

Frankly, we do give a damn:  
The relationship between profanity and honesty

Gilad Feldman

Department of Work and Social Psychology

Maastricht University

Maastricht, 6200MD, the Netherlands

[gilad.feldman@maastrichtuniversity.nl](mailto:gilad.feldman@maastrichtuniversity.nl)

Huiwen Lian

Department of Management

Hong Kong University of Science and Technology

Clearwater Bay, Kowloon, Hong Kong

[huiwen@ust.hk](mailto:huiwen@ust.hk)

\*Michal Kosinski

Graduate School of Business,

Stanford University, California

[michalk@stanford.edu](mailto:michalk@stanford.edu)

\*David Stillwell

Judge School of Business,

University of Cambridge, United Kingdom

[d.stillwell@jbs.cam.ac.uk](mailto:d.stillwell@jbs.cam.ac.uk)

\*Third author equal contribution

Word count:

Manuscript – 4780; Abstract – 150

Accepted for publication at Social Psychological and Personality Science (Oct 23, 2016)

### **Abstract**

There are two conflicting perspectives regarding the relationship between profanity and dishonesty. These two forms of norm-violating behavior share common causes, and are often considered to be positively related. On the other hand, however, profanity is often used to express one's genuine feelings, and could therefore be negatively related to dishonesty. In three studies, we explored the relationship between profanity and honesty. We examined profanity and honesty first with profanity behavior and lying on a scale in the lab (Study 1; N = 276), then with a linguistic analysis of real-life social interactions on Facebook (Study 2; N = 73,789), and finally with profanity and integrity indexes for the aggregate level of U.S. states (Study 3; N = 50 states). We found a consistent positive relationship between profanity and honesty; profanity was associated with less lying and deception at the individual level, and with higher integrity at the society level.

*Keywords:* profanity; honesty; cursing; integrity

Frankly, we do give a damn:

The relationship between profanity and honesty

“Frankly my dear, I don’t give a damn.” (*Gone with the Wind*, 1939)

## **Introduction**

Profane as it is, this memorable line by the character Rhett Butler in the film *Gone with the Wind* profoundly conveys Butler’s honest thoughts and feelings. However, it was the use of this profane word that led to a \$5,000 fine against the film’s production for violating the Motion Picture Production Code. This example reveals the conflicting attitudes that most societies hold toward profanity, reflected in a heated debate taking place in online forums and media in recent years—with passionate views on both sides. For example, the website *debate.org*, which conducts online polls and elicits general public opinions on popular online debates, has many comments on the issue, with a 50-50 tie between the two views (“Are people who swear more honest?”, 2015). This public debate reflects an interesting question and mirrors the academic discussion regarding the nature of profanity. On the one hand, profane individuals are widely perceived as violating moral and social codes, and thus deemed untrustworthy and potentially antisocial and dishonest (Jay, 2009). On the other hand, profane language is considered as more authentic and unfiltered, thus making its users appear more honest and genuine (Jay, 2000). These opposing views on profanity raise the question of whether profane individuals tend to be more or less dishonest.

## **Profanity**

Profanity refers to obscene language including taboo and swear words, which in regular social settings are considered inappropriate and in some situations unacceptable. It often includes sexual references, blasphemy, objects eliciting disgust, ethnic-racial-gender slurs, vulgar terms, or offensive slang (Mabry, 1974). The interest in understanding the psychological roots of the use of profanity dates back to as far as the early 20<sup>th</sup> century (Patrick, 1901), yet the literature in this domain is scattered across different scientific fields with only recent attempts to connect the findings into a unified framework (Jay, 2009).

The reasons for using profanity depend on the person and the situation, yet profanity is commonly related to the expression of emotions such as anger, frustration, or surprise (Jay & Janschewitz, 2008). The spontaneous use of profanity is usually the unfiltered genuine expression of emotions, with the most extreme type being the bursts of profanity (i.e. *coprolalia*) accompanying the Tourette syndrome (Cavanna & Rickards, 2013). The more controllable use of profanity often helps to convey world views or internal states; or is used to insult an object, a view, or a person (Jay, 2009). Speech involving profane words has a stronger impact on people than regular speech, and has been shown to be processed on a deeper level in people's minds (Jay, Caldwell-Harris, & King, 2008).

The context is important for understanding profanity. Profanity can sometimes be interpreted as antisocial, harmful, and abusive—if, for example, intended to harm or convey aggression and hostile emotions (Stone, McMillan, & Hazelton, 2015). It also violates the moral foundations of purity (Sylwester & Purver, 2015) and the common norm for speech, suggestive of the potential to engage in other antisocial behaviors that violate norms and morality. However, profanity may also be seen as a positive if it does not inflict harm but acts as a reliever of stress

or pain in a cathartic effect (Robbins et al., 2011; Vingerhoets, Bylsma, & de Vlam, 2013).

Profane language can serve as a substitute for potentially more harmful forms of violence (Jay, 2009), and can alert others to one's own emotional state or the issues that one cares about deeply (Jay, 2009). Profanity is also used to entertain, attract (Kaye & Sapolsky, 2009), and influence audiences (Scherer & Sagarin, 2006), as illustrated by the frequent use of profane language in comedy, mass media, and advertising (Sapolsky & Kaye, 2005). Profanity has even been used by presidential candidates in American elections (Slatcher, Chung, Pennebaker, & Stone, 2007), as recently illustrated by Donald Trump, who has been both hailed for authenticity and criticized for moral bankruptcy (Sopan, 2015).

### **Dishonesty**

In its most basic form, dishonesty involves the conscious attempt by a person to convince others of a false reality (Abe, 2011). In this work, we operationalize dishonesty as a generalized personal inclination to obscure the truth in natural, everyday life situations. The most common type of such dishonesty is represented by "white lies" or "social lies" that people tell themselves or others in order to appear more desirable or positive (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Granhag & Vrij, 2005). While most people claim to be honest most of the time (Aquino & Reed, 2002; Halevy, Shalvi, & Verschuere, 2013), research suggests that minor cases of dishonesty are quite common (DePaulo & Kashy, 1998; Hofmann, Wisneski, Brandt, & Skitka, 2014; Serota, Levine, & Boster, 2010), especially when people believe that dishonesty is harmless or justifiable (Fang & Casadevall, 2013), or that they can avoid any penalties (Gino, Ayal, & Ariely, 2009). In other words, people tend to rationalize their own dishonesty (Ayal & Gino, 2012) and perceive it as less severe (Peer, Acquisti, & Shalvi, 2014) or non-existent (Mazar, Amir, & Ariely, 2008).

### **The relationship between profanity and dishonesty**

There are two opposing perspectives on the relationship between profanity and dishonesty. As dishonesty and profanity are both considered deviant (Bennett & Robinson, 2000) and immoral (Buchtel et al., 2015), they are generally perceived as a reflection of a disregard for societal normative expectations (Kaplan, 1975), low moral standards, lack of self-control, or negative emotions (Jay, 1992, 2000). In this regard, profanity appears to be positively related to dishonesty, explaining why people who swear are perceived as untrustworthy (Jay, 1992) and why swear words are often associated with deceit (Rassin & Van Der Heijden, 2005). Previous work has also linked the use of swear words to the dark triad personality traits—namely narcissism, Machiavellianism, and psychopathy—all indicative of social deviance and a higher propensity for dishonesty (Holtzman, Vazire, & Mehl, 2010; Sumner, Byers, Boochever, & Park, 2012). Swearing has also been shown to hold a negative relationship with the personality traits of conscientiousness and agreeableness, which are considered the more socially aware and moral aspects of personality (Kalshoven, Den Hartog, & De Hoogh, 2011; Mehl, Gosling, & Pennebaker, 2006; Walumbwa & Schaubroeck, 2009).

On the other hand, profanity can be positively associated with honesty. It is often used to express one's unfiltered feelings (e.g. anger, frustration) and sincerity. Innocent suspects, for example, are more likely to use swear words than guilty suspects when denying accusations (Inbau, Reid, Buckley, & Jayne, 2011). Accordingly, people perceive testimonies containing swear words as more credible (Rassin & Van Der Heijden, 2005).

### **The present investigation**

This work explores the relationship between profanity and honesty to address the paradoxical perspectives in the existing literature. Study 1 examined the relationship between

profanity use and honesty on a lie scale. Study 2 examined behavior in real-life naturalistic setting by analyzing behavior on Facebook: looking at the relationship between users' profanity rate and honesty in their online status updates, as indicated by a linguistic detection of deception. Study 3 extended to society level by exploring the relationship between state-level profanity rates and state-level integrity. The supplementary materials include power analyses, procedures, and stimuli used in the three studies, and data and code were made available on the Open Science Framework (<https://osf.io/z9jbm/>).

### **Study 1 – Honesty on a lie scale**

We began our investigation with a test for the relationship between profanity and honesty, captured by a widely used lie scale.

#### **Method**

##### **Participants and procedure**

A total of 307 participants were recruited online using Amazon Mechanical Turk. Of the sample, 31 participants failed attention checks (10%) and were excluded from the analysis, leaving a sample of 276 ( $M_{age} = 40.71$ ,  $SD_{age} = 12.75$ ; 171 females). The exclusion of participants had no significant impact on the reported effect sizes or p-values below. Participants self-reported profanity use in everyday life: given the opportunity to use profanity, rated reasons for the use of profanity, and answered a lie scale.

##### **Measures**

***Profanity use behavioral measure.*** In two items, participants were asked to list their most commonly used and favorite profanity words: “Please list the curse words you [1 – use; 2 – like] the most (feel free, don’t hold back).” By giving participants an opportunity to curse freely, we expected that the daily usage and enjoyment of profanity would be reflected in the total

number of curse words written. Participants' written profanity was counted and coded by the first author and a coder unrelated to the project, who was unaware of the study hypotheses and data structure. The interrater reliability was .91 (95% CI [.87, .94]) for most commonly used curse words and .93 (95% CI [.91, .97]) for favorite curse words, indicating a very high level of agreement.

***Profanity self-reported use.*** To supplement the behavioral measures, we also added self-reported use of profanity. Participants self-reported their everyday use of profanity (Rassin & Muris, 2005) using three items: "How often do you curse (swear / use bad language)" (1) "verbally in person (face to face)," (2) "in private (no one around)," and (3) "in writing (e.g. texting/messaging/posting online/emailing)" (1 = *Never*; 2 = *Once a year or less*; 3 = *Several times a year*; 4 = *Once a month*; 5 = *2–3 times a month*; 6 = *Once a week*; 7 = *2–3 times a week*; 8 = *4–6 times a week*; 9 = *Daily*; 10 = *A few times a day*) ( $\alpha = .84$ ).

***Reasons for profanity use.*** Following Rassin and Muris (2005), we also asked participants to rate reasons for their use of profanity (0 = *Never a reason for me to swear*; 5 = *Very often a reason for me to swear*), and asked questions regarding the general perceived reasons for using profanity (0 = *Not at all*; 5 = *To a very large extent*) (see supplementary materials).

***Honesty.*** Honesty was measured using the lie subscale of the Eysenck Personality Questionnaire Revised short scale (Eysenck, Eysenck, & Barrett, 1985). The lie subscale is one of the most common measures for assessing individual differences in lying for socially desirable responding (Paulhus, 1991). The lie scale includes 12 items, such as: "If you say you will do something, do you always keep your promise no matter how inconvenient it might be?" and "Are all your habits good and desirable ones?" (dichotomous Yes/No scale). In these examples,



positive answers are considered unrealistic and therefore most likely a lie ( $\alpha = .79$ ). The lie scale was reversed for the honesty measure.

## Results

The means, standard deviations, and correlations for the honesty and profanity measures are detailed in Table 1. Honesty was positively correlated with all profanity measures, meaning that participants lied less on the lie scale if they wrote down a higher number of frequently used ( $r = .20, p = .001$ ; CI [.08, .31]) and liked curse words ( $r = .13, p = .032$ ; CI [.01, .24]), or self-reported higher profanity use in their everyday lives ( $r = .34, p < .001$ ; CI [.23, .44]), even when controlling for age and gender (behavioral 1: partial  $r = .20, p = .001$ ; CI [.08, .31]; behavior 2: partial  $r = .12, p = .049$ ; CI [.001, .24]; self-report: partial  $r = .32, p < .001$ ; CI [.21, .42]).

Table 1

*Study 1: Means, standard deviations, and correlations for variables.*

	Mean	SD	1	2	3	4	5
1 – Honesty	7.63	3.00	(.79)				
2 – Profanity self-report	6.51	2.56	.34***	(.84)			
3 – Profanity behavioral 1	4.09	2.61	.20**	.46***	(-)		
4 – Profanity behavioral 2	1.60	1.62	.13*	.41***	.45***	(-)	
5 – Age	40.71	12.75	-.13*	-.34***	-.05	-.08	(-)
6 – Gender	1.62	.49	-.06	-.03	-.07	-.04	.08

*Note:* N = 276. Gender coding: 1 = male, 2 = female; \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ . Scale alpha coefficients are on the diagonal. Profanity behavioral 1 = number of most frequently used curse words written; Profanity behavioral 2 = number of most liked curse words written.

We asked participants to rate their reasons for use of profanity. The reasons that received the highest ratings were the expression of negative emotions ( $M = 4.09, SD = 1.33$ ), habit ( $M = 3.08, SD = 1.82$ ), and an expression of true self ( $M = 2.17, SD = 1.73$ ). Participants also indicated that in their personal experience, profanity was used for being more honest about their feelings ( $M = 2.69, SD = 1.72$ ) and dealing with their negative emotions ( $M = 2.57, SD = 1.64$ ). Profanity received a lower rating as a tool for insulting others ( $M = 1.41, SD = 1.53$ ), as well as for being perceived as intimidating or insulting ( $M = 1.12, SD = 1.36$ ). This supports the view that people regard profanity more as a tool for the expression of their genuine emotions, rather than being antisocial and harmful.

### **Study 2 – Naturalistic deceptive behavior on Facebook**

Study 1 provided initial support for a positive relationship between profanity use and honesty, with the limitations of lab settings. Study 2 was constructed to extend Study 1 to a naturalistic setting—using a larger sample, more accurate measures of real-life use of profanity, and a different honesty measure.

With a stellar growth, Facebook has become the world's most dominant social network and is strongly embedded in its users' overall social lives (Manago, Taylor, & Greenfield, 2012; Wilson, Gosling, & Graham, 2012). Online social networking sites such as Facebook now serve as an extension of real-life social context, allowing individuals to express their actual selves (Back et al., 2010). Facebook profiles have been found to provide fairly accurate portrayals of their users' personalities and behaviors (Kosinski, Stillwell, & Graepel, 2013; Schwartz et al., 2013; Wang, Kosinski, Stillwell, & Rust, 2012), including socially undesirable aspects (Garcia & Sikström, 2014) such as self-promotion (Waggoner, Smith, & Collins, 2009; Weisbuch, Ivcevic,

& Ambady, 2009), narcissism (Buffardi & Campbell, 2008), and low self-esteem (Zywica & Danowski, 2008).

In this work, we detected dishonesty by analyzing Facebook users' status updates that were used to broadcast messages to their online social network. Using language to tap into people's psyches dates back to Freud (1901), who analyzed patients' slips of the tongue, and Lacan (1968), who argued that the unconscious manifests itself in language use. A growing body of literature has since demonstrated that the language that people use in their daily lives can reveal hidden aspects of their personalities, cognitions, and behaviors (Pennebaker, Mehl, & Niederhoffer, 2003). The linguistic approach is especially useful in the case of dishonesty, which—though prevalent—is frowned upon when detected, and therefore leads those who are acting dishonestly to try to hide it from others (Hancock, 2009; Toma, Hancock, & Ellison, 2008). In the case of Facebook, the dishonesty we refer to is not necessarily blunt deception aimed at exploiting or harming others, but rather a mild distortion of the truth intended to construe a more socially desirable appearance (Whitty, 2002; Whitty & Gavin, 2001).

## **Method**

### **Participants and procedure**

A total of 153,716 participants were recruited using the myPersonality Facebook application (Kosinski, Matz, Gosling, Popov, & Stillwell, 2015). Participants voluntarily chose to use this application and provided opt-in consent to record their Facebook profiles, including demographic data and their status updates (more information about the myPersonality Facebook application is available at: <http://mypersonality.org>). This analysis is limited to users who used the English version of Facebook, had more than 50 Facebook status updates, and had more than 30 friends (an indication of being an active Facebook user). The final sample included 73,789

participants (62.0% female,  $M_{\text{age}} = 25.34$ ,  $M_{\text{network-size}} = 272.37$ ;  $M_{\text{statusupdates}} = 201.28$ ,  $SD_{\text{statusupdates}} = 167.33$ ;  $M_{\text{words}} = 3181.82$ ,  $SD_{\text{words}} = 3014.44$ ).

### Measures

We used Linguistic Inquiry and Word Count (LIWC; Tausczik & Pennebaker, 2010) in order to analyze participants' status updates. The analysis was conducted by aggregating all the status updates of every participant into a single file, and executing a LIWC analysis on each user's combined status updates. The LIWC software reported the percentages of the words in each LIWC category out of all of the words used in the combined status updates, as follows:

$$\text{LIWC category rate user } X = \frac{\text{User } X \text{ word count for LIWC category in all status updates}}{\text{User } X \text{ word count in all status updates}}$$

**Honesty.** The honesty of the status updates written by the participants was assessed following the approach introduced by Newman and colleagues (2003) using LIWC. Their analyses showed that liars use fewer first-person pronouns (e.g. I, me), fewer third-person pronouns (e.g. she, their), fewer exclusive words (e.g. but, exclude), more motion verbs (e.g. arrive, go), and more negative words (e.g. worried, fearful) (Newman, Pennebaker, Berry, & Richards, 2003). The explanation was that dishonest people subconsciously try to (1) dissociate themselves from the lie and therefore refrain from referring to themselves; (2) prefer concrete over abstract language when referring to others (using someone's name instead of "he" or "she"); (3) are likely to feel discomfort by lying and therefore express more negative feelings; and (4) require more mental resources to obscure the lie and therefore end up using less cognitively demanding language, which is characterized by a lower frequency of exclusive words and a higher frequency of motion verbs. Equation and usage rates in this study are summarized in Table 2.

Table 2

*Study 2: Word analysis of LIWC categories and keywords.*

LIWC dimensions	Sample LIWC keywords	Honesty coefficients $\beta$ s	Percentage (M)	Percentage (SD)
1 <sup>st</sup> -person pronouns	I, me, mine	0.260	4.21%	1.71%
3 <sup>rd</sup> -person pronouns	She, her, him, they, their	0.250	0.84%	.33%
Exclusive words	But, without, exclude	0.419	1.78%	.63%
Motion verbs	Arrive, car, go	-0.259	1.57%	.53%
Anxiety words	Worried, fearful, nervous	-0.217	0.21%	.14%

Newman et al. (2003) achieved up to 67% accuracy when detecting lies, which was significantly higher than the 52% near-chance accuracy achieved by human judges. Their approach has been successfully applied to behavioral data (Slatcher et al., 2007) and to Facebook status updates (Feldman, Chao, Farh, & Bardi, 2015). Other studies have since found support for these LIWC dimensions as being indicative of lying and dishonesty (Bond & Lee, 2005; Hancock, Curry, Goorha, & Woodworth, 2007; see meta-analyses by DePaulo et al., 2003 and Hauch, Masip, Blandón-Gitlin, & Sporer, 2012).

To calculate the honesty score, we first computed LIWC scores to obtain participants' use rate of first-person pronouns, third-person pronouns, exclusive words, motion verbs, and anxiety words, and then applied average regression coefficients from Newman et al. (2003). Here we note that we focused on anxiety words rather than general negative words (which include anxiety, anger, and sadness) due to two considerations. First, it has been suggested that anxiety words may be more predictive of honesty than overall negative emotions (Newman et al., 2003). Second, measuring honesty using negative emotions with anger words may bias the profanity-honesty correlations because anger has been shown to have a strong positive relation with profanity. Holtzman et al. (2010) reported a correlation of .96 between anger and profanity, and

Yarkoni (2010) found swearing to be strongly associated with anger but not with anxiety, which is not surprising given the conclusion by Jay and Janschewitz (2008) that profanity is mostly used to express anger<sup>1</sup>.

**Profanity.** We used the LIWC dictionary of swear words (e.g. damn, piss, fuck) to obtain the participants' use rate of profanity. This approach was previously used to analyze swearing patterns in social contexts (e.g. Holtgraves, 2011; Mehl & Pennebaker, 2003). Profanity use rates were calculated per each participant using LIWC, with rates indicating the percentage of swear words used in all status updates by the participant overall. Profanity use rates were then log-transformed to normalize distribution ( $\ln(\text{profanity}+1)$ ).

## Results

The descriptive statistics and zero-order correlations of all variables are provided in Table 3. The mean of profanity use was 0.37% ( $SD = 0.43\%$ ; 7,969 [10.8%] used no profanity at all), which is in line with previous findings (Jay, 2009). Profanity and honesty were found to be significantly and positively correlated ( $N = 73,789$ ;  $r = .20$ ,  $p < .001$ ; 95% CI [.19, .21]; see Figure 1 for an aggregated plot), indicating that those who used more profanity were more honest in their Facebook status updates. Controlling for age, gender, and network size resulted in a slightly stronger effect (partial  $r = .22$ ,  $p < .001$ ; 95% CI [.21, .22]).

---

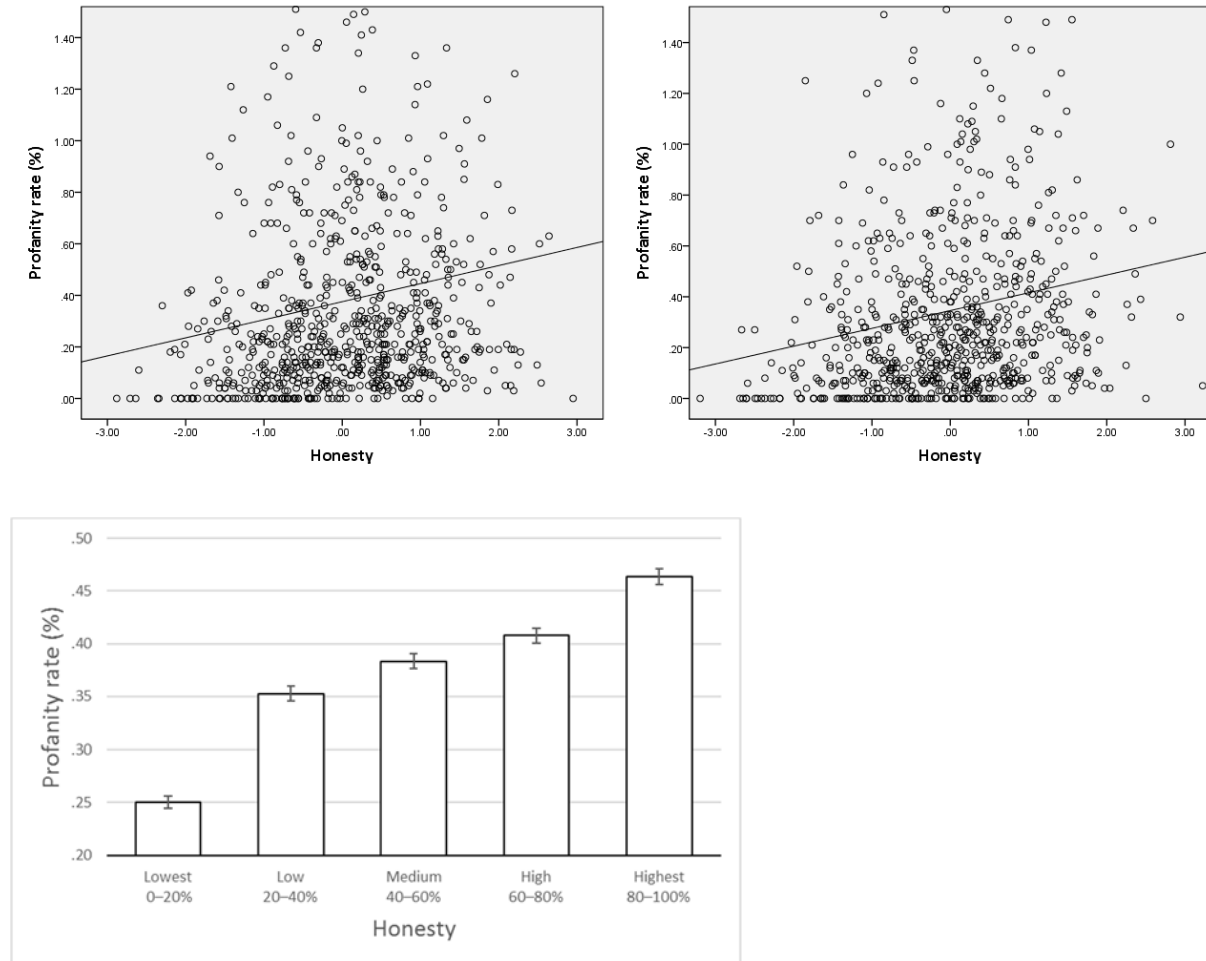
<sup>1</sup> The supplementary materials include further details and a report of the results using the original equation of negative emotions including anger ( $r = .02$ ,  $p < .001$ ; 95% CI [.01, .03]; with controls: partial  $r = .04$ ,  $p < .001$ ; 95% CI [.03, .05]).

Table 3

*Study 2: Descriptive statistics of honesty, profanity, and demographics.*

	Mean	SD	Skewness	Kurtosis	Honesty	Profanity	Age	Gender
Honesty (raw)	0 1.60	1 .60	.03	.02	(-)	.22		
Profanity (raw)	.28 .37	.26 .43	1.37 2.51	2.00 9.49	.20	(-)		
Age	25.34	8.78	1.90	3.96	-.05	-.18	(-)	
Gender	.62	.49	-.49	-1.76	.12	-.23	.08	(-)
Network size (raw)	5.30 272.37	.79 249.71	-.03 4.18	-.25 39.82	.18	-.09	-.13	.00(ns)

*Note:* Gender coding: 0 = male, 1 = female; (ns) indicates a non-significant correlation coefficient; remaining coefficients were significant at  $p < .001$  level; honesty was standardized, profanity and network size were log-transformed. Males used more profanity than females (diff = .12 [.12, .13],  $t(46884.67) = 59.26$ ,  $p < .001$ ,  $d = .47$ ), and were less honest (diff = -.14 [-.15, -.13],  $t = -31.69$ ,  $p < .001$ ,  $d = -.23$ ). Raw lines indicate statistics for variables before transformations or standardizing. Values above the diagonal are partial correlations controlling for age, gender, and network size.



*Figure 1.* Study 2: The relationship between profanity and honesty (model 2). The first two scatterplots are of two randomly chosen 1% subsets of the total population (plot1: N = 750; plot2: N = 721). The third graph is a plot of aggregated honesty groups, and average profanity was computed for five equal groups of participants based on their honesty. The honesty score was standardized to the mean of 0 and standard deviation of 1. Error bars indicate a 95% confidence interval. The profanity rate is in percentages (e.g., .25 is 0.25% use).

### Study 3 – State-level integrity

Studies 1 and 2 demonstrated that the use of profanity is a predictor of honesty at the individual level. Study 3 sought to extend these findings by taking a broader view and examining the possible implications that individual differences in use of profanity have for society (as suggested by Back & Vazire, 2015). If the use of profanity is indeed positively related to



honesty, then it can be argued that societies with higher profanity rates may be characterized by a higher appreciation for honesty and genuineness. Study 3 examined whether the state-level use of profanity is predictive of state-level integrity as reported by the State Integrity Index 2012.

### **Measures**

**State-level profanity.** State-level profanity scores were computed by averaging the profanity scores of the American participants in Study 2 (29,701 participants) across the states. The state profanity scores are detailed in Table 4.

**State-level integrity.** State-level integrity was obtained from the State Integrity Investigation 2012 (SSI2012), the year that the myPersonality data collection was concluded. Estimating state levels of integrity and corruption is a complicated and controversial issue. For example, corruption was sometimes measured with the number of corruption convictions per state, yet a higher conviction rate can be indicative of better policing and thus lower corruption. We therefore used an index of integrity that is less affected by possible conflicting interpretations of crime and conviction statistics: the SSI2012. The SSI2012 ranks the states on 14 broad integrity criteria, including stance on honesty and transparency, the presence of independent ethics commissions; and executive, legislative, and judicial accountability. State integrity scores are detailed in Table 4. More information about how the State Integrity scores were obtained can be found in the supplementary materials.

Table 4

*Study 3: State-level profanity and integrity rates.*

State	Profanity rate	Integrity	State	Profanity rate	Integrity	State	Profanity rate	Integrity
Alabama	34	72	Maine	33	56	Oregon	36	73
Alaska	42	68	Maryland	46	61	Pennsylvania	42	71
Arizona	41	68	Massachusetts	46	74	Rhode Island	44	74
Arkansas	29	68	Michigan	41	58	South Carolina	29	57
California	44	81	Minnesota	39	69	South Dakota	38	50
Colorado	39	67	Mississippi	33	79	Tennessee	32	76
Connecticut	52	86	Missouri	37	72	Texas	38	68
Delaware	51	70	Montana	35	68	Utah	26	65
Florida	41	71	Nebraska	42	80	Vermont	35	69
Georgia	36	49	Nevada	47	60	Virginia	40	55
Hawaii	45	74	New Hampshire	36	66	Washington	36	83
Idaho	31	61	New Jersey	50	87	West Virginia	34	68
Illinois	45	74	New Mexico	34	62	Wisconsin	39	70
Indiana	35	70	New York	46	65	Wyoming	34	52
Iowa	40	87	North Carolina	37	71			
Kansas	39	75	North Dakota	37	58			
Kentucky	37	71	Ohio	39	66			
Louisiana	35	72	Oklahoma	33	64			

*Note:* Integrity is the SSI2012 index. Profanity rates were aggregated to the state level from the Study 2 Facebook profanity rates for American participants.

## Results

A scatterplot of profanity and integrity rates for all states is provided in Figure 2. We found a positive relationship between profanity and integrity on a state level ( $N = 50$ ;  $r = .35$ ,  $p =$

.014; CI [.08, .57]). States with a higher profanity rate had a higher integrity score.<sup>2</sup> For example, two of the three states with the highest profanity rate, Connecticut and New Jersey, were also two of the three states with the highest integrity scores on the index.

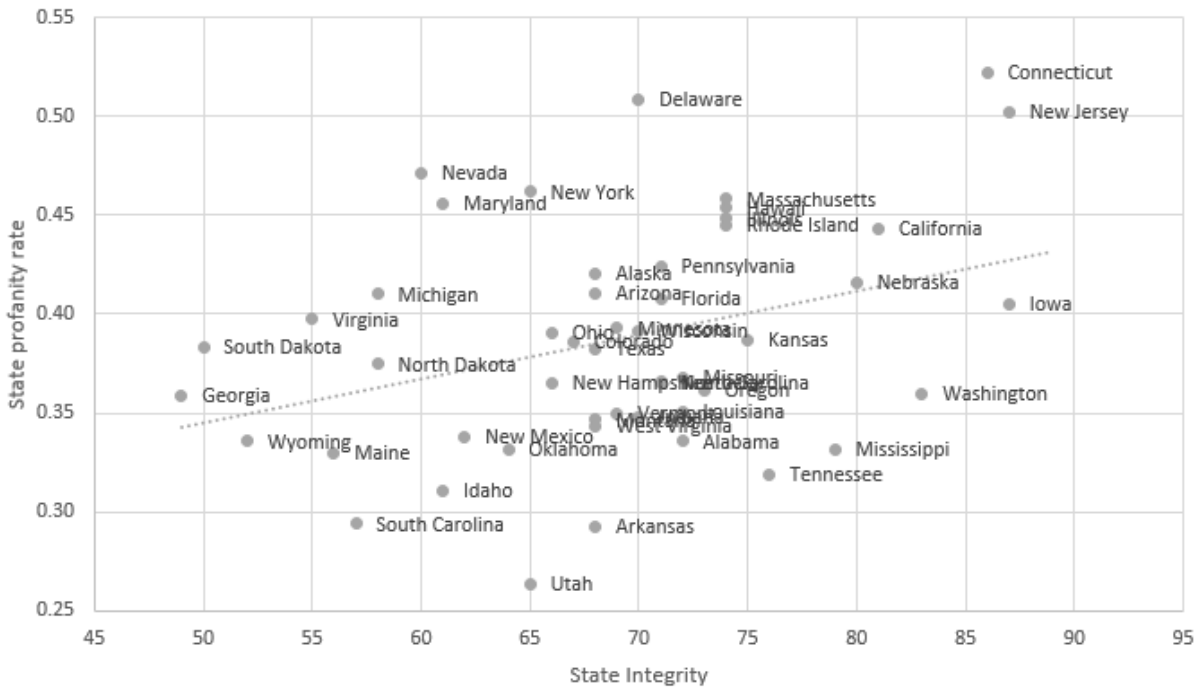


Figure 2. Study 3: scatterplot presenting integrity and profanity rates across 50 U.S. states.

We also conducted a spatial regression analysis to address possible spatial-dependence regional confounds (Ward & Gleditsch, 2008). We calculated spatial distance matrices (Merryman, 2008) for the distance between states’ centroids using the following formula for

<sup>2</sup> We noted problems in using crime and conviction rates in the methods, but ran several robustness checks. Higher state average of profanity use was negatively correlated with state rates of property crime ( $r = -.30, p = .032$ ), burglary ( $r = -.31, p = .029$ ), larceny theft ( $r = -.34, p = .015$ ), and rape ( $r = -.24, p = .093$ )—obtained from the FBI website.

Euclidean distance between state A and state B ( $y$  and  $x$  denote the y-coordinate and x-coordinate, respectively):

$$d(x_A, y_A; x_B, y_B) = \sqrt{(y_A - y_B)^2 + (x_A - x_B)^2}$$

We then inverted the distances ( $1/X$ ) to form a proximity measure, multiplied the proximity matrix by the state profanity column, and divided by the sum to create a measure of spatial lag – a spatial weighted profanity per each state (Webster & Duffy, 2016). Excluding Hawaii and Alaska for their geographical isolation, the spatial profanity measure had a correlation of  $r = .55$  with the state profanity measure ( $N = 48$ ;  $p < .001$ ; CI [.32, .72]; Moran  $I$  statistic = .15,  $p < .001$ ), indicative of spatial dependence. After controlling for the spatial profanity the partial correlation between profanity and integrity was  $r = .33$  ( $p = .025$ , CI [.05, .56]).

### General Discussion

We examined the relationship between the use of profanity and dishonesty, and showed that profanity is positively correlated with honesty at an individual level, and with integrity at a society level. Table 5 provides a summary of the results. Study 1 showed that participants with higher profanity use were more honest on a lie scale and in Study 2 profanity was associated with more honest language patterns in Facebook status updates. In Study 3, state-level profane language usage was positively related to integrity at the state level.

Table 5

*Summary of the results.*

#	Sample size	Sample type	Level of analysis	Profanity measure(s)	Honesty measure	Effect
1	276	American English native MTurk workers	Individual	1-2: Counts of written profanity / 3: Self-report	Eysenck Personality Questionnaire Revised short scale	.20/.13/.34
2	73,789	English version Facebook users	Individual	Rate of profanity in language used in status updates	Derivative of standard LIWC dimensions (Newman et al., 2003)	.20 (.22)
3	50 (48)	States in the U.S.	State	Average profanity in language used in status updates	State Integrity Investigation 2012 index	.35 (.33)

*Note.* Effects in parentheses are effects while controlling for other factors (Study2: age, gender, network-size; Study 3: Spatial distance).

### **Challenges in studying profanity and dishonesty in naturalistic settings**

The empirical investigation of the relationship between dishonesty and profanity poses a unique challenge. The behavioral ethics literature has been successful in devising ways to examine unethical behavior in the lab, yet observing dishonesty and unethical behavior in the field remains an ongoing challenge, and so far only a few studies were able to devise innovative methods to overcome that challenge (e.g. Hofmann et al., 2014; Piff, Stancato, Côté, Mendoza-Denton, & Keltner, 2012). The indirect linguistic approach for the detection of dishonesty with an analysis of spoken and written language patterns paves the way for more behavioral ethics research on actual dishonest behavior in the field.

Unlike behavioral ethics, the study of profanity is still very much in its infancy (Jay, 2009). Profanity is a much harder construct to measure and even more difficult to effectively

elicit or manipulate, whether it is in the lab or in the field. The relatively low use rates of profanity decrease even further when people know that they are observed or that their behavior is studied. Therefore, to be able to gain an understanding of profanity use, it is important that the behavior observed is genuine and in naturalistic settings. The current investigation has been able to address this challenge by applying a linguistic analysis approach to a unique large-scale naturalistic behavior dataset.

The linguistic approach to detecting dishonesty used in Study 2 has been used and verified in a number of previous studies (e.g. Feldman et al., 2015; Slatcher et al., 2007). In Study 2, the linguistic analysis showed that men tended to be more dishonest than women, which is in line with a large body of literature presenting similar findings (Childs, 2012; Dreber & Johannesson, 2008; Friesen & Gangadharan, 2012). Also, those with larger networks had a higher likelihood for dishonesty and a lower likelihood for profanity, which supports the notion of dishonesty online as a means of creating a more socially desirable profile. Both findings contribute to the construct validity of the linguist honesty measure by demonstrating previously established nomological networks. The consistency in the direction and effect size of the profanity-honesty relationship across the three studies further raises confidence in this approach to measuring dishonesty.

### **Extending to society level**

Our research offers a first look at the use of profanity at a society level. Using the large-scale sample of American participants from Study 2, we were able to devise state-level rates of profanity for use in Study 3. Addressing calls for psychological research to attempt to examine the social implications of psychological findings (Back, 2015; Back & Vazire, 2015), we used this measure in order to examine whether the positive relationship between profanity and honesty

found at the individual level could be extended to the society level. Such an attempt involves many challenges, as there are many variables that may intervene or offer competing explanations for a detected relationship. Yet we believe that this is an important first attempt to provide a baseline for further investigation. The consistent findings across the studies suggest that the positive relation between profanity and honesty is robust, and that the relationship found at the individual level indeed translates to the society level.

### **Implications and future directions**

We briefly note several limitations in the current research and these are further discussed in the supplementary materials with implications and future directions. First, the three studies were correlational, thus preventing us from drawing any causal conclusions. Second, the dishonesty we examined in Studies 1 and 2 was mainly about self-promoting deception to appear more desirable to others, rather than blunt unethical behavior. We therefore caution that the findings should not be interpreted to mean that the more a person uses profanity the less likely he or she will engage in more serious unethical or immoral behaviors. Third, the measures in Study 2 were proxies using an aggregation of linguistic analysis of online behavior using Facebook over a long period of time. Finally, Simpson's Paradox (Simpson, 1951) points to conceptual and empirical differences in testing a relationship on different levels of analysis, and therefore the state-level findings of Study 3 are conceptually broader than the findings in Studies 1-2.

These limitations notwithstanding, our research is a first step in exploring the profanity-honesty relationship, and we believe that the consistent effect across samples, methodology, and levels of analysis contributes to our understanding of the two constructs and paves the way for future research. Future studies could build on our findings to further study the profanity-honesty

relationship using experimental methods to establish causality and incorporating real-life behavioral measures with a wider range of dishonest conduct including unethical behavior.

### **Conclusion**

We set out to provide an empirical answer to competing views regarding the relationship between profanity and honesty. In three studies, at both the individual and society level, we found that a higher rate of profanity use was associated with more honesty. This research makes several important contributions by taking a first step to examine profanity and honesty enacted in naturalistic settings, using large samples, and extending findings from the individual level to a look at the implications for society.



### References

- Abe, N. (2011). How the Brain Shapes Deception: An Integrated Review of the Literature. *The Neuroscientist*, 17(5), 560–574.
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440.
- “Are people who swear more honest?” *Debate.org*. Retrieved December 1, 2015, from <http://www.debate.org/opinions/are-people-who-swear-more-honest>
- Ayal, S., & Gino, F. (2012). Honest Rationales for Dishonest Behavior. In M. Mikulincer & P. G. Shaver (Eds.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil* (pp. 149–166). Washington, DC: American Psychological Association.
- Back, M. D. (2015). Opening the Process Black Box: Mechanisms Underlying the Social Consequences of Personality. *European Journal of Personality*, 29(2), 91–96.
- Back, M. D., & Vazire, S. (2015). The Social Consequences of Personality: Six Suggestions for Future Research. *European Journal of Personality*, 29(2), 296–307.
- Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook Profiles Reflect Actual Personality, Not Self-Idealization. *Psychological Science*, 21(3), 372–374.
- Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85(3), 349–360.
- Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Linguistic classification of prisoners’ truthful and deceptive natural language. *Applied Cognitive Psychology*, 19(3), 313–329.

- Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., Chen, S. X., & Bond, M. H. (2015). Immorality East and West: Are Immoral Behaviors Especially Harmful, or Especially Uncivilized? *Personality and Social Psychology Bulletin, 41*(10), 1382–1394.
- Buffardi, L. E., & Campbell, W. K. (2008). Narcissism and Social Networking Web Sites. *Personality and Social Psychology Bulletin, 34*(10), 1303–1314.
- Cavanna, A. E., & Rickards, H. (2013). The psychopathological spectrum of Gilles de la Tourette syndrome. *Neuroscience & Biobehavioral Reviews, 37*(6), 1008–1015.
- Childs, J. (2012). Gender differences in lying. *Economics Letters, 114*(2), 147–149.
- DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74*(1), 63–79.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979–995.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118.
- Dreber, A., & Johannesson, M. (2008). Gender differences in deception. *Economics Letters, 99*(1), 197–199.
- Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences, 6*(1), 21–29.
- Fang, F. C., & Casadevall, A. (2013). Why We Cheat. *Scientific American Mind, 24*(2), 30–37.
- Feldman, G., Chao, M. M., Farh, J. L., & Bardi, A. (2015). The motivation and inhibition of breaking the rules: Personal values structures predict unethicity. *Journal of Research in Personality, 59*, 69–80.
- Freud, S. (1901). *The Psychopathology of Everyday Life*. SE.

- Friesen, L., & Gangadharan, L. (2012). Individual level evidence of dishonesty and the gender effect. *Economics Letters*, *117*(13), 624–626.
- Garcia, D., & Sikström, S. (2014). The dark side of Facebook: Semantic representations of status updates predict the Dark Triad of personality. *Personality and Individual Differences*, *67*, 92–96.
- Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and Differentiation in Unethical Behavior: The Effect of One Bad Apple on the Barrel. *Psychological Science*, *20*(3), 393–398.
- Granhag, P. A., & Vrij, A. (2005). Deception Detection. In N. Brewer & K. D. Williams (Eds.), *Psychology and Law: An Empirical Perspective* (pp. 43–92). New York, NY: Guilford Press.
- Halevy, R., Shalvi, S., & Verschuere, B. (2013). Being Honest About Dishonesty: Correlating Self-Reports and Actual Lying. *Human Communication Research*, *40*(1), 54–72.
- Hancock, J. T. (2009). Digital deception: Why, when and how people lie online. In A. N. Joinson, K. Y. A McKenna, T. Postmes, & U. Reips (Eds.), *Oxford Handbook of Internet Psychology*, (pp.289–301). Oxford, UK: Oxford University Press.
- Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication. *Discourse Processes*, *45*(1), 1–23.
- Hauch, V., Masip, J., Blandón-Gitlin, I., & Sporer, S. L. (2012). Linguistic Cues to Deception Assessed by Computer Programs: A Meta-Analysis. In *Proceedings of the EACL 2012 Workshop on Computational Approaches to Deception Detection* (pp. 1–4). Avignon, France: Association for Computational Linguistics.

- Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*(6202), 1340–1343.
- Holtgraves, T. (2011). Text messaging, personality, and the social context. *Journal of Research in Personality*, *45*(1), 92–99.
- Holtzman, N. S., Vazire, S., & Mehl, M. R. (2010). Sounds like a narcissist: Behavioral manifestations of narcissism in everyday life. *Journal of Research in Personality*, *44*(4), 478–484.
- Jay, T. (1992). *Cursing in America: A psycholinguistic study of dirty language in the courts, in the movies, in the schoolyards and on the streets*. Amsterdam, the Netherlands: John Benjamins Publishing Company.
- Jay, T. (2000). *Why We Curse: A neuro-psycho-social theory of speech*. Amsterdam, the Netherlands: John Benjamins Publishing Company.
- Jay, T. (2009). The Utility and Ubiquity of Taboo Words. *Perspectives on Psychological Science*, *4*(2), 153–161.
- Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research. Language, Behaviour, Culture*, *4*(2), 267–288.
- Jay, T., Caldwell-Harris, C., & King, K. (2008). Recalling Taboo and Nontaboo Words. *The American Journal of Psychology*, *121*(1), 83–103.
- Kalshoven, K., Den Hartog, D. N., & De Hoogh, A. H. (2011). Ethical Leader Behavior and Big Five Factors of Personality. *Journal of Business Ethics*, *100*(2), 349–366.
- Kaplan, H. B. (1975). Increase in self-rejection as an antecedent of deviant responses. *Journal of Youth and Adolescence*, *4*(3), 281–292.

- Kaye, B. K., & Sapolsky, B. S. (2009). Taboo or Not Taboo? That is the Question: Offensive Language on Prime-Time Broadcast and Cable Programming. *Journal of Broadcasting & Electronic Media*, 53(1), 22–37.
- Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist*, 70(6), 543.
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805.
- Lacan, J. (1968). *The Language of the Self*. (A. Wilden, Trans.). Baltimore, MD: Johns Hopkins Press.
- Mabry, E. A. (1974). DIMENSIONS OF PROFANITY. *Psychological Reports*, 35(1), 387–391.
- Manago, A. M., Taylor, T., & Greenfield, P. M. (2012). Me and my 400 friends: The anatomy of college students' Facebook networks, their communication patterns, and well-being. *Developmental Psychology*, 48(2), 369–380.
- Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*, 45(6), 633–644.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, 90(5), 862–877.
- Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: a psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology*, 84(4), 857–870.

- Merryman, S. (2008). USSWM: Stata module to provide US state and county spatial weight (contiguity) matrices. *Statistical Software Components*.
- Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying Words: Predicting Deception from Linguistic Styles. *Personality and Social Psychology Bulletin*, *29*(5), 665–675.
- Patrick, G. T. W. (1901). The psychology of profanity. *Psychological Review*, *8*(2), 113–127.
- Paulhus, D. L. (1991). Measurement and Control of Response Bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press, Inc.
- 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), 202–217.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, *54*(1), 547–577.
- Piff, P. K., Stancato, D. M., Côté, S., Mendoza-Denton, R., & Keltner, D. (2012). Higher social class predicts increased unethical behavior. *Proceedings of the National Academy of Sciences*, *109*(11), 4086–4091.
- Rassin, E., & Muris, P. (2005). Indecisiveness and the interpretation of ambiguous situations. *Personality and Individual Differences*, *39*(7), 1285–1291.
- Rassin, E., & Heijden, S. V. D. (2005). Appearing credible? Swearing helps! *Psychology, Crime & Law*, *11*(2), 177–182.
- Inbau, F. E., Reid, J., Buckley, J., & Jayne, B. (2011). *Criminal Interrogation and Confessions* (5th ed.). Burlington, MA: Jones & Bartlett Publishers.

Robbins, M. L., Focella, E. S., Kasle, S., López, A. M., Weihs, K. L., & Mehl, M. R. (2011).

Naturalistically observed swearing, emotional support, and depressive symptoms in women coping with illness. *Health Psychology, 30*(6), 789–792.

Sapolsky, B. S., & Kaye, B. K. (2005). The Use of Offensive Language by Men and Women in Prime Time Television Entertainment. *Atlantic Journal of Communication, 13*(4), 292–303.

Scherer, C. R., & Sagarin, B. J. (2006). Indecent influence: The positive effects of obscenity on persuasion. *Social Influence, 1*(2), 138–146.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... Ungar, L. H. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE, 8*(9), e73791.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The Prevalence of Lying in America: Three Studies of Self-Reported Lies. *Human Communication Research, 36*(1), 2–25.

Simpson, E. H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society, Series B (Methodological), 13*(2), 238–241.

Slatcher, R. B., Chung, C. K., Pennebaker, J. W., & Stone, L. D. (2007). Winning words: Individual differences in linguistic style among U.S. presidential and vice presidential candidates. *Journal of Research in Personality, 41*(1), 63–75.

Sopan, D. (2015, September 30). Donald Trump uses profanity to describe Bush-Rubio relationship. *CBS News*. Retrieved from <http://www.cbsnews.com/news/donald-trump-uses-profanity-to-describe-bush-rubio-relationship>

Stone, T. E., McMillan, M., & Hazelton, M. (2015). Back to swear one: A review of English language literature on swearing and cursing in Western health settings. *Aggression and Violent Behavior, 25*, 65–74.

- Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). Predicting Dark Triad Personality Traits from Twitter Usage and a Linguistic Analysis of Tweets. *Proceedings from ICMLA '12: The 2012 11th International Conference on Machine Learning and Applications, 2*, 386–393.
- Sylwester, K., & Purver, M. (2015). Twitter Language Use Reflects Psychological Differences between Democrats and Republicans. *PLOS ONE, 10*(9), e0137422.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology, 29*(1), 24–54.
- Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating Fact from Fiction: An Examination of Deceptive Self-Presentation in Online Dating Profiles. *Personality and Social Psychology Bulletin, 34*(8), 1023–1036.
- Vingerhoets, A. J., Bylsma, L. M., & de Vlam, C. (2013). Swearing: A Biopsychosocial Perspective. *Psihologijske Teme, 22*(2), 287–304.
- Waggoner, A. S., Smith, E. R., & Collins, E. C. (2009). Person perception by active versus passive perceivers. *Journal of Experimental Social Psychology, 45*(4), 1028–1031.
- Walumbwa, F. O., & Schaubroeck, J. (2009). Leader personality traits and employee voice behavior: Mediating roles of ethical leadership and work group psychological safety. *Journal of Applied Psychology, 94*(5), 1275–1286.
- Wang, N., Kosinski, M., Stillwell, D., & Rust, J. (2012). Can Well-Being be Measured Using Facebook Status Updates? Validation of Facebook's Gross National Happiness Index. *Social Indicators Research, 115*(1), 1–9.



- Ward, M. D., & Gleditsch, K. S. (2008). *Spatial Regression Models* (Vol. 155). Thousand Oaks, CA: Sage Publications, Inc.
- Weisbuch, M., Ivcevic, Z., & Ambady, N. (2009). On being liked on the web and in the “real world”: Consistency in first impressions across personal webpages and spontaneous behavior. *Journal of Experimental Social Psychology, 45*(3), 573–576.
- Webster, D., & Duffy, D. (2016). Losing faith in the intelligence–religiosity link: New evidence for a decline effect, spatial dependence, and mediation by education and life quality. *Intelligence, 55*, 15-27.
- Whitty, M., & Gavin, J. (2001). Age/Sex/Location: Uncovering the Social Cues in the Development of Online Relationships. *Cyberpsychology & Behavior, 4*(5), 623–630.
- Whitty, M. T. (2002). Liar, liar! An examination of how open, supportive and honest people are in chat rooms. *Computers in Human Behavior, 18*(4), 343–352.
- Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science, 7*(3), 203–220.
- Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of research in personality, 44*(3), 363-373.
- Zywica, J., & Danowski, J. (2008). The Faces of Facebookers: Investigating Social Enhancement and Social Compensation Hypotheses; Predicting Facebook™ and Offline Popularity from Sociability and Self-Esteem, and Mapping the Meanings of Popularity with Semantic Networks. *Journal of Computer-Mediated Communication, 14*(1), 1–34.

## **Supplementary materials**

### Contents

Power analyses .....	2
Study 1 .....	2
Study 2 .....	2
Study 3 .....	2
Materials used .....	3
Study 1 .....	3
Measures.....	3
Study 2 .....	6
LIWC analysis procedure and code .....	6
Honesty measure .....	7
Study 3 .....	8
Implications and future directions.....	10

## Power analyses

### Study 1

According to Richard, Bond Jr., and Stokes-Zoota (2003) the average correlation in social psychology is  $r = \sim .21$ .

Study 1 served as a preliminary test for the profanity-honesty relationship, and as indicated in the introduction we had no prior indications of the direction or the strength of the relationship. Using the  $r = .21$  estimate, G\*Power 3.1.9.2 power analyses with an alpha of 0.05 and power of .95 resulted in a required sample of 284. The final sample after exclusions included 276 participants.

In Study 1, we observed effects of  $r = .34 / .20 / .13$ , which on average is close to the Richard et al. (2003) estimates.

### References:

Richard, F. D., Bond Jr, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, 7(4), 331.

### Study 2

Given the large sample size (73,789) the power is very close to 1.00.

### Study 3

Study 3 examined state-level variables, and consisted of the entire population of the 50 US states. The posthoc power analysis using G\*Power 3.1.9.2 post-hoc power analyses with an alpha of 0.05 and a sample of 50 for the effect observed in Studies 1 and 2 of  $r = .20$  indicated a power of .41 (one tail).

The observed effect was  $r = .35$ .

## Materials used

### Study 1

#### Measures

##### *Profanity behavioral measure*

#### Guidelines for coding:

Your task is to code the number of curse phrases written by the participants. Participants were asked to first write down a list of all the curse phrases that they use the most, and then write down their favorite curse phrases. Since we did not give any indication of the number of curse phrases we expected, the total number of curse words is used as a behavioral proxy for their tendency to use profanity in their everyday lives. Therefore, we need to count the number of curse phrases provided by each participant in two categories – (1) used the most, and (2) liked the most.

The instructions given to the participants were:

- **used the most:** Please list the curse words you use the most (feel free, don't hold back). Please enter the curses separated by a comma (,) or a semicolon (;). If you do not use curse words please write "NONE"
- **liked the most:** Please list the curse words you like the most (feel free, don't hold back). Please enter the curses separated by a comma (,) or a semicolon (;). If you do have any favorite curse words please write "NONE"

For each participant enter the value in the corresponding field:

- count\_most: word count for the use\_most SPSS field.
- count\_like: word count for the use\_like SPSS field.

Things to note:

- Phrases are expected to be separated by commas or semicolon, BUT this isn't always the case. This is why automated coding isn't possible.
- NONE is not a counted word. If participants indicated NONE then please code 0 (zero).
- A curse phrase can be more than one word, for example "god damn" is a single curse phrase (code 1), not two (do not code 2).
- Phrases with a repeating word that are different are counted separately. For example, "fuck, fuck off, fuck you" are three curse phrases, not one.
- If a curse phrase repeats with different spellings, then do not count it again. If a curse phrase repeats using a completely different word, then count it again. For example, fuck and phuck are counted as 1. Poop and shit are counted as 2.

- Some participants add explanations, such as “mostly just”. A complicated example is: “Fuck. (I usually spell it Phuck) and Shit probably Does poop count? I use poop on occasion.” Is counted as 1 for fuck (do not count Phuck as different), 1 for shit, and 1 for poop, ignoring all other text.
- Some participants combined a few curse phrases/words together to form an unfamiliar curse, these should be separated to single curse words. For example: “shit dammit hell no” is counted as 1 for “shit”, 1 for “dammit”, 1 for “hell no”, 3 overall. However, complete phrases that bind together are counted as 1, for example “Fucking cunt nigger” or “cunt face whore” are both counted as a single curse phrase since there is a connection between them to form one coherent curse.
- No answers should be coded as 99 (missing values), and NOT as 0 (zero).
- Spelling does not matter, count the curse words even if spelling is wrong.

#### *Profanity self-reported measure*

In this section we're interested in the use of profanity - cursing, swearing, and the use of bad language.

1. How often do you curse (swear / use bad language) verbally in person (face to face)?
2. How often do you curse (swear / use bad language) in writing (e.g., texting/messaging/posting online/emailing)?
3. How often do you curse (swear / use bad language) verbally in private (no one around)?

Scale for the three items:

1. Never
2. Once a Year or Less
3. Several Times a Year
4. Once a Month
5. 2-3 Times a Month
6. Once a Week
7. 2-3 Times a Week
8. 4-6 Times a Week
9. Daily
10. A few times a day

#### *Reasons for profanity*

Which of the following might be a reason for you to curse (swear / use bad language) ?

Scale: Never a reason to swear for me (0) - Very often a reason to swear for me (5)

1. Habit
2. Strengthening my argument
3. Expressing positive emotions (e.g., surprise, enthusiasm, or admiration)
4. Expressing negative emotions (e.g., anger, frustration, or pain)
5. Insulting or shocking others
6. Expressing my true self
7. Being honest about my feelings

Please answer the following questions using the scale (0 = Not at all ; 5 = To a very large extent) :

1. Does the use of swearwords strengthen your argument?
2. Do you get intimidated by others if they swear?
3. Do you succeed in expressing positive emotions by means of swearing?
4. Do you feel relieved after swearing in reaction to negative emotions?
5. Do you feel you shock or insult people by means of swearing?
6. Do you feel swearing allows you to be your true self?
7. Do you feel swearing allows you to be more honest about your feelings?

*Eysenck Personality Questionnaire Revised short scale*

(Eysenck, Eysenck, & Barrett, 1985):

1. Does your mood often go up and down?
2. Do you take much notice of what people think?
3. Are you a talkative person?
4. If you say you will do something, do you always keep your promise no matter how inconvenient it might be?
5. Do you ever feel 'just miserable' for no reason?
6. Would being in debt worry you?
7. Are you rather lively?
8. Were you ever greedy by helping yourself to more than your share of anything?
9. Are you an irritable person?
10. Would you take drugs which may have strange or dangerous effects?
11. Do you enjoy meeting new people?
12. Have you ever blamed someone for doing something you knew was really your fault?
13. Are your feelings easily hurt?
14. Do you prefer to go your own way rather than act by the rules?
15. Can you usually let yourself go and enjoy yourself at a lively party?
16. Are all your habits good and desirable ones?
17. Do you often feel 'fed-up'?
18. Do good manners and cleanliness matter much to you?
19. Do you usually take the initiative in making new friends?
20. Have you ever taken anything (even a pin or button) that belonged to someone else?
21. Would you call yourself a nervous person?
22. Do you think marriage is old-fashioned and should be done away with?
23. Can you easily get some life into a rather dull party?
24. Have you ever broken or lost something belonging to someone else?
25. Are you a worrier?
26. Do you enjoy co-operating with others?
27. Do you tend to keep in the background on social occasions?
28. Does it worry you if you know there are mistakes in your work?
29. Have you ever said anything bad or nasty about anyone?
30. Would you call yourself tense or 'highly-strung'?
31. Do you think people spend too much time safeguarding their future with savings and insurances?
32. Do you like mixing with people?
33. As a child were you ever cheeky to your parents?

34. Do you worry too long after an embarrassing experience?
35. Do you try not to be rude to people?
36. Do you like plenty of bustle and excitement around you?
37. Have you ever cheated at a game?
38. Do you suffer from 'nerves'?
39. Would you like other people to be afraid of you?
40. Have you ever taken advantage of someone?
41. Are you mostly quiet when you are with other people?
42. Do you often feel lonely?
43. Is it better to follow society's rules than go your own way?
44. Do other people think of you as being very lively?
45. Do you always practice what you preach?
46. Are you often troubled about feelings of guilt?
47. Do you sometimes put off until tomorrow what you ought to do today?
48. Can you get a party going?

Lie items:

- YES: 4, 16, 45
- NO: 8, 12, 20, 24, 29, 33, 37, 40, 47

#### *Attention checks*

We also added two attention checks to the lie scale randomly mixed with the scale items:

1. Are balls round? (Yes/No)
2. Do rich people have less money than poor people? (Yes/No)

Correct answers for inclusion: 1 – Yes, 2 – No.

## Study 2

All the myPersonality data including the results of the LIWC dictionary analyses are available for download at: <http://mypersonality.org/wiki/doku.php>

Information about the LIWC use in myPersonality:

[http://mypersonality.org/wiki/doku.php?id=list\\_of\\_variables\\_available#facebook\\_status\\_updates](http://mypersonality.org/wiki/doku.php?id=list_of_variables_available#facebook_status_updates)

Additional information (including alphas) is available from the [LIWC website](#) and the [LIWC manual](#).

#### *LIWC analysis procedure and code*

For the LIWC analyses, we used Autoit scripts to split the main file to separate text files of combined status updates per each participant which were imported in batch mode for analysis with LIWC (by selecting multiple text files).

A sample Autoit script code is (directory names and CSV file structure would need to be adjusted to your code):

```
$Limit = 25407613
$Count = 1
$LogFile = "C:\LIWC\output.txt"
While 1
    $r = FileReadLine("C:\LIWC\myPersonality.csv", $Count)
    If @error = -1 Then
```

```

FileWriteLine($LogFile,"Error at read - " & $Count & @CRLF)
ExitLoop
Else
$input = StringSplit($r, ",")
FileWriteLine($LogFile,$r & @CRLF)
Local $userid = $input[1]
Local $statusupdates = $input[2]
If $statusupdates<> "" Then
FileWriteLine("C:\LIWC\statusupdates\" & $userid & ".txt", ""
& $statusupdates & "")
FileWriteLine($LogFile,"AN " & $Count & @CRLF)
EndIf
EndIf
$Count = $Count + 1
WEnd

```

### Honesty measure

The original model by Newman, Pennebaker, Berry, and Richards (2003) used the LIWC category of negative emotions. The negative emotions category includes, anxiety, anger, and sadness. In their article Newman et al. (2003) wrote (p. 672):

“The “negative emotion” category in LIWC contains a subcategory of “anxiety” words, and it is possible that anxiety words are more predictive than overall negative emotion. However, in the present studies, anxiety words were one of the categories omitted due to low rate of use”.

The purpose of our investigation was to assess the relationship between profanity and honesty, yet profanity is positively associated with anger. Holtzman, Vazire, and Mehl (2010) reported a correlation of .96 (!). Jay and Janschewitz (2008) summarized that “the main purpose of cursing is to express emotions, especially anger and frustration.” (p. 267). Therefore, if we were to measure honesty using anger words as the negative emotions category then the profanity-honesty correlations would be strongly biased towards a negative relationship. This is less of a problem with anxiety, as – for example - Yarkoni (2010) found swearing to be strongly associated with anger but not anxiety (see Table 3, p. 368).

In light of the issue regarding the use of anger words in the strong profanity-anger associations and the Newman et al. (2003) note, we decided to focus only on anxiety in as negative emotions and these are the findings reported in the main manuscript.

For transparency, we also report the use of the original equation here:

#### Usage rates table

LIWC dimensions	Sample LIWC keywords	Honesty coefficients $\beta$ s	Percentage (M)	Percentage (SD)
Model 1: Negative emotions	Hurt, ugly, nasty	-0.217	1.93%	.77%
Model 2: Anxiety	Worried, fearful, nervous	-0.217	0.21%	.14%



### Reporting with both equations

Model 1 using negative emotions:  $r = .02, p < .001$ ; 95% CI [.01, .03];

Model 2 using anxiety words:  $r = .20, p < .001$ ; 95% CI [.19, .21]

Controlling for age, gender, and network size:

Model 1: partial  $r = .04, p < .001$ ; 95% CI [.03, .05].

Model 2: partial  $r = .22, p < .001$ ; 95% CI [.21, .22].

### *Study 2: Descriptive statistics of honesty, profanity, and demographics.*

	Mean	SD	Skewness	Kurtosis	1	2	3	4	5
1. Honesty 1 (raw)	0 1.23	1 .56	.09	.01	(-)		.04		
2. Honesty 2 (raw)	0 1.60	1 .60	.03	.02	.97	(-)	.22		
3. Profanity (raw)	.28 .37	.26 .43	1.37 2.51	2.00 9.49	.02	.20	(-)		
4. Age	25.34	8.78	1.90	3.96	.02	-.05	-.18	(-)	
5. Gender	.62	.49	-.49	-1.76	.16	.12	-.23	.08	(-)
6. Network size (raw)	5.30 272.37	.79 249.71	-.03 4.18	-.25 39.82	-.16	.18	-.09	-.13	.00(ns)

### Study 3

The State Integrity Investigation 2012 data is available online:

<https://www.publicintegrity.org/accountability/state-integrity-investigation/state-integrity-2012>

The FBI crime statistics are available online: <https://www.fbi.gov/stats-services/crimestats>

Information about the State Integrity Investigation 2012 is available on

<https://www.publicintegrity.org/2012/03/19/8429/methodology>.

Following are key quotes regarding the methodology:

The State Integrity Investigation mobilized a network of state reporters to generate quantitative data and qualitative reporting on the health of the anti-corruption framework at the state level. Reporters scored 330 questions about their state's accountability and transparency framework — what we call Corruption Risk Indicators — by combining extensive desk research with thousands of original interviews of state government experts. Among those interviewed were current and former officials with experience inside the bureaucracy, as well as private sector researchers, journalists and executives from civil society and 'good government' organizations.

To identify the project's Corruption Risk Indicators, staff from Global Integrity (<http://www.globalintegrity.org/>) and the Center for Public Integrity contacted nearly 100 state-level organizations working in the areas of good government and public sector reform nationwide. We asked them a simple question: what issue areas mattered most in their state when it came to the risk of significant corruption occurring in the public sector? The outcome was a list of questions, rooted in the day-to-day realities of state government operations. In addition, Global Integrity and the Center for Public Integrity included additional indicators that the two organizations had previously fielded in similar projects. Specifically, these indicators were drawn from the Center's States of Disclosure project and Global Integrity's Global Integrity Report and Local Integrity Initiative efforts. [...]

All indicators were scored by the lead state reporters based on original interviews and desk research. The data were then blindly reviewed by a peer reviewer for each state who was asked to flag indicators that appeared inaccurate, inconsistent, biased, or otherwise deserving of correction. [...]

## Implications and future directions

To meet the Social Psychological and Personality Science word count restrictions, the implications and future directions section has been summarized in the main manuscript and is elaborated below.

We note several limitations in the current research. The three studies were correlational, thus preventing any causal interpretation of the findings. Therefore, we cannot draw conclusions on whether honesty is driving profanity or profanity is driving more honest behavior (and it is possible that both may be acting simultaneously); however, theoretically, it seems more plausible that honesty is driving the higher use of profanity as an expression of a genuine self. Future studies may aim for a research design that would allow for the drawing of casual conclusions.

In Study 2, we restricted our examination to the use of mild profanity in interactions with friends and family on Facebook, and in this context, dishonesty is mainly self-promoting deception to make oneself appear more desirable. Therefore, the conclusion that can be drawn is that those who use mild profanity in familiar social surroundings do tend to express themselves in a more genuine manner, possibly because they care less about having to appease others and promote themselves. Since the most common forms of dishonesty are white lies (DePaulo & Kashy, 1998), and the most common form of profanity is mild profanity in social settings (Jay, 2009), there is much merit in these findings. However, we caution that these findings should not be interpreted to mean that the more a person uses profanity the less likely he or she will engage in more serious unethical or immoral behaviors.

We also recognize that our dishonesty measure in Study 2 is only a proxy indicating a likelihood of dishonesty based on linguistic analysis, rather than directly observing dishonesty. Although this is currently the most accurate tool available for measuring honesty online (as far as we know), future research may attempt to capture dishonesty using more direct methods. As detecting lies and evaluating individuals' dishonesty are challenging, we consider our work as an important first step in examining the relation between profanity and dishonesty, and we call for further empirical work to examine this in more depth.

In addition, we note that in Study 2, our measures of profanity and honesty reflect an aggregation of behaviors over a period of time. As we aimed to understand whether people who curse more are more or less dishonest in general, we consider our method to be appropriate. However, future research may examine the context-specific associations between the two constructs in order to see whether individuals are being honest or trying to hide the truth when they curse. Furthermore, different dimensions of deception can be used for different purposes (Hancock, 2007), and future research may further attempt a more nuanced understanding of the link between profanity and different types of deception.

Lastly, in Study 3, we examined the relationship between profanity use and dishonesty at a society level. Simpson's Paradox (Simpson, 1951) warns about the conceptual and empirical differences in testing a relationship on different levels, and that variables hold different meaning across levels or when aggregated. It was therefore important that we examine and find consistent results across levels; however, we caution that the results from Study 3 can only provide limited support for the individual-level assessments in Studies 1 and 2. Another point to consider regarding measuring society-level variables is that the assessment of dishonesty at a society level is complicated and other proxies could have been used (such as crime rates). However, we believe that the integrity index is the closest proxy to the type of honesty that we examined in Studies 1 and 2. Examining

behavioral variables at a society level also introduces a new level of complexity, and there are many societal factors that may be taken into account. Our research has taken the first step in establishing a main effect, and future research can extend this investigation by looking at other forms of dishonesty and testing for moderating and mediating factors (economy, culture, etc.).

The view that those who use profanity more frequently are more honest is mainly based on the premise that a higher use of profanity is indicative of having fewer social filters and indicative of a more genuine self. For the individual, this can be reflected in lower agreeableness and social desirability, and therefore less lying in order to appease others. In Studies 1 and 2, the dishonesty indicators were of subtle deception that involved no direct harm to others, rather than of unethical behavior with a clear violation of rules or moral codes. Quite possibly, socially desirable deception on Facebook may in fact represent the implicit social norm, and those who disregard it as non-conformists. Therefore, the behavior of a person who bends the truth to appear more desirable to researchers (Study 1), or to friends and family on Facebook (Study 2), is not necessarily indicative of more extreme unethical acts (Gaspar, Levine, & Schweitzer, 2015). An interesting research extension would be to examine whether profanity would also be associated with blunter dishonesty, which involves harmful and unethical behavior.

Study 2 is based on a Facebook sample and examined online behavior. There may be concerns about whether findings from online behavior can be generalized to real-life behavior. However, research of behavior on Facebook has shown that online behavior is reflective of real life (Back et al., 2010; Wilson et al., 2012), and can predict real-life outcomes (Bond et al., 2012; Qiu, Lin, Ramsay, & Yang, 2012; Saslow, Muise, Impett, & Dubin, 2013; Seder & Oishi, 2012). The findings of Study 2 were consistent with those of Study 1, where we measured real-life behavior, supporting the generalizability of the findings. Future research may set out to further test whether the positive relation between profanity and honesty found with a Facebook population can be generalized to a population in real life.