm pretty easily manipulated	.780	
I'm pretty good at working out when someone is trying to fool me*		.587
I'm not that good at reading the signs that someone is trying to manipulate me		.659
I'm pretty poor at working out if someone is tricking me		.738
It usually takes me a while to 'catch on' when someone is deceiving me		.817
I'm usually quick to notice when someone is trying to cheat me*		.564
I quickly realize when someone is pulling my leg*		.591

Note. * denotes a reverse-scored item, P = Persuadability, I = Insensitivity

Study 3: Exploratory and Confirmatory Factor Analysis of the 2-Factor Gullibility Scale

Results and Discussion

Exploratory Factor Analysis

A PAF analysis with oblimin rotation was conducted on the 12 items (see supplemental material for details). The point of inflection on the scree plot suggested a two-factor solution explaining 56.50% of the variance with only one item loading significantly onto both factors. Only two factors had eigenvalues over one, meeting Kaiser's criterion. A parallel analysis with Principal Components analysis using the rawpar.sps program by O'Connor (2000) also suggested a two-factor solution.

Confirmatory Factor Analysis

Table 5 presents the factor loadings for each of the items onto their respective factors. Overall, the factor structure found in Study 2 was replicated.

Table 5

Study 3: Standardized Regression Weights for the Final 12-item Gullibility Scale

	P	I
I guess I am more gullible than the average person	.783	
If anyone is likely to fall for a scam, it's me	.686	
My friends think I'm easily fooled	.790	
My family thinks I am an easy target for scammers	.605	
People think I'm a little naïve	.663	
Overall, I'm pretty easily manipulated	.809	
I'm pretty good at working out when someone is trying to fool me*		.675
I'm not that good at reading the signs that someone is trying to manipulate me		.723
I'm pretty poor at working out if someone is tricking me		.812
It usually takes me a while to 'catch on' when someone is deceiving me		.802
I'm usually quick to notice when someone is trying to cheat me*		.643
I quickly realize when someone is pulling my leg*		.563

Note. * denotes a reverse-scored item, P = Persuadability, I = Insensitivity

Measurement Invariance

To ensure measurement invariance of the scale across the groups (i.e., males and females, students and community members), the four samples from Studies 1-3 were combined. Testing for measurement invariance allows researchers to determine if items are perceived the same way by members of different groups (Cheung & Rensvold, 2002) such as participant type or gender. This combined sample (N=1523) provided adequate sample size for the smaller groups such as males (N=323) and community members (N=481) whereas the within-study samples were smaller. A multiple-groups factor analysis in AMOS on the final 12-item scale for males and females as well as students and community members found there was multiple-groups

measurement invariance (gender: χ^2=4.144, p=.941, participant type: χ^2=6.770, p=.747). This suggests that although there may sometimes be significant differences between group members in particular studies, when this difference is examined with larger samples through a Multigroup CFA, this difference disappears.

Test-Retest Reliability

A new sample of 60 undergraduate students, 15 males and 45 females with an average age of 22.07 years (*SD*=2.00), completed the 12-item Gullibility Scale at two time-points, 12 weeks apart, as part of a classroom exercise. The Intraclass Correlation Coefficient between the two time-points was .80 (95% CI .69-.86, p<.0005). This figure is above the recommended .70 (Terwee et al., 2007) demonstrating that the Gullibility Scale has excellent test-retest reliability over 3 months.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*, 396–402. https://doi.org/10.3758/BF03200807

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., … de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*, 84–136. https://doi.org/10.2307/270754