

What is honesty? Laypersons interpret high lie-scale scores
as reflecting intentional dishonesty:

Rejoinder to de Vries et al.'s (2017) comment on Feldman et al. (2017)

Gilad Feldman

First affiliation: Department of Psychology

University of Hong Kong

Hong Kong SAR

giladfel@gmail.com

Second affiliation: Department of Work and Social Psychology

Maastricht University

Maastricht, 6200MD, The Netherlands

gilad.feldman@maastrichtuniversity.nl

Word count: Abstract – 142; Manuscript excluding tables – 2975.

Please note: Manuscript has been accepted to *Social Psychological and Personality Science* on September 21th 2017. This is an unedited pre-print. Final version may include minor corrections.

Corresponding author

Gilad Feldman, Department of Psychology, University of Hong Kong, Hong Kong SAR,
giladfel@gmail.com

AUTHOR BIOS:

Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on morality, personal values, lay-beliefs, and judgment decision-making.

Declaration of Conflict of Interest:

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

financial disclosure/funding:

The author(s) received no financial support for the research and/or authorship of this article.

Acknowledgments or Author Notes:

[None]

Authorship declaration

Single author submission.

Abstract

Scholars have recently been questioning the original premise of lie-scales as measuring dishonesty for social desirability, with some even claiming that lie-scales may in fact represent higher honesty (e.g., de Vries et al., 2017 commentary on Feldman, Lian, Kosinski, & Stillwell, 2017). In a pre-registered experiment, I examined the relationship between lie-scales and honesty by directly assessing laypersons' perceptions of honesty reflected in lie-scales. Overall, laypersons perceived higher lie-scale scores as reflecting higher dishonesty and higher lie-scale scorers as more intentionally dishonest and socially sensitive, endorsing a negative (and rejecting a positive) relationship between lie-scales and honesty. These findings provide empirical support for lie-scales as capturing dishonesty. I discuss implications and possible directions for addressing the complexity inherent in the construct of honesty and the need for an integration of the seemingly contradictory findings to advance the debate regarding lie-scales and honesty.

Keywords: lie scale; impression management; honesty; dishonesty; truth

What is honesty? Laypersons interpret high lie-scale scores as reflecting intentional dishonesty:

Rejoinder to de Vries et al.'s (2017) comment on Feldman et al. (2017)

Introduction

There is an ongoing debate regarding what lie scales measure. Lie scales, sometimes referred to as impression management, other-deception, or the Gamma-factor scales, are associated with mixed and confusing findings that make it difficult to interpret how well they capture honesty or what aspects of personality they reflect (Uziel, 2010). Traditionally, lie-scales were constructed to measure deceit for appearing more socially desirable. Later accounts interpreted high scores on lie-scales as capturing aspects of response styles and personality traits. The “defensiveness approach” argued lie-scales are capturing defensiveness of the self, reflecting distortions in self-evaluations and avoidance of social disapproval (Crowne, 1979). The “adjustment approach” considered high lie-scale scores as reflecting healthy social functioning, and such response styles as representing individuals who emphasize social adjustment and coexistence with others in society. This evolving debate exemplifies the different interpretations regarding lie-scales and lying. Do lie-scales indeed capture lying, as their title suggests, or perhaps lie-scales have been completely misunderstood and misinterpreted and are actually measuring something else altogether, possibly even measuring honesty.

This debate has recently reemerged with a commentary criticizing the use and interpretation of a lie-scale as a measure of honesty in a study of the relationship between profanity and honesty (Feldman et al., 2017). In the first out of three studies, my colleagues and I found that two behavioral and one self-reported measures of profanity had a negative relationship with lie-scale score using the Eysenck Personality Questionnaire Revised (Eysenck,

Eysenck, & Barrett, 1985), and based on the original premise of these scale and a large body of supporting evidence (e.g., Davis, Thake, & Weekes, 2012) our interpretation of this evidence was that those who were higher on profanity use were more honest. De Vries et al. (2017) had strong objections to this interpretation and summarized recent evidence which suggests that lie-scales capture honesty, rather than dishonesty. They therefore interpreted our evidence to mean that “honest people tend to use less - not more - profanity” (commentary title). They concluded by calling for a “moratorium on the use of lie scales as measures of deception and dishonesty” (p. 5).

In Table 1 I briefly address the main points in the commentary. In the target article, we called for further debate and empirical evidence and cautioned against drawing strong conclusions, and I am happy to see further evidence on the topic. Yet, the commentary's accompanying criticism is often misplaced or unwarranted, the evidence suggestive, and the strong conclusions premature. In this reply I will focus on the main criticism – that high lie-scale scores cannot be interpreted as reflecting dishonesty. I argue that the debate on lie-scales is far from conclusion, and much more needs to be clarified and examined regarding lie-scales and honesty. I first review weaknesses in the arguments against interpreting lie-scales as dishonesty, and I then provide strong experimental evidence showing that laypersons understand and interpret high lie-scale scores as being intentionally dishonest and reflecting social sensitivity. Finally, I discuss implications for the lie-scale and honesty literatures, and suggest directions for integration.

Table 1

Brief responses to the de Vries et al commentary critics

Commentary arguments	Response
"People who use profanities, both online and offline, may feel exonerated by the study "	The comment's use of "exonerated" and the opening note seems to reflect a stance against profanity as being a negative behavior. We were cautious to refrain from any positive or negative judgment of profanity or its users.
"Political representatives may even be stimulated to increase their use of profanities hoping that it may increase the public's perception of their integrity."	The target article did not address or discuss impressions of honesty. We specifically cautioned against such interpretations in the general discussion.
"Feldman et al. (2017) [...] gathered three self-report measures of people's profanity use"	Correction. We used one self-reported measure of profanity and two behavioral measures of profanity.
Recent evidence has indicated that an interpretation of Impression Management Scales as reflecting less honesty is fundamentally flawed	The recent evidence has not questioned previous evidence as being flawed, but presented new evidence about links between impression management scales and other factors that is suggestive of a different interpretation.
"If Impression Management (or Lie) Scales actually did measure lying on items such as "Do you always practice what you preach?," the correlation between self-ratings and ratings obtained from knowledgeable others would have to be negative."	Two issues with this assumption: (1) the very purpose of impression-management is to affect impressions in others, (2) people are very bad at detecting lies, and there is ample evidence that close others are even more biased and easily deceived (e.g., McCornack & Parks, 1986; Levine & McCornack, 1992). Therefore, self-other agreements are not a good measure of objective truth.
"Balanced Inventory of Desirable Responding (BIDR) Impression Management Scale (Paulhus, 2002), strongly correlated (e.g., $r = .61$; Davies, French, & Keogh, 1998) with, and similar in content to, the Lie Scale used in Feldman et al.'s (2017) Study 1"	I welcome reporting of related evidence, yet note that the target article did not use the BIDR nor the Davies et al. (1998) scale. The items and the response scale in the BIDR are very different. Davies et al. changed the scale range used by the lie-scale from yes-no to a Likert scale. See the general discussion.
Lie Scale used by the target article has been found to be positively associated with trait honesty, as measured by HEXACO-PI-R honesty-humility	I welcome reporting of related evidence, yet note that lie-scales and trait honesty-humility may not capture the same type of honesty. See the general discussion.

Commentary arguments	Response
Scores on Impression Management Scales have been found to be negatively correlated with scores on objective behavioral indicators of dishonesty	<p>The target article specifically cautioned against interpreting lie-scales as unethicity, and about generalizing findings to mean anything about unethicity. Lie-scales and unethicity measures do not capture the same type of honesty. See the general discussion.</p> <p>I welcome reporting of related evidence, yet caution that the evidence provided is suggestive, as the profanity measure is confounded with the impression management scale (BIDR) and was not meant to measure profanity.</p> <p>The commentary authors complained about the target's Study 2 construct validity but make very strong claims using a re-definition of a single item from an unrelated scale typically used for a different purpose using data archive not originally collected for the purpose of the intended profanity-honesty investigation.</p> <p>Lastly, a correlation of -.09 using a problematic self-reported profanity proxy measure against an unethicity proxy for group aggregates based on probabilities estimations (25%) should - at best - be interpreted with caution and followed by further research before drawing strong conclusions.</p>
"In summary, the above findings strongly suggest that the conclusions of Feldman et al. (2017) are incorrect"	<p>The strong statement is unwarranted given the evidence provided. The commentary reflects one of several perspectives and a specific body of evidence led by the commentary authors, which is not fully representative of the ongoing debate on lie-scales and the complexities of honesty. See general discussion for possible directions for an integration of the mixed findings.</p>

Lie scales and honesty

Lie scales measure unlikely answers to extreme statements such as in answering “yes” to “are *all* your habits good and desirable ones?” or answering “no” to questions such as “have you *ever* said anything bad or nasty about anyone?” (Eysenck & Eysenck, 1975; Eysenck et al., 1985). The premise was that those who answer yes-no to the questions above – for whatever reason - are most likely being deceitful, possibly in service of some purpose or reflecting certain traits.

There is an ongoing debate regarding the meaning of lie-scales, and the question of what lie-scales or impression-management measures capture is still in debate (Uziel, 2010). Scholars in the “adjustment view” (e.g., de Vries, Zettler, & Hilbig, 2014; Dunlop, Morrison, Koenig, & Silcox, 2012; Zettler, Hilbig, Moshagen, & de Vries, 2015) have come to question the original premise of lie-scales as measuring dishonesty and the “defensiveness view” (e.g., Davis et al., 2012) based on two types of correlational evidence. First, they demonstrated that there is a positive correlation between lie-scales and various measures of honesty, such as the HEXACO honesty-humility personality trait, and measures of unethical behavior. Second, they demonstrated high correlations between self ratings and close others’ ratings of lie-scales items. Both are important contributions, yet they are not necessarily evidence against lie-scales as representing dishonesty, and I explain further below.

What is honesty? Discussion of honesty categories

Lie-scales are at the intersection of two conflicting interpretations of honesty – honesty that is meant to express true-self, and honesty that is meant to reflect coexistence with others in society (Erickson, 1995; see related discussion in Feldman et al., 2017). Lie-scales were designed to measure the sense of being honest about and true to oneself, reflecting authenticity-honesty. However, honesty trait measures, such as the honesty-humility dimension of the HEXACO personality scale (Lee & Ashton, 2004) measure sincerity, fairness, greed-avoidance, and modesty, all socially desirable and focusing on social aspects of honesty in capturing attitudes towards manipulation of others and towards blunt unethical behavior violating social norms and laws. These types of honesty are very different from authenticity-honesty. A suggested typology of the different types of honesty and examples from related scales are provided in Table 2.

To give an example for a possible misalignment, authentic-honest persons could act unethically if their true selves are self-centered and deviate from the set norms. In such a case, a person may cheat but would answer more truthfully about his cheating behavior. Similarly, persons can adhere to social-moral codes, yet in doing so may present themselves dishonestly or conform by doing things that they do not wish to do and would not admit to not wanting. In such a case, a socially-honest person may be insincere about own wants, needs, desires, and emotions. It is therefore possible to differentiate between at least two types of honesty, a distinction that is not clearly mentioned or addressed in the lie-scales debate. Therefore, studies examining authenticity-honesty and studies examining social-honesty are not necessarily studying the same kind of honesty. Findings of lie-scale showing a negative relationship with unethical behavior or positive relationship with honesty-humility traits do not necessarily mean that lie-scales do not measure lying or that they may even measure honesty. It is possible that lie-scales simply capture a different type of honesty.

Demonstrating self-other agreement on lie-scales is also problematic as evidence against lie-scales, since the very purpose of impression-management is to affect impressions in others, and there is ample evidence that others are biased and easily deceived (e.g., McCornack & Parks, 1986; Levine & McCornack, 1992). I return to this point in the discussion.

The present investigation

It is therefore important to directly examine whether lie-scales indeed capture dishonesty, and specifically test authentic-dishonesty as the type of honesty they were meant to assess. I take a different approach than that taken so far regarding lie-scales. Rather than making assumptions about what laypersons mean when they refer to honesty and do when answering lie-scales, or indirectly interpreting lie-scales based on correlates or using self-other comparisons, my

approach uses experimental folk psychology methodology (Knobe et al., 2012; Knobe & Nichols, 2013; Nichols, 2011; see example regarding honesty in Turri & Turri, 2015) to more directly assess laypersons' cognition regarding honesty and lie-scales.

In a pre-registered experiment, I examined laypersons' interpretations of high and low scores and scorers in terms of honesty, to test (1) whether high versus low scores answers are perceived as more or less honest answers, and (2) whether persons with a high versus low scores are perceived as more or less honest people.

Table 2

Honesty measures summary

Measure	Sample items	General explanation	Scale example
Lie scale	Are <i>all</i> your habits good and desirable ones?" Have you <i>ever</i> said anything bad or nasty about anyone? (0 – <i>No</i> ; 1 – <i>Yes</i>)	Social desirability others-oriented deception	Eysenck Personality Questionnaire Revised (Eysenck et al., 1985):
Honesty-humility: Sincerity	"If I want something from a person I dislike, I will act very nicely toward that person in order to get it" "I wouldn't use flattery to get a raise or promotion at work, even if I thought it would succeed" (1 – <i>Strongly disagree</i> ; 5 – <i>Strongly agree</i>)	"low scorers will flatter others or pretend to like them in order to obtain favors, whereas high scorers are unwilling to manipulate others"	HEXACO-PI-R (Lee & Ashton, 2004)
Honesty-humility: Fairness	I would be tempted to buy stolen property if I were financially tight If I knew that I could never get caught, I would be willing to steal a million dollars (1 – <i>Strongly disagree</i> ; 5 – <i>Strongly agree</i>)	"Low scorers are willing to gain by cheating or stealing, whereas high scorers are unwilling to take advantage of other individuals or of society at large"	HEXACO-PI-R (Lee & Ashton, 2004)
Unethicality	Participants roll a six-faced die under a cup with a peeking hole only available to the participants. Participants self-report the outcome. Rewards are given based on rolls, presenting participants with an opportunity to misreport results to gain higher rewards, without any risk of getting caught.	Behavioral measure of lying for monetary gain.	e.g., die-under-cup paradigm (Shalvi, Dana, Handgraaf, & De Dreu, 2011)

Experiment

Pre-registration and open-science

I pre-registered the experiment on the Open Science Framework and data collection was launched later that day. Pre-registration, power analyses, and all materials are available in the supplementary materials and shared together with data and code on the Open Science Framework (<https://osf.io/7gnfe/>).

Participants

A total of 302 American Amazon Mechanical Turk (MTurk) participants ($M_{age} = 36.79$, $SD_{age} = 12.21$; 150 female) were recruited online using TurkPrime.com (Litman, Robinson, & Abberbock, 2017). MTurk participants have considerable experience in taking part in surveys and experiments, and I therefore expected that some may have taken the lie-scale before. To prevent any possible biases, I pre-registered the exclusion of participants who indicated that they were familiar with or have taken the scale before. I excluded 40 participants in total (13.2%), leaving a sample of 262 ($M_{age} = 36.88$, $SD_{age} = 12.22$; 129 female), although exclusions had no effect over the results (full results reported in the supplementary).

Procedure and measures

The experiment included three parts. The first part was a within-subject experiment with each participant rating two lie-scale answer profiles – a high-score on the lie-scale and a low-score on the lie scale (see Table 3). The presentation order of the two profiles was randomized. Participants were asked to rate for each of the answer profiles whether a participant who answered in such a way: (1) answered truthfully, (2) is generally an honest person, (3) cares what others think of him/her, (4) understood the questions, (5) read the questions carefully, (6)

believes that his/her answers reflect the truth (intentionality), (7) answered as most people would, and (8) answered randomly (1 – *Extremely unlikely*; 7 – *Extremely likely*).

Table 3

Part 1: Description of answer profiles

Highest lie-scale score (Participant X)	Lowest lie-scale score (Participant Y)
If you say you will do something, do you always keep your promise no matter how inconvenient it might be? <u>yes</u>	If you say you will do something, do you always keep your promise no matter how inconvenient it might be? <u>no</u>
Are all your habits good and desirable ones? <u>yes</u>	Are all your habits good and desirable ones? <u>no</u>
Do you always practice what you preach? <u>yes</u>	Do you always practice what you preach? <u>no</u>
Were you ever greedy by helping yourself to more than your share of anything? <u>no</u>	Were you ever greedy by helping yourself to more than your share of anything? <u>yes</u>
Have you ever blamed someone for doing something you knew was really your fault? <u>no</u>	Have you ever blamed someone for doing something you knew was really your fault? <u>yes</u>
Have you ever taken anything (even a pin or button) that belonged to someone else? <u>no</u>	Have you ever taken anything (even a pin or button) that belonged to someone else? <u>yes</u>
Have you ever broken or lost something belonging to someone else? <u>no</u>	Have you ever broken or lost something belonging to someone else? <u>yes</u>
Have you ever said anything bad or nasty about anyone? <u>no</u>	Have you ever said anything bad or nasty about anyone? <u>yes</u>
As a child were you ever cheeky to your parents? <u>no</u>	As a child were you ever cheeky to your parents? <u>yes</u>
Have you ever cheated at a game? <u>no</u>	Have you ever cheated at a game? <u>yes</u>
Have you ever taken advantage of someone? <u>no</u>	Have you ever taken advantage of someone? <u>yes</u>
Do you sometimes put off until tomorrow what you ought to do today? <u>no</u>	Do you sometimes put off until tomorrow what you ought to do today? <u>yes</u>

In the second part, both answer profiles were presented side by side as representing two persons. Participants were asked to compare the two and were forced to choose which of the two persons: (1) is a more honest person, (2) has answered the questions more honestly, (3) cares more what others think of him/her, (4) better represents the average person's answers.

In the third part, I ran a between-subject design experiment and participants were randomly assigned to one of two conditions for the two conflicting interpretations of the lie-scale as either representing low or high honesty. Participants were asked to rate their agreement with the interpretation presented in their assigned condition (1 - *Strongly disagree*; 7 - *Strongly agree*).

Results

Means, standard deviations, and summary statistics for Part 1 are provided in Table 4. Participants perceived persons with low lie-scale score as answering more truthfully ($d = 1.55$), as believing their answers are truthful ($d = 1.13$), as being more honest ($d = 0.59$), and as caring less about what others think of them ($d = 2.10$). Both answer profiles were perceived as reflecting a good understanding of the items and as intentional non-random responding (understanding: $M = 6.11/6.24$; careful reading: $M = 5.71/6.12$; random answering: $M = 1.95/2.03$; $d = 0.05$ to 0.32 ; on 1-7 scale; one-sample t-test comparing to 4 as neutral answer: all $|t| > 19.77$, $p < .001$).

Table 4

Part 1: Means, standard deviations, and summary statistics

	High lie-score		Low lie-score		<i>r</i>	<i>t</i>	<i>p</i>	<i>d_z</i>	<i>d_{rm}</i>	<i>CIL</i>	<i>CIH</i>
	<i>M_X</i>	<i>SD_X</i>	<i>M_Y</i>	<i>SD_Y</i>							
Answered items truthfully	3.50	2.10	6.10	1.19	-0.32	15.41	<.001	0.95	1.55	1.35	1.74
Believes answers reflect truth	4.28	1.96	6.10	1.20	-0.23	11.65	<.001	0.72	1.13	0.94	1.31
An honest person	4.16	1.92	5.25	1.75	-0.31	5.94	<.001	0.37	0.59	0.41	0.76
Cares what others think	6.17	1.20	3.13	1.66	-0.07	23.26	<.001	1.44	2.10	1.89	2.31
Understood the questions	6.11	1.19	6.24	1.19	0.37	1.58	.115	0.10	0.11	-0.06	0.28
Read questions carefully	5.71	1.40	6.12	1.11	0.26	4.30	<.001	0.27	0.32	0.15	0.49
Answered randomly	2.03	1.44	1.95	1.47	0.39	0.80	.422	0.05	0.05	-0.12	0.22
How most people would answer	4.43	1.88	3.81	1.82	-0.43	3.20	.002	0.20	0.34	0.17	0.51

Note. N = 262 pairs in a with-subject design; X = Participant X with highest lie-scale score, Y = Participant Y with lowest lie-scale score; SD = Standard deviation, *r* = correlation between X and Y, *d_{rm}* = Cohen's d effect-size, *d_{rm}* = Cohen d corrected for paired samples (see Lakens, 2013). CIL/CIH = low and high bound for *d_{rm}* 95% confidence intervals. Scales: 1 – *Extremely unlikely*; 7 – *Extremely likely*.

To examine the comparison between the two profiles in Part 2, I ran a binomial test contrasting proportions to a 50%-50% random split choice. Participants rated the low lie-scale score as reflecting a more honest person (187/262, 71% CI[65%,76%], binomial $z = 6.86$, $p < .001$), as answering items more truthfully (204/262, 78% CI[72%,82%], binomial $z = 8.96$, $p < .001$), and as caring less about what others think of them (42/262, 16% CI[12%,50%], binomial $z = -10.94$, $p < .001$). Ratings of the two profiles as representing the average person did not significantly differ from a 50-50 split (137 versus 125, 52% versus 48%, binomial $z = .68$, $p = .497$).

Finally, participants showed higher agreement with the interpretation that high lie-scale scores represent lower honesty (N = 131, $M = 5.60$, $SD = 1.57$ versus N = 129, $M = 3.27$, $SD =$

2.14, $t(234.12) = 10.03$, $p < .001$, $d = 1.24$ CI[0.97, 1.50]; one-sample t-test comparing means to 4 as neutral answer: both $|t| > 3.87$, $p < .001$).

Discussion

Laypersons perceived high lie-scale scores as less honest than low lie-scale scores and high lie-scale scorers as less honest people than low lie-scale scorers. Compared to low lie-scale scores, high lie-scale scorers were interpreted as intentionally deceitful and as reflecting caring more about what others think of them. These results were consistent across rating high and low lie-scale scores independently and contrasts in a comparison of the two answer profiles. Finally, laypersons endorsed the interpretation that high lie-scale scores reflect lower honesty and rejected the interpretation that high lie-scale scores reflect higher honesty.

These findings support the original premise of lie-scales as capturing dishonesty for socially desirability. These calls to question the very strong statements made about lie-scales as clearly not reflecting dishonesty, mainly based on correlates with other factors and self-other agreement (Uziel, 2014; de Vries et al., 2014). At the very least, our findings exemplify the need for a more complex and nuanced understanding of lie-scales than the clear-cut conclusions by de Vries et al., (2017) arguing that lie-scales most definitely do not reflect dishonesty. Furthermore, laypersons viewed high lie-scale scores as reflecting a clear understanding of the items, a clear strategic agenda in answering, and intentional dishonesty, contrary to the untested argument that laypersons process the “spirit or gist of the item [...] rather than exact wording” (De Vries et al., 2017, footnote 7).

Which raises the question - why would there be a negative relationship between lie-scales and measures of honesty-humility traits and ethical behavior?

There are two possible explanations, and both could be relevant. The first, as suggested in the introduction, is that lie-scales capture a different type of honesty than that measured by honesty-humility traits and unethicality measures. I will not reiterate this point, but I call scholars to differentiate between the types of honesty and to clearly identify which type of honesty they aim to capture and by what measure, noting generalizability of their findings, as we have in the target article. If we are to accept the possibility of different types of honesty, then it is likely that those with high lie-scale scores are high on authenticity-honesty, and possibly low on the social-honesty.

The second explanation is that the methodological approach in assessing honesty reflected by lie-scales using traits and self-other agreements has limitations in evaluating authenticity-honesty that lie-scales were originally meant to assess. Measuring honesty about oneself presents scholars with a unique challenge. Take, for example, the paradoxical dilemma embedded in lying self-reports (e.g., Halevy, Shalvi, & Verschuere, 2014; Peer, Acquisti, & Shalvi, 2014) - are people who self-report themselves as liars honest about lying? are people who self-report to being honest lying to appear honest? How, then, do we assess whether such self-reports about lying are honest? One approach taken by personality scholars was to examine correlates between such self-reports and various individual differences and behaviors. This approach would be useful to the extent that the correlated factors indeed capture the same construct, yet honesty-humility may not capture the same construct. A second approach was to assess self-other agreements. Yet again, such an approach also embeds a paradox. If the very purpose of impression management is to create more positive impressions in others, then how would asking about others' impressions helpful? The assumption of that approach was that close others are likely to know a person well and their opinions would therefore reflect an objective

truth, yet ample research has shown that people are generally bad at detecting and correcting for deception or for impression management, and that close others are even more easily deceived and are generally more biased (e.g., McCornack & Parks, 1986; Levine & McCornack, 1992). Close others are frequently lied to, even if lying is merely meant to protect them and their feelings, and are often surprised when learning that cheating or deceit has taken place (Drouin, Tobin, & Wygant, 2014; Ennis, Vrij, & Chance, 2008), most vividly exemplified by reactions to admissions of partners faking sexual interest, pleasure, or experience in close romantic relationships (e.g., Muehlenhard & Shippee, 2010).

An integration of the findings regarding lie-scales is possible with (1) a typology of types of honesty with a clear and careful use of the term, and by (2) recognizing and addressing the limitations of the correlates and self-other agreement approaches in understanding honesty-authenticity and lie-scales. To successfully achieve such integration further studies are needed employing different methods and including measures for the different types of honesty.

An integration would also require an explanation of findings for a negative relationship between lie-scales and honesty-humility or unethicity. Why would moral people lie about who they are? First, moral people may self-justify lies for social-moral reasons and may therefore even come to perceive such lying as acceptable or needed. Violations of morals and honesty are deemed acceptable if they can be easily justified (Shalvi, Gino, Barkan, & Ayal, 2015), especially if allegedly serving others (Gino, Ayal, & Ariely, 2013), and can help reduce ethical dissonance between the moral-self and dishonest behavior (Barkan, Ayal, & Ariely, 2015). Conformity to social moral norms may serve as a powerful reason for justifying white lies to the extent that such a lie could even be seen as moral. Second, moral people are more sensitive to others' views, behaviors, and criticism and tend to put others' concerns above their own. As such,

they are more likely to engage in impression-management to influence others' perceptions of them as being more moral, social, and positive.

Conclusion

Those low on lie-scales are perceived to be more honest about who they are and their true-selves, challenging the criticism that lie-scales cannot be reflective of dishonesty and that laypersons answer lie-scales do not perceive lie-scale items as lying. The debate on lie-scales and honesty demonstrates the complexity of defining and examining honesty. This raises the need for (1) a typology of honesty categories (e.g., differentiating between authenticity-honesty and social-honesty), (2) a clear identification of the type of honesty assessed or referenced in empirical studies, and (3) further research to better understand honesty and the relationships between honesty categories. The novel folk-psychology empirical approach used in this investigation can be used to further examine laypersons' honesty perceptions and behavior.

References

- Barkan, R., Ayal, S., & Ariely, D. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology, 6*, 157-161.
- Crowne, D. P. (1979). *The Experimental Study of Personality*. Halsted Press.
- de Vries, R. E., Hilbig, B. E., Zettler, I., Dunlop, P. D., Holtrop, D. J., Lee, K., & Ashton, M. C. (2017). Honest people tend to use less—not more—profanity: Comment on Feldman et al.'s (in press) Study 1. *Social Psychological and Personality Science*. Manuscript accepted for publication.
- de Vries, R. E., Zettler, I., & Hilbig, B. E. (2014). Rethinking trait conceptions of social desirability scales: Impression management as an expression of honesty-humility. *Assessment, 21*, 286-299.
- Davies, M. F., French, C. C., & Keogh, E. (1998). Self-deceptive enhancement and impression management correlates of EPQ-R dimensions. *The Journal of Psychology, 132*, 401-406.
- Davis, C. G., Thake, J., & Weekes, J. R. (2012). Impression managers: Nice guys or serious criminals?. *Journal of Research in Personality, 46*(1), 26-31.
- Drouin, M., Tobin, E., & Wygant, K. (2014). “Love the Way You Lie”: Sexting deception in romantic relationships. *Computers in Human Behavior, 35*, 542-547.
- Dunlop, P. D., Morrison, D. L., Koenig, J., & Silcox, B. (2012). Comparing the Eysenck and HEXACO models of personality in the prediction of adult delinquency. *European Journal of Personality, 26*, 194-202.
- Ennis, E., Vrij, A., & Chance, C. (2008). Individual differences and lying in everyday life. *Journal of Social and Personal Relationships, 25*, 105-118.

- Erickson, R. J. (1995). The importance of authenticity for self and society. *Symbolic Interaction, 18*, 121-144.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. Hodder & Stoughton, London.
- Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences, 6*, 21-29.
- Feldman, G., Lian, H., Kosinski, M., & Stillwell, D. (2017). Frankly, we do give a damn: The relationship between profanity and honesty. *Social Psychological and Personality Science, 1-11*. doi: 10.1177/1948550616681055. Manuscript accepted for publication.
- Gino, F., Ayal, S., & Ariely, D. (2013). Self-serving altruism? The lure of unethical actions that benefit others. *Journal of economic behavior & organization, 93*, 285-292.
- Halevy, R., Shalvi, S., & Verschuere, B. (2014). Being honest about dishonesty: Correlating self-reports and actual lying. *Human Communication Research, 40*, 54-72.
- Knobe, J., & Nichols, S. (Eds.). (2013). *Experimental philosophy (Vol. 2)*. Oxford University Press.
- Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkissian, H., & Sommers, T. (2012). Experimental philosophy. *Annual Review of Psychology, 63*, 81-99.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*.
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate behavioral research, 39*, 329-358.

- Levine, T. R., & McCornack, S. A. (1992). Linking love and lies: A formal test of the McCornack and Parks model of deception detection. *Journal of Social and Personal Relationships, 9*, 143-154.
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49*, 433-442.
- McCornack, S. A., & Parks, M. R. (1986). Deception detection and relationship development: The other side of trust. *Annals of the International Communication Association, 9*, 377-389.
- Muehlenhard, C. L., & Shippee, S. K. (2010). Men's and women's reports of pretending orgasm. *Journal of Sex Research, 47*, 552-567.
- Nichols, S. (2011). Experimental philosophy and the problem of free will. *Science, 331*, 1401-1403.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. *The role of constructs in psychological and educational measurement, 4969*.
- Peer, E., Acquisti, A., & Shalvi, S. (2014). "I cheated, but only a little": Partial confessions to unethical behavior. *Journal of Personality and Social Psychology, 106*(2), 202.
- Shalvi, S., Dana, J., Handgraaf, M. J., & De Dreu, C. K. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes, 115*, 181-190.
- Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications doing wrong and feeling moral. *Current Directions in Psychological Science, 24*(2), 125-130.
- Turri, A., & Turri, J. (2015). The truth about lying. *Cognition, 138*, 161-168.

- Uziel, L. (2010). Rethinking social desirability scales from impression management to interpersonally oriented self-control. *Perspectives on Psychological Science*, 5, 243-262.
- Uziel, L. (2014). Impression Management (“Lie”) Scales Are Associated With Interpersonally Oriented Self-Control, Not Other-Deception. *Journal of personality*, 82(3), 200-212.
- Zettler, I., Hilbig, B. E., Moshagen, M., & de Vries, R. E. (2015). Dishonest responding or true virtue? A behavioral test of impression management. *Personality and Individual Differences*, 81, 107-111.

Lie-scales & Honesty: Supplementary materials

Contents

Power analyses	2
Open Science.....	2
Procedure and data disclosures.....	2
Data collection	2
Conditions reporting	2
Data exclusions	2
Variables reporting	2
Pre-registration	3
Hypotheses	3
Methods.....	3
Design.....	3
Planned sample.....	3
Exclusion criteria	4
Materials used in the experiment.....	6
Intro.....	6
Quiz and familiarity questions	6
Part 1.....	6
Part 2.....	7
Part 3.....	8
Collected and not analyzed.....	8

Power analyses

See pre-registration section.

Open Science

Data and code is on the Open Science Framework. Files are available for review using the following link: <https://osf.io/7gnfe/>

Procedure and data disclosures

Data collection

Data collection was completed before conducting an analysis of the data.

Conditions reporting

All collected conditions are reported.

Data exclusions

There were no data exclusions beyond the exclusions reported in the manuscript.

The full sample reports are:

Table 2

Part 1: Means, standard deviations, and summary statistics

	M X	SD X	MY	SD Y	<i>r</i>	<i>t</i>	<i>p</i>	<i>d_z</i>	<i>d_{rm}</i>
An honest person	4.12	1.88	5.21	1.76	-0.29	6.49	.000	0.37	0.60
Answered items truthfully	3.43	2.08	6.05	1.25	-0.27	16.87	.000	0.97	1.55
Believes answers reflect truth	4.23	1.93	6.08	1.22	-0.18	13.06	.000	0.75	1.16
Cares what others think	6.16	1.18	3.16	1.68	-0.05	24.80	.000	1.43	2.06
Understood the questions	6.07	1.20	6.24	1.18	0.42	2.31	.021	0.13	0.14
Read questions carefully	5.68	1.44	6.08	1.14	0.31	4.53	.000	0.26	0.31
Answered randomly	2.02	1.43	1.95	1.47	0.40	0.76	.445	0.04	0.05
How most people would answer	4.44	1.86	3.84	1.82	-0.46	3.32	.001	0.19	0.33

Note. N = 302 pairs in a with-subject design; X = Participant X with highest lie-scale score, Y = Participant Y with lowest lie-scale score; SD = Standard deviation, *r* = correlation between X and Y, *d* = pairs Cohen's *d* effect-size.

When comparing the two profiles in Part 2, participants rated the low lie-scale score as reflecting a more honest person (215/302, 71%, Binomial $p < .001$), as answering items more truthfully (235/302, 78%, binomial $p < .001$), and as caring less about what others think of them (48/302, 16%, binomial $p < .001$). The two profiles were rated as likely to represent the average person (52% versus 48%, binomial $p = .605$).

Finally, participant showed higher agreement with the interpretation that high lie scale scores represent less honesty (N = 151, $M = 5.60$, $SD = 1.60$ versus N = 151, $M = 3.25$, $SD = 2.07$, $t(282.30) = 10.99$, $p < .001$, $d = 1.27$).

Variables reporting

All variables collected for this study are reported and included in the provided data.

Pre-registration

We pre-registered the experiment on the Open Science Framework and data collection was launched later that day.

Hypotheses

Description of essential elements

1. Laypersons interpret a profile with high scores on the lie-scale as indicating higher dishonesty than a profile with low scores on the lie-scale.
2. Laypersons interpret a person with high scores on the lie-scale as indicating a generally more honest person than a person with low scores on the lie-scale.

Methods

Design

1. Part 1: Rating high-scorers and low-scorers. Within-subject design, all participants rating both scorers.
 - a. IV: low-scorer versus high-scorers on the lie-scale
 - b. DVs:
 - i. honesty in answering items (“has answered all items truthfully”)
 - ii. generally an honest person (“is generally an honest person”)
 - iii. social desirability (“cares what others think of”)
 - c. Controls (to rule out alternatives)
 - i. understanding the survey
 - ii. read carefully the survey
 - iii. intentionality (“believes that answers reflect the truth”)
 - iv. answered randomly
 - d. Additional analyses (exploratory)
 - i. Reflects the average person
2. Part 2: Comparing high-scorers and low-scorers
 - a. IV: no manipulation, comparing to a 50-50 random split
 - b. DVs:
 - i. honesty in answering items (“has answered all items truthfully”)
 - ii. generally an honest person (“is generally an honest person”)
 - iii. social desirability (“cares what others think of”)
 - c. Additional analyses (exploratory)
 - i. Reflects the average person
3. Part 3: Is high-score on lie scale higher or lower dishonesty
 - a. IV: between-subject manipulation of order framing (less truthful versus more truthful) in order to control for confirmation bias.
 - b. DV:
 - i. Agreement with interpretation

Planned sample

4. Amazon Mechanical Turk (MTurk) will be used to recruit 200 American participants to take part in the research to obtain data collection.
5. The sample size was determined by using the G-power analysis based on the calculation of minimum sample size for power .99 and alpha .05 for a weak-medium effect ($d = .35$; $g = .15$)

Comparing between the high-score versus low-score in a with-in design. N = 131:

t tests - Means: Difference between two dependent means (matched pairs)

Analysis: A priori: Compute required sample size
Input: Tail(s) = One
 Effect size dz = 0.35
 α err prob = 0.05
 Power (1- β err prob) = 0.99
Output: Noncentrality parameter δ = 4.0059331
 Critical t = 1.6566594
 Df = 130
 Total sample size = 131
 Actual power = 0.9903610

Comparing between a choice of whether the high-scorer or the low-scorer is more honest. Comparing the choice to a random 50-50 split, N = 173:

Exact - Proportion: Difference from constant (binomial test, one sample case)

Analysis: A priori: Compute required sample size
Input: Tail(s) = One
 Effect size g = 0.15
 α err prob = 0.05
 Power (1- β err prob) = 0.99
 Constant proportion = 0.5
Output: Lower critical N = 98.0000000
 Upper critical N = 98.0000000
 Total sample size = 173
 Actual power = 0.9906740
 Actual α = 0.0470574

Exclusion criteria

6. The following will be excluded from the sample:
 - a. Participants who are familiar with this scale.
 - b. Participants who have taken the scale before.
7. We will also determine exclusions based on:
 - a. Participants indicating a low proficiency of English (self-report < 5, scale of 1-7)
 - b. Participants who self-report not being serious about filling in the survey (self-report < 4, scale of 1-5).

In any case, we will report exclusions in detail with results for full sample and results following exclusions (in either the manuscript or the supplementary).

Procedure

A Qualtrics survey will be used for this study.

See attached exported Qualtrics survey for full procedure and materials.

C. Analysis plan

Confirmatory analyses

1. Part 1:
 - Dependent samples t-test comparing high-scorers to low-scorers on all DVs.
 - Ruling out controls by showing no significant differences and with equivalence test, not different from NULL, weaker than a weak effect ($d = .2$)
2. Part 2:
 - Binomial Z comparison to a 50-50 split random answer on all DVs.

3. Part 3:

- Each of the conditions significantly different from mid cut off point (4). T-test for single sample.
- Reversing one of the conditions, combining and testing difference from cut-off point (4).

Recommended elements

Specify contingencies and assumptions, such as:

4. No missing data (Qualtrics verifications)
5. No reliability criteria for item inclusion in scale.
6. No data transformations.

Answer the following final questions:

Has data collection begun for this project?

- No, data collection has not begun

If data collection has begun, have you looked at the data?

- No

Materials used in the experiment

Intro

In this study, we are asking for your help with a scale we used on other participants in a different study. We are interested in your interpretations of people's answers to this scale. We presented participants with a scale asking participants the following questions: If you say you will do something, do you always keep your promise no matter how inconvenient it might be? (yes/no) Are all your habits good and desirable ones? (yes/no) Do you always practice what you preach? (yes/no) Were you ever greedy by helping yourself to more than your share of anything? (yes/no) Have you ever blamed someone for doing something you knew was really your fault? (yes/no) Have you ever taken anything (even a pin or button) that belonged to someone else? (yes/no) Have you ever broken or lost something belonging to someone else? (yes/no) Have you ever said anything bad or nasty about anyone? (yes/no) As a child were you ever cheeky to your parents? (yes/no) Have you ever cheated at a game? (yes/no) Have you ever taken advantage of someone? (yes/no) Do you sometimes put off until tomorrow what you ought to do today? (yes/no)

Quiz and familiarity questions

How many items are on this scale? (open text question with validation 12)

What are the possible answers for each of the scale items? (validation for yes/no)

1. Yes or No (1)
2. Likert scale from 1 to 7 (2)
3. Free text answers (3)

Have you ever answered this exact scale before?

1. Yes (23)
2. No (24)

Have you ever seen or heard of this exact scale before?

1. Yes (23)
2. No (24)

Part 1

Within-subject design with randomized order of display.

Intro:

We would like your opinion regarding what answers on this scale reflect. We will now present you with two sample participants - Participant X and Participant Y - and their answers, and you will be asked to evaluate the participants based on their answers to this scale.

High lie-scale score

Participant X responded to the scale in the following way:

1. If you say you will do something, do you always keep your promise no matter how inconvenient it might be? **yes**
2. Are all your habits good and desirable ones? **yes**

3. Do you always practice what you preach? **yes**
4. Were you ever greedy by helping yourself to more than your share of anything? **no**
5. Have you ever blamed someone for doing something you knew was really your fault? **no**
6. Have you ever taken anything (even a pin or button) that belonged to someone else? **no**
7. Have you ever broken or lost something belonging to someone else? **no**
8. Have you ever said anything bad or nasty about anyone? **no**
9. As a child were you ever cheeky to your parents? **no**
10. Have you ever cheated at a game? **no**
11. Have you ever taken advantage of someone? **no**
12. Do you sometimes put off until tomorrow what you ought to do today? **no**

Low lie-scale score

Participant Y responded to the scale in the following way:

1. If you say you will do something, do you always keep your promise no matter how inconvenient it might be? **no**
2. Are all your habits good and desirable ones? **no**
3. Do you always practice what you preach? **no**
4. Were you ever greedy by helping yourself to more than your share of anything? **yes**
5. Have you ever blamed someone for doing something you knew was really your fault? **yes**
6. Have you ever taken anything (even a pin or button) that belonged to someone else? **yes**
7. Have you ever broken or lost something belonging to someone else? **yes**
8. Have you ever said anything bad or nasty about anyone? **yes**
9. As a child were you ever cheeky to your parents? **yes**
10. Have you ever cheated at a game? **yes**
11. Have you ever taken advantage of someone? **yes**
12. Do you sometimes put off until tomorrow what you ought to do today? **yes**

Questions

We would like to ask for your opinion regarding this participant.

Please indicate how likely it is that the participant:

1. Is generally an honest person
2. Has answered all items in this scale truthfully
3. Cares about what others would think of him/her
4. Has understood the questions
5. Has read the questions carefully
6. Believes that his/her answers reflect the truth
7. Answered in a way most people would answer this scale
8. Answered this scale randomly

(1 – Extremely unlikely; 7 – Extremely likely)

Part 2

Now, we would like you to compare between the two: Participant X Participant X responded to the scale in the following way:

[The two conditions from part 1 were presented next to each other in a table]

Which of the two participants:

1. Is generally a more honest person
2. Has answered items on this scale more truthfully
3. Cares more about what others would think of him/her
4. Better represents the average person's answer

Part 3

Between-subject design

High lie-score = less truthful

[Were again presented with a table comparing the two participants from Part 1, as in Part 2]

This scale was designed to measure honesty. If you compare the two participants above - this suggested scale considers Participant X who scored high on this scale to be less truthful than Participant Y who scored low on the scale. Do you agree with this interpretation?

Strongly disagree (1) to Strongly agree (7)

High lie-score = more truthful

[Were again presented with a table comparing the two participants from Part 1, as in Part 2]

This scale was designed to measure honesty. If you compare the two participants above - this suggested scale considers Participant X who scored high on this scale to be more truthful than Participant Y who scored low on the scale. Do you agree with this interpretation?

Strongly disagree (1) to Strongly agree (7)

Collected and not analyzed

We also asked an open question – “To the best of your understanding, guess if you can't tell, what do you think this scale measures? (briefly, in one sentence, write only one thing)”. This question was meant to assess possible issues if findings did not meet pre-registration. We have not analyzed responses to this question.