Authentic First Impressions Relate to Interpersonal, Social, and Entrepreneurial Success

David M. Markowitz, Maryam Kouchaki, Francesca Gino,

Jeffrey T. Hancock, and Ryan L. Boyd

Accepted Manuscript: Social Psychological and Personality Science

Abstract

This paper examines how verbal authenticity influences person perception. Our work combines human judgments and natural language processing to suggest verbal authenticity is a positive predictor of interpersonal interest (Study 1: 294 dyadic conversations), engagement with speeches (Study 2: 2,655 TED talks), entrepreneurial success (Study 3: 478 Shark Tank pitches), and social media engagements (Studies 4a-c; N = 387,039 Tweets). We find that communicating authenticity is associated with increased interest in and perceived connection to another person, more comments and views for TED talks, receiving a financial investment from investors, and more social media likes and retweets. Our work is among the first to evaluate how authenticity relates to person perception and manifests naturally using verbal data.

Keywords: authenticity, impression formation, natural language processing, first impressions

Authentic First Impressions Relate to Interpersonal, Social, and Entrepreneurial Success

People quickly form impressions of others using limited informational cues. For example, people make judgments of attractiveness and trustworthiness after brief exposure to a person's face (Todorov et al., 2009; Willis & Todorov, 2006). Emotional displays also influence our impression of others (Grandey et al., 2005). For example, an authentic smile during a service encounter leads the service provider to appear friendlier.

In addition to facial and emotional cues, communication is an important vehicle for impression formation. Hearing a person speak can influence physical attractiveness impressions (Snyder et al., 1977) and judgments of their competence, intelligence, and thoughtfulness as compared to written text (Schroeder & Epley, 2015). This body of research suggests that, in addition to nonverbal cues, people use communication patterns to judge others on key interpersonal attributes, such as warmth and competence.

One dimension many people care about when it comes to managing impressions is authenticity. People want to act in ways that are consistent with their "true self" and care about being perceived as authentic (Kernis & Goldman, 2006), for good reason. Perceptions of authenticity are linked to positive outcomes across a variety of domains, including hospitality (Grandey et al., 2005), morality (Gino et al., 2015), medicine (Árnason, 1994), and academia (Archer, 2008). Observers' perceptions of authenticity, which are consequential (Lehman et al., 2019), rely on external cues. Here, we consider how language can be a vehicle for transmitting authenticity cues and how perceptions of authenticity in language relate to interpersonal, social, and entrepreneurial outcomes. We argue that a speaker's words affect perceptions of their authenticity and that specific language features indicate authentic versus inauthentic communication. In four studies, we use natural language processing and human judgments to examine these main hypotheses.

Impression Formation Through Language

People often rely on verbal cues when forming impressions of others, and the amount of language required for person perception is relatively small. For instance, effective salespeople can be identified within twenty seconds of their verbal output (Ambady et al., 2006), and judges can predict if a couple will divorce by using emotion markers from the beginning of conflict discussions (Carrère & Gottman, 1999). Several language patterns relate to impression formation in short first interactions. Prior work suggests interest in a speed-dating partner can be predicted by similar use of style words (Ireland et al., 2011). The more two people matched on style words (e.g., articles, pronouns), the greater interest they expressed in meeting the partner for a future date. In another study, positive job interview performance was associated with an increase in collective pronouns and fewer filler words (Naim et al., 2018).

Language, then, can be a powerful vehicle for impression formation. Drawing on this main idea, we examine how people make authenticity judgments of others based on words and how verbal authenticity relates to meaningful social and entrepreneurial outcomes. This work is theoretically important, as typical evaluations of authenticity use self-report data to rate targets (Grandey et al., 2005; Krumhuber et al., 2007). Investigations using language data in impression formation are less common, though prior work suggests words are important conduits of psychological information and can indicate a range of psychological dynamics, including authenticity (Jordan et al., 2018; Pennebaker, 2011). We draw on this empirical foundation to evaluate how content words (e.g., nouns, verbs) and style words (e.g., articles, pronouns) reveal authentic communication patterns and how authenticity markers associate with positive interpersonal, social, and entrepreneurial outcomes.

Conceptualizing Verbal Authenticity

Authenticity describes the perception that a person is acting consistently with their true and core values (Gecas, 1991; Kernis & Goldman, 2006). Verbal authenticity extends this general definition by indicating a person who communicates their true and core self with words to reveal their psychological focus. This "words as attention" model has been used in psychology to understand how people attend to specific topics or aspects of the self (Boyd & Schwartz, 2021). Words can reflect social and psychological processes, such as authenticity, and should therefore have the potential to influence perceptions of authenticity.

Authentic people communicate original ideas (Moulard et al., 2015). While authenticity is often linked to "free and natural expression of [one's] core feelings, motives, and inclinations" (Kernis & Goldman, 2006, p. 299), it is distinct from constructs such as honesty: An individual can communicate authentically (and be perceived as authentic) without telling the truth. For example, President Donald Trump is a prolific liar (Kessler et al., 2021), but he communicates in a manner that supporters perceive as authentic because his speech is spontaneous, without reservation, unfiltered, and compatible with his values (Montgomery, 2017).

The current work follows the psychology of language tradition by using an established measure of verbal authenticity, validated and used in past work, to examine how authenticity can be transmitted through language patterns and associate with positive outcomes (Arenas, 2018; Bulkeley & Graves, 2018; Dalvean, 2017; Markowitz & Griffin, 2020; Sell & Farreras, 2017; Tahmasbi & Rastegari, 2018). Our work focuses on verbal authenticity as revealed by what people say (verbal content) and how they say it (verbal style), which extends how researchers currently think about authenticity cues. We expect verbal authenticity to associate with positive behavioral outcomes (e.g., social engagements and entrepreneurial investment) because perceiving a target as authentic often leads to a range of positive self-report outcomes, such as trust and subjective well-being (Kifer et al., 2013; Portal et al., 2019). Our work therefore contributes to authenticity theory by proposing verbal content and style words as cues to indicate how authenticity is communicated and relayed to others, facilitating positive impression formation and outcomes in diverse settings.

We report results from four studies. All preregistrations, data, and statistical code are available online (https://osf.io/y3mpx/?view_only=8205647eaeba4cd5abdf50c6b22b4953). Descriptive statistics and full model outputs for key variables across studies are available in the supplement. Additional study materials are available from the authors upon request due to copyright concerns regarding certain stimuli.

Study 1: Verbal Authenticity Predicts Positive Interpersonal Perceptions

Our first study assessed the idea that language patterns can serve as a vehicle to transmit authenticity, leading to positive social impressions. Such an effect would be consistent with prior research suggesting people form positive impressions of those who are perceived as authentic; we expect a similar trend to occur through language. Instead of asking people to make explicit authenticity judgments about their conversational partner, we evaluated how verbal authenticity in conversation corresponds to person perceptions.

Method

We first conducted a pilot study (N = 224 individuals; 112 dyads) to determine the effect size required for our full investigation. As noted in the online supplement, after participants chatted with a randomly assigned partner online, verbal authenticity in the chat was positively associated with a perceptions composite variable for each dyad, $r(110) = .187$, $p = .049$, 95% CI [0.001, 0.359]. We used this effect size to power our preregistered study at 80% ($\alpha = .05$, twotailed), requiring at least 222 pairs.

The conversations occurred on Chatplat.com. We recruited participants from CloudResearch (*N* = 700 individuals; 350 dyads) and paired them automatically; these individuals were strangers who had never met before. Consistent with our preregistration, we excluded dyads from the analysis if one or more members did not complete the study or failed to follow its instructions (*n* = 112 individuals; 56 dyads).

Each person in our final sample of 588 participants (*n* = 294 dyads) received the following instructions:

> *Today you will be paired with another MTurker in a chatroom. You will have a discussion about what the one thing you would change about yourself would be. It's very important that you do not leave the chat until it says "This chat has now expired." However, if you are idle for more than 30 seconds, you may be removed from the chat room.*

Dyads chatted for 7-10 minutes. Next, each participant answered four questions about their partner: (1) "How much did you like the other MTurker?" (2) "How much would you like to talk to the other MTurker again in the future?" (3) "How close did you feel to the other MTurker?" and (4) "How connected did you feel to the other MTurker?" These questions were measured on 7-point scales ranging from (1) Not at all to (7) Very much. Conversations were analyzed at the dyad level, since both participants received identical instructions and should therefore have similar chat experiences. We averaged responses to each of the four perception questions to form a general evaluation score for the pair. The average perception ratings for the four questions were highly reliable as a collection (Cronbach's $\alpha = 0.95$); therefore, we formed a composite by adding the standardized (*z*-scored) values of the items.

Automated Text Analysis

Conversations between each pair were combined into a single file. We analyzed the texts with Linguistic Inquiry and Word Count (LIWC:(Pennebaker et al., 2015), a program that calculates the percentage of words in a given text across social and psychological dimensions. For example, the phrase “I am happy today” contains four words; LIWC counts several categories in its dictionary, including first-person singular pronouns (e.g., *I*; 25%) and positive emotion terms (e.g., *happy*; 25%).

We evaluated the relationship between interpersonal perceptions and authentic speech as measured by LIWC. The LIWC authenticity index is a composite score of language variables (Jordan et al., 2018), derived from empirical studies, and reflects honest, unfiltered, and spontaneous speech. Dimensions that positively load onto the authenticity index include self-references (e.g., *I*), insight words (e.g., *aware*), differentiation words (e.g., *but*), and relativity terms (e.g., *above*). Dimensions that negatively load onto the authenticity index include discrepancies from reality (e.g., *must*) and third-person singular pronouns (e.g., *she*). The supplementary materials contain the authenticity index formula. Each dyad received an authenticity score ranging from 0 (low authenticity) to 100 (high authenticity). Examples of texts with high and low levels of authenticity are provided in the supplementary materials.

Results

Consistent with our pilot data, verbal authenticity positively predicted the overall composite score, $r(292) = .127$, $p = .029$, 95% CI [0.012, 0.238]. Specifically, verbal authenticity positively predicted average liking of the partner ($p = .006$) and interpersonal connection ($p = .034$). How much people would like to talk to their partner in the future ($p = .054$) and interpersonal closeness ($p = .190$) were in the expected direction but did not reach significance.

These results validate the use of language patterns to assess the connection between verbal authenticity and impression formation. Indeed, the language findings are consistent with work suggesting that authenticity is associated with positive interpersonal perceptions (Árnason, 1994; Kifer et al., 2013). Next, we examine how verbal authenticity in TED talks relates to how viewers react to the video content, with the underlying theory suggesting a more authentic verbal communication style should lead to more online impact. Social psychological and consumer behavior research often attempts to understand the characteristics that lead to online engagement, influence, and interactivity (Cialdini, 2006; Thaler & Sunstein, 2009). Online engagement and interactivity are often indicators of positive impression formation (Sundar, 2007) and therefore, we examine how language patterns of authenticity in TED talks facilitate positive impression formation and relate to increased online behavior.

Study 2: Verbal Authenticity Predicts Social Impact and Engagement

Method

Data Collection and Database Descriptives

Note, this dataset was also used in other research (Boyd et al., 2020; Markowitz & Shulman, 2021; Meier et al., 2020a, 2020b), but there was no overlap in empirical aims across papers. TED talk transcripts were collected from June 2006 through March 2018 (N = 2,655 speeches). We also gathered talk metadata, including the film date, the number of keyword tags for each speech (M = 7.72 tags, SD = 4.28 tags), and the talk impact, defined as the number of views (M = 1,804,030 views, SD = 2,688,951 views) and comments for each talk (M = 14,850.04 comments, SD = 11,676.06 comments).[1] Consistent with our prior study, TED transcripts were

[1] Note, comments were available on TED.com at the time of data collection but were removed some time after. They were retained in this analysis and the results are substantively unchanged if views were used as the only metric of talk impact.

analyzed with LIWC, and each text received an authenticity score.

Talk Impact

Two metadata dimensions constructed a dependent measure of impact: the number of views and comments per video. Because the raw number of views and comments were not normally distributed, we made several adjustments to create an impact score for each talk. First, we calculated the number of days between the film date and the date of data collection. The number of views and comments were then divided by the number of elapsed days to create a relative view and comment score for each talk. These data were then natural log-transformed out of skewness concerns ($\text{skewness}_{\text{Views}} = 4.81$, $\text{kurtosis}_{\text{Views}} = 39.36$; $\text{skewness}_{\text{Comments}} = 15.28$, $\text{kurtosis}_{\text{Comments}} = 461.48$). This transformation was successful ($\text{skewness}_{\text{Views}} = 0.06$, $\text{kurtosis}_{\text{Views}} = -0.47$; $\text{skewness}_{\text{Comments}} = -0.60$, $\text{kurtosis}_{\text{Comments}} = -0.73$). Log-transformed views and comments were highly correlated, ($r(2653) = .709$, $p < .001$); therefore, we combined them into an index called *impact* by adding the standardized values of each dimension.

Results

The data were analyzed using a mixed effects regression with several controls as random intercepts. Consistent with prior work (Markowitz & Shulman, 2021), we controlled for talk topic as a random intercept by identifying the primary keyword for each talk and then dummy-coding the results based on keyword tags with more than 100 mentions (TED Fellows, Business, Africa, Art, Culture, TEDx). The remaining keywords were assigned to an “other” category. We also controlled for the topical breadth of the talk by counting the total number of keyword tags associated with each speech as a fixed effect. Finally, we controlled for data non-independence (e.g., TED presenters may have given multiple talks) with a random intercept for speaker. An expanded rationale for covariates in Studies 2-4 is in the supplement.

Verbal authenticity and impact were positively related (*B* = 0.008, *SE* = 0.002, *t* = 4.25, *p* < .001, R^2c = 0.46, 95% CI [0.004, 0.012]). TED talks with higher verbal authenticity received more comments and more views than TED talks with lower verbal authenticity.

The evidence thus far suggests verbal authenticity relates to forming positive first impressions of another person online (Study 1) and more online impact (Study 2). There are two noteworthy critiques of Study 2, however. First, people can view or comment on TED talks for different reasons; the video of a speaker with an authentic style may have received many comments or views because the content was controversial, for example. It is therefore difficult to understand if greater impact stemmed from a positive experience facilitated by authenticity or a negative experience of authenticity from the perceiver's point of view. Second, it is reasonable to suggest that viewing and commenting on a TED talk are lightweight and low-stakes evaluations. A more crucial test of the link between verbal authenticity and impression formation might involve greater stakes, where communicating inauthentically has direct negative consequences. We address this possibility in Study 3 by evaluating how first impressions and person perceptions of authenticity are reflected in the language patterns of investment pitches.

Study 3: Verbal Authenticity Predicts Entrepreneurial Investment

We used pitches from a popular television show, *Shark Tank*, to examine the link between verbal authenticity and receiving a financial investment. In this show, entrepreneurs "pitch" their business or product to investors and hope to go into business with them.

Method

Data Collection

We collected *Shark Tank* episodes from Season 1 (2009) through Season 10 (2019) by accessing publicly available online videos. Some episodes per season were unavailable;

therefore, our database includes a selection of *Shark Tank* pitches, not the entire archive.[2]

We isolated pitches to create videos for each product, service, or company (N = 478 total videos). Pitches consisted of entrepreneurs communicating why they need funding to investors, or "sharks," who receive equity in return for their investment. Because we were predominantly interested in first impressions and authenticity, our videos included the initial pitch and excluded any subsequent negotiation. Most pitches were 259.25 seconds long (SD = 62.07 seconds), though video length was not statistically different across successful or unsuccessful investment outcomes [Welch's $t(427.38) = -0.39$, $p = .694$, Cohen's $d = 0.04$].

We extracted financial elements of each pitch: *the ask* (e.g., how much money pitchers requested), *equity* (e.g., the percentage of the company sharks would receive in return for their investment), and *valuation* (e.g., the company's worth, calculated by the ask and equity variables). For example, an entrepreneur who asks for a \$10,000 investment for 25% of their company values the business at \$40,000. As expected, bivariate correlations between the ask and valuation, $r(476) = .893$, $p < .001$, and equity and valuation, $r(476) = -.713$, $p < .001$, were strong and significant (variables natural log-transformed). We therefore only used valuation as a fixed effect covariate to prevent collinearity issues in regression analysis. Valuation was natural log transformed out of skewness concerns, and this transformation was successful (raw skewness = 3.01, raw kurtosis = 10.55; transformed skewness = 0.17, transformed kurtosis = -0.48).

We also aggregated metadata from online sources to obtain covariates for each pitch, including the number of pitchers (e.g., people making a pitch to the Sharks) and the business category. On average, businesses that received an investment tended to have more pitchers than those that did not [Welch's $t(472.54) = 2.12$, $p = .035$, Cohen's $d = 0.19$]. This finding motivated

[2] Approximately 88 pitches are aired per season (Levin, 2019). Our data consisted of pitches from the first 10 seasons of the show, suggesting we collected nearly half of the available pitches.

us to control for number of pitchers in our statistical models. There were 14 business categories (see the supplementary materials for details).

Data Transcription and Preprocessing

We used a professional human transcription service to convert the pitch audio data to text. Each speaker, including the pitcher(s) and investor(s), received a unique identifier. We then performed several steps to prepare the data for analysis. First, we removed investor speeches (from the Sharks), since our goal was to associate entrepreneur speech with investment outcome. Second, we removed speaker identifiers. Finally, we combined the text from all speakers of a pitch into one document, making our unit of analysis the entire opening text per pitch.

Automated Text Analyses

On average, entrepreneurs used 514.41 words per pitch (SD = 142.32 words), and the number of words was significantly different across investment outcomes. Pitchers who received an investment ($n = 261$) spoke more than pitchers who did not ($n = 217$), [Welch's $t(447.91) = 2.27$, $p = .023$, Cohen's $d = 0.21$]. As described, LIWC2015 was used to quantify an initial verbal authenticity score from the transcripts.

Human Coding

Two independent coders watched the entire *Shark Tank* corpus and made perception judgments about the pitchers. Coders were instructed to make judgments about the pitchers as a whole and to rate, in general, how much they were perceived as *attractive*, *authentic*, *sincere*, *genuine*, *competent*, *skillful*, *capable*, *warm*, *trustworthy*, and *friendly* (Fiske, 2018). Coders made their ratings on a scale of (1) Not at all to (7) Extremely. Ratings were reliable on each individual measure (Cronbach's αs $> .62$). We also averaged the ratings to form a perceptions score, called the *perceptions index*. Ratings on the perceptions index were highly reliable within coder as well (Cronbach's αs > .92). Please see the supplement for all reliability statistics.

Custom Authenticity Measure (NLP)

To expand beyond LIWC's generalized authenticity measure, we created a language-based model of verbal authenticity as well. The benefits of this process were two-fold: (1) to better understand which cues were being used by coders in their evaluations of authenticity, and (2) to establish the extent to which context-specific cues of authenticity can be inferred computationally. To create an automated generalization of authenticity cues from the human coding process, we conducted a series of Natural Language Processing (NLP) procedures, creating a final authenticity estimator specific to the *Shark Tank* corpus.

We first extracted all common words and two-word phrases (i.e., "n-grams"; two-grams retained where normalized pointwise mutual information exceeded 0.50; (Bouma, 2009) used by speakers that appeared at least 25 times across the entire corpus (N = 1,080). Then, we performed a ridge regression to estimate each n-gram's combined contribution to the human-coded authenticity score (Schwartz et al., 2013); lambda values were determined via cross-validation (Friedman et al., 2021). This modeling technique uses a process known as "regularization" that helps to reduce overfitting typical of a standard multiple linear regression (for a detailed discussion, see: (McNeish, 2015) and produces a linear model that approximated the human-coded authenticity ratings from the transcribed language data of the pitches. The analysis revealed an authenticity estimator that captured nearly half of the variance in the human coding (r = 0.689, p < .001; r^2 = 0.474; 10-fold cross-validated r^2 = 0.110, p < .001). The custom authenticity metric was also positively associated with LIWC authenticity (Table 1). The code used to create the custom NLP measure is available from the authors.

Results

Our dependent variable, investment outcome, was binary (1 = received an investment, 0 = did not receive an investment). We used a logistic regression model to predict investment outcome from LIWC authenticity and several controls (the number of pitchers, product category, natural log transformed valuation) as fixed effects. LIWC authenticity was natural log transformed after these data were more nearly normal in a histogram relative to raw scores.

The probability of receiving an investment on *Shark Tank* increased with more verbal authenticity via LIWC authenticity (B = 0.42, SE = 0.20, z = 2.08, p = .037, Odds Ratio (OR) = 1.52, Nagelkerke's R^2 = 0.063, 95% CI = [0.027, 0.811]). Human coding of the pitchers revealed a consistent signal as well. Pitchers who received an investment were rated more positively on the overall perceptions index (B = 0.79, SE = 0.11, z = 7.34, p < .001, OR = 2.20, Nagelkerke's R^2 = 0.213, 95% CI = [0.582, 1.003]). Isolating the authenticity item from the perceptions index revealed a consistent pattern, with human-coded authenticity positively predicting investment outcome (B = 0.53, SE = 0.09, z = 5.77, p < .001, OR = 1.70, Nagelkerke's R^2 = 0.149, 95% CI = [0.355, 0.715]). LIWC authenticity also positively predicted human-coded authenticity (B = 0.55, SE = 0.11, t = 5.03, p < .001, R^2 = 0.094, 95% CI = [0.336, 0.766]). A correlation matrix of all authenticity measures from this study is available in Table 1.

Table 1

Correlation Matrix of Key Authenticity Measures: Study 3

	Investment Outcome	LIWC authenticity	Perceptions index	Human-coded authenticity
Investment Outcome	--			
LIWC authenticity	.031	--		
Perceptions index	.364**	.150**	--	
Human-coded authenticity	.290**	.193**	.840**	--
Custom authenticity measure (NLP)	.283**	.259**	.776**	.855**

Note. Correlations with the *investment outcome* variable are point-biserial correlations (1 = received an investment, 0 = did not receive an investment). *LIWC authenticity* is the natural log-transformed verbal authenticity measure; *perceptions index* is the average of ten trait judgments made by human coders; *human-coded authenticity* is the human-coded judgments of authenticity; and *custom authenticity measure (NLP)* is the NLP-derived measure of authenticity specific to Shark Tank data. These correlations do not include controls from the regression models. ***p* < .01.

Custom Authenticity Measure (NLP)

We used a logistic regression model to predict investment outcomes from our custom authenticity measure (NLP) and the prior controls. We observed that such authentic language was positively related to receiving an investment [B = 1.51, SE = 0.26, z = 5.71, p < .001, OR = 4.53, Nagelkerke's R^2 = 0.147, 95% CI = [1.004, 2.044]. Results controlling for LIWC authenticity are reported in the online supplement.

What features from our custom measure were most strongly related to perceptions of authenticity from human coding? Table 2 displays that politeness and language reflecting compliments (e.g., *thank* you) tended to relate positively to authenticity. Finally, pitchers who used many financial terms or inflated their self-presentation (e.g., the *CEO of* their company, seeking *an investment*) tended to be perceived as less authentic. It could be that pitchers who used these terms were viewed as using canned entrepreneurial language that appears to be less genuine, rehearsed, or conceited.

Although we found positive support for the idea that language can transmit authenticity cues in three studies, language patterns were also communicated with other cues (e.g., nonverbals, synchronicity). Therefore, our final study evaluated the relationship between authenticity and positive social outcomes in a setting that is predominantly text-based. Specifically, we examined how Tweet authenticity associates with social media engagements.

Table 2

Terms Positively and Negatively Associated with Perceptions of Authenticity

Positively Associated with Authenticity		Negatively Associated with Authenticity	
Term	Estimate	Term	Estimate
Thousands	151.16	lot	-129.57
little bit	83.94	world's	-86.19
United	63.03	lot of	-52.00
Grew	48.71	least	-43.18
samples for	33.00	150,000	-37.81
Able	32.57	amount	-35.72
Type	31.40	front of	-29.44
300,000	30.74	moment	-28.05
11	29.10	ceo of	-27.20
Mom	28.91	mine	-25.82
one day	28.81	companies	-25.23
Thank	27.69	developed	-25.19
400,000	27.24	wants to	-24.39
Thinking	26.81	name	-24.31
Bigger	26.39	an investment	-24.11
Total	26.29	body	-24.03
some samples	25.94	test	-23.83
90	25.71	30,000	-23.77
Millions	25.60	coming	-23.48
Charge	25.52	exciting	-23.24

Note. Terms come from the custom authenticity measure, and beta weights indicate the strength of association between the term and perceptions of authenticity. Therefore, the word *thousands* is most strongly associated with receiving a Shark Tank investment in the positive direction and *lot* is most strongly associated with receiving a Shark Tank investment in the negative direction.

Studies 4a-c: Verbal Authenticity Predicts Social Engagements

Method

Tweets from 53 Republican political figures (Study 4a), 15 journalists (Study 4b), and 3 news outlets (Study 4c) were obtained from prior work (Markowitz & Shulman, 2021). There was no overlap in empirical aims across papers, and the relationship between authenticity and engagement was not previously examined. Engagements were defined as the standardized sum of natural log transformed likes and retweets. Our hypothesis and analytic plans were preregistered for political Tweets (https://osf.io/y3mpx/?view_only=8205647eaeba4cd5abdf50c6b22b4953), though the other samples were exploratory. Our expanded rationale for selecting these data can be found in the online supplement.

Results

Modeling approaches were identical to prior work (Markowitz & Shulman, 2021). Authenticity was positively associated with social engagements for political Tweets, B = 4.311e-04, SE = 6.673e-05, t = 6.46, p < .001, R^2c = 0.58, 95% CI [0.0003, 0.0006]. The effect was marginally significant for journalist Tweets (p = .091) but not significant for news Tweets (p = .802), though all coefficients were positive. Social media posts generally received more engagements when the verbal output was written authentically compared to inauthentically.

Meta-Analysis

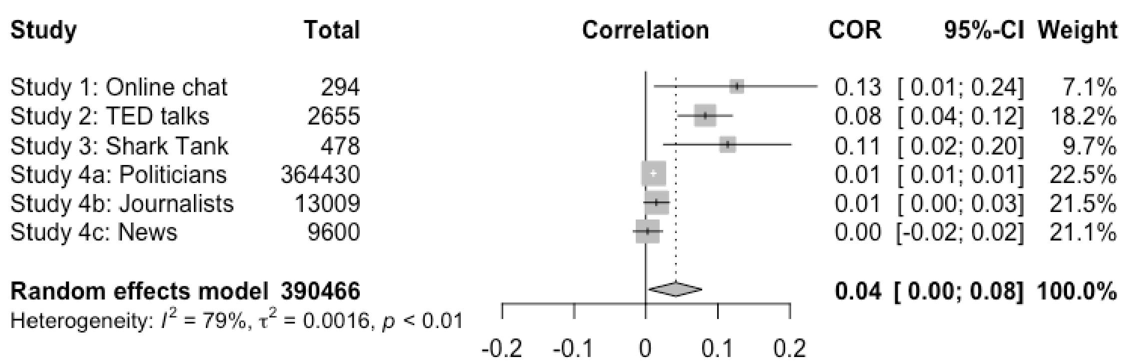
We estimated the effect size across all studies using a random effects model. As shown in Figure 1, the effect is small but significant (*corr* = .041, z = 2.18, p = .029; 95% CI [0.0042; 0.0781]). Please see the online supplement for backend details.

We suspect the small overall effect obtained from this meta-analysis is a result of LIWC's simple word-counting approach and the large amount of heterogeneity across samples.

As argued by Biber et al. (2007), communicators have discourse communities that they subscribe to and often shift their language style to meet the expectations of such communities (e.g., the audience of a politician is quite different than the audience of a TED talk, which likely impacts one's communication patterns). Such shifting contextual elements might be a major contributor to heterogeneity and differences in effect sizes across studies. We expect that if samples with contextual consistencies were analyzed (e.g., audiences, speakers, goals of the communicators), the effect size estimate would be larger.

Figure 1

Meta-Analysis Across Studies and Samples



General Discussion

The evidence from four studies suggests verbal authenticity predicts positive perceptions of a target across diverse settings. That is, verbal authenticity is linked to positive interpersonal perceptions (Study 1), social engagement (Studies 2 and 4a-c), and financial investments (Study 3). Human coding and natural language processing support the idea that verbal authenticity can relate to positive impressions of a target and meaningful behavioral outcomes, such as receiving money from investors. Taken together, our findings suggest authenticity is not only transmitted via natural language but also has important psychosocial consequences.

Theoretical and Applied Contributions

As compared to most assessments of authenticity via self-report, our investigations demonstrate language as a behavioral vehicle to communicate authenticity cues in interpersonal, social, and entrepreneurial settings: Authenticity is revealed by what people say and how they say it. The amount of verbal behavior appraised in each case was limited, yet people still formed positive perceptions of a target when communicators spoke authentically as compared to inauthentically. Thus, the data needed to form judgments about one's conversation partner, online video or text, and investment pitch required little elaboration. Crucially, our data suggests authenticity can be captured by top-down dictionaries that include style words (e.g., pronouns) and bottom-up, custom measures of mostly content words (Study 3). Therefore, we expand the theory of what authenticity means, what it can be indicated by (e.g., content and style), and how it relates to person perception. Relatedly, we also provide a constellation of supporting evidence (e.g., a validated verbal authenticity measure, a custom authenticity measure, human perceptions) to conclude that authenticity can lead to positive interpersonal, social, and entrepreneurial outcomes. The settings and coding procedures used to analyze these data were vastly different, yet they all support these positive effects of authenticity.

Our work also has practical implications outside of academic research. Corporations and public leaders (e.g., politicians) often strive to appear authentic (McShane & Cunningham, 2012), since this quality is associated with positive perceptions of their brand. Effective strategies for achieving this goal, however, are highly variable, context-dependent, and often nonverbal. Testing the value of authenticity and its effects are also rare. We offer language as a lens to clarify how authenticity is associated with positive perceptions of a target. Authentic language has both direct and indirect consequences for impression formation. Direct consequences include positive perceptions of a partner and receiving money from investors;

indirect consequences include increased social engagement online. We believe both outcomes are important to measure, as the wealth of recorded text data increases and is associated with meaningful outcomes, such as money and social interest.

Limitations and Future Directions

Some of the studies in this package did not rely on language alone, however. As noted, nonverbal cues might impact perceptions of authenticity in the data we examined (e.g., TED talks). Therefore, while Studies 4a-c concluded that authentic language in isolation can associate with positive social consequences, other studies contained a mixture of verbal and nonverbal cues that might additively or independently impact perceptions of a target. Finally, some of the effects in this paper were only robust after including covariates, and the meta-analysis revealed the effect is reliable, though generally small. Future work should further investigate the role of such covariates in analyses of authenticity and use more sophisticated word-counting approaches to evaluate the relationships in this paper.

References

Ambady, N., Krabbenhoft, M. A., & Hogan, D. (2006). The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology*, *16*(1), 4–13. https://doi.org/10.1207/s15327663jcp1601_2

Archer, L. (2008). Younger academics' constructions of "authenticity", "success" and professional identity. In *Studies in Higher Education* (Vol. 33, Issue 4, pp. 385–403). https://doi.org/10.1080/03075070802211729

Arenas, E. S. (2018). Exploring pornography in Widad Benmoussa's poetry using LIWC and corpus tools. *Sexuality and Culture*, *22*(4), 1094–1111. https://doi.org/10.1007/s12119-018-9513-7

Árnason, V. (1994). Towards authentic conversations. Authenticity in the patient-professional relationship. *Theoretical Medicine*, *15*(3), 227–242. https://doi.org/10.1007/BF01313339

Biber, D., Connor, U., & Upton, T. A. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. John Benjamins Publishing Company.

Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. *Proceedings of the Biennial GSCL Conference*, 31–40.

Boyd, R. L., Blackburn, K. G., & Pennebaker, J. W. (2020). The narrative arc: Revealing core narrative structures through text analysis. *Science Advances*, *6*(32), eaba2196. https://doi.org/10.1126/sciadv.aba2196

Boyd, R. L., & Schwartz, H. A. (2021). Natural language analysis and the psychology of verbal behavior: The past, present, and future states of the field. *Journal of Language and Social Psychology*, *40*(1), 21–41. https://doi.org/10.1177/0261927X20967028

Bulkeley, K., & Graves, M. (2018). Using the LIWC program to study dreams. *Dreaming*, *28*(1), 43–58. https://doi.org/10.1037/drm0000071

Carrère, S., & Gottman, J. M. (1999). Predicting divorce among newlyweds from the first three minutes of a marital conflict discussion. *Family Process*, *38*(3), 293–301. https://doi.org/10.1111/j.1545-5300.1999.00293.x

Cialdini, R. B. (2006). *Influence: The psychology of persuasion*. Harper Business.

Dalvean, M. (2017). Changes in the style and content of Australian election campaign speeches from 1901 to 2016: A computational linguistic analysis. *ICAME Journal*, *41*(1), 5–30. https://doi.org/10.1515/icame-2017-0001

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, *27*(2), 67–73. https://doi.org/10.1177/0963721417738825

Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N., & Qian, J. (2021). *glmnet: Lasso and elastic-net regularized generalized linear models* (4.1-1).

Gecas, V. (1991). The self-concept as a basis for a theory of motivation. In J. A. Howard & P. L. Callero (Eds.), *The self-society dynamic: Cognition, emotion, and action* (pp. 171–187). Cambridge University Press.

Gino, F., Kouchaki, M., & Galinsky, A. D. (2015). The moral virtue of authenticity: How inauthenticity produces feelings of immorality and impurity. *Psychological Science*, *26*(7), 983–996. https://doi.org/10.1177/0956797615575277

Grandey, A. A., Fisk, G. M., Mattila, A. S., Jansen, K. J., & Sideman, L. A. (2005). Is “service with a smile” enough? Authenticity of positive displays during service encounters. *Organizational Behavior and Human Decision Processes*, *96*(1), 38–55. https://doi.org/10.1016/j.obhdp.2004.08.002

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science, 22*(1), 39–44. https://doi.org/10.1177/0956797610392928

Jordan, K. N., Pennebaker, J. W., & Ehrig, C. (2018). The 2016 U.S. presidential candidates and how people tweeted about them. *SAGE Open*, *8*(3), 215824401879121. https://doi.org/10.1177/2158244018791218

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. In *Advances in Experimental Social Psychology* (Vol. 38, pp. 283–357). https://doi.org/10.1016/S0065-2601(06)38006-9

Kessler, G., Kelly, M., Rizzo, S., Shapiro, L., & Dominguez, L. (2021). *A term of untruths*. The Washington Post. https://www.washingtonpost.com/politics/interactive/2021/timeline-trump-claims-as-president/

Kifer, Y., Heller, D., Perunovic, W. Q. E., & Galinsky, A. D. (2013). The good life of the powerful: The experience of power and authenticity enhances subjective well-being. *Psychological Science*, *24*(3), 280–288. https://doi.org/10.1177/0956797612450891

Krumhuber, E., Manstead, A. S. R., & Kappas, A. (2007). Temporal aspects of facial displays in person and expression perception: The effects of smile dynamics, head-tilt, and gender. *Journal of Nonverbal Behavior*, *31*(1), 39–56. https://doi.org/10.1007/s10919-006-0019-x

Lehman, D. W., O’Connor, K., Kovács, B., & Newman, G. E. (2019). Authenticity. *Academy of Management Annals*, *13*(1), 1–42. https://doi.org/10.5465/ANNALS.2017.0047

Levin, G. (2019). *“Shark Tank”: All your burning questions, answered*. USA TODAY. https://www.usatoday.com/story/life/tv/2019/01/04/we-answer-your-shark-tank-burning-questions/2463620002/

Markowitz, D. M., & Griffin, D. J. (2020). When context matters: How false, truthful, and genre-related communication styles are revealed in language. *Psychology, Crime & Law*, *26*(3), 287–310. https://doi.org/10.1080/1068316X.2019.1652751

Markowitz, D. M., & Shulman, H. C. (2021). The predictive utility of word familiarity for online engagements and funding. *Proceedings of the National Academy of Sciences*, *118*(18), e2026045118. https://doi.org/10.1073/pnas.2026045118

McNeish, D. M. (2015). Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. *Multivariate Behavioral Research*, *50*(5), 471–484. https://doi.org/10.1080/00273171.2015.1036965

McShane, L., & Cunningham, P. (2012). To thine own self be true? Employees’ judgments of the authenticity of their organization’s corporate social responsibility program. *Journal of Business Ethics*, *108*(1), 81–100. https://doi.org/10.1007/s10551-011-1064-x

Meier, T., Boyd, R. L., Mehl, M. R., Milek, A., Pennebaker, J. W., Martin, M., Wolf, M., & Horn, A. B. (2020a). (Not) lost in translation: Psychological adaptation occurs during speech translation. *Social Psychological and Personality Science*, *12*(1), 131–142. https://doi.org/10.1177/1948550619899258

Meier, T., Boyd, R. L., Mehl, M. R., Milek, A., Pennebaker, J. W., Martin, M., Wolf, M., & Horn, A. B. (2020b). Stereotyping in the digital age: Male language is “ingenious”, female language is “beautiful” – and popular. *PLOS ONE*, *15*(12), e0243637. https://doi.org/10.1371/journal.pone.0243637

Montgomery, M. (2017). Post-truth politics? Authenticity, populism and the electoral discourses of Donald Trump. *Journal of Language and Politics*, *16*(4), 619–639. https://doi.org/10.1075/jlp.17023.mon

Moulard, J. G., Garrity, C. P., & Rice, D. H. (2015). What makes a human brand authentic? Identifying the antecedents of celebrity authenticity. *Psychology & Marketing*, *32*(2), 173–186. https://doi.org/10.1002/mar.20771

Naim, I., Tanveer, M. I., Gildea, D., & Hoque, M. E. (2018). Automated analysis and prediction of job interview performance. *IEEE Transactions on Affective Computing*, *9*(2), 191–204. https://doi.org/10.1109/TAFFC.2016.2614299

Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. Bloomsbury Press.

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic Inquiry and Word Count: LIWC2015*. Pennebaker Conglomerates.

Portal, S., Abratt, R., & Bendixen, M. (2019). The role of brand authenticity in developing brand trust. *Journal of Strategic Marketing*, *27*(8), 714–729. https://doi.org/10.1080/0965254X.2018.1466828

Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological Science*, *26*(6), 877–891. https://doi.org/10.1177/0956797615572906

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E. P., & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLoS ONE*, *8*(9), e73791. https://doi.org/10.1371/journal.pone.0073791

Sell, J., & Farreras, I. G. (2017). LIWC-ing at a century of introductory college textbooks: Have the sentiments changed? *Procedia Computer Science*, *118*, 108–112. https://doi.org/10.1016/j.procs.2017.11.151

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, *35*(9), 656–666. https://doi.org/10.1037/0022-3514.35.9.656

Sundar, S. S. (2007). Social psychology of interactivity in human-website interaction. In A. N. Joinson, K. Y. A. McKenna, T. Postmes, & U.-D. Reipes (Eds.), *The Oxford handbook of internet psychology*. Oxford University Press.

Tahmasbi, N., & Rastegari, E. (2018). A Socio-contextual approach in automated detection of cyberbullying. *Proceedings of the 51st Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2018.269

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin Books.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*(6), 813–833. https://doi.org/10.1521/soco.2009.27.6.813

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x