How High Arousal Language Shapes Micro versus Macro Influencers' Impact

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Cascio Rizzo, L., Villarroel Ordenes, F.J., Pozharliev, R., De Angelis, M., Costabile, M. (2024). How High Arousal Language Shapes Micro versus Macro Influencers' Impact. JOURNAL OF MARKETING, 88(4), 107-128 [10.1177/00222429231207636].

Availability:
This version is available at: https://hdl.handle.net/11585/981468 since: 2024-09-06

Published:
DOI: http://doi.org/10.1177/00222429231207636

(Article begins on next page)

DOI: 10.1177/00222429231207636

JOURNAL of Marketing

How Aroused Language Shapes Micro versus Macro Influencers' impact

Journal:	*Journal of Marketing*
Manuscript ID	JM-22-0437.R4
Manuscript Type:	Revised Submission
Research Topics:	Consumer Engagement, Influencer Marketing, Social Media, Persuasion Knowledge, Word of Mouth
Methods:	Multimethod, Natural Language Processing, Online Experiments, Image analysis , Audio Data analysis

How High Arousal Language Shapes Micro versus Macro Influencers' Impact

Author Note

Giovanni Luca Cascio Rizzo is a doctoral student in marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (glcasciorizzo@luiss.it). Francisco Villarroel Ordenes is an assistant professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (fvillarroel@luiss.it). Rumen Pozharliev is assistant professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (rpozharliev@luiss.it). Matteo De Angelis is a professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (mdeangelis@luiss.it). Michele Costabile is a professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (mcostabile@luiss.it). Please address correspondence to Giovanni Luca Cascio Rizzo. This article is based on the lead author's dissertation.

Acknowledgements

The authors acknowledge their collaboration on this research with Buzzoole S.p.a.. They also thank Ann Kronrod, Andrea Luangrath, Stephan Ludwig, and Stefano Puntoni for their valuable feedback, and participants at the EMAC 2023, Rotterdam School of Management lunchclub, and SIM doctoral consortium 2022 for helpful comments.

Data collection information

The field data for Study 1 and its follow-up study were collected in winter 2021 by the first author. The data for Studies 2, 3a, and 3b were collected by the fourth author in fall 2022 and spring 2023. The exploratory study was collected by the third author in spring 2023. Study 4 was collected by the second author in winter 2022. All experiments used Prolific participants located in the United States and were designed by the first author. The field data measures of Study 1 were developed by the first and second authors. Analysis for all studies was performed by the first author. The data are currently stored in a project directory on the Open Science Framework.

How High Arousal Language Shapes Micro versus Macro Influencers' Impact

Abstract

Influencer marketing is a popular strategy to connect with consumers. However, influencers' use of overly high arousal language in promoting products (e.g., "it's totally AMAZING!") has raised questions about their true motivations. This article investigates how high arousal language in micro versus macro influencers' sponsored posts might shape engagement. Six studies, combining automated text, image, video, and audio analyses of thousands of Instagram and TikTok posts with preregistered controlled experiments, demonstrate that high arousal language increases engagement with micro influencers, but it decreases engagement with macro influencers, seemingly because it makes micro (macro) influencers appear more (less) trustworthy. Yet the negative effect of arousal for macro influencers can be mitigated if their posts provide counterbalanced valence (e.g., both positive and negative assessments) or if they indicate an informative, rather than commercial, goal. These findings deepen understanding of how language arousal shapes consumer responses, reveal a psychological mechanism through which language arousal affects perceptions, and provide actionable insights for crafting more effective social media content.

Keywords: language arousal, micro and macro influencers, unstructured data analysis, engagement, social media, persuasion knowledge

Influencer marketing has become a prominent strategy; more than 90% of brands enlist micro or macro influencers to connect with consumers and achieve a variety of marketing goals, from creating awareness to increasing sales (Kupfer et al. 2018; Leung, Gu, and Palmatier 2022; Santora 2022). Despite the popularity and relevance of influencer marketing, its effectiveness varies, largely depending on the engagement that influencers attract on social media (Hughes, Swaminathan, and Brooks 2019). Some posts garner engagement, stimulating interest in the products endorsed, while others do not. So, what makes some sponsored posts more engaging than others?

One possibility is that engagement depends on how trustworthy influencers seem. Companies rely on influencers because consumers are wary of advertising (Leung et al. 2022). Yet growing consumer awareness that influencers get paid to promote products may raise questions about influencers' motives (Cascio Rizzo et al. 2023). Moreover, anecdotal evidence suggests that high arousal claims (e.g., "it's totally amazing!") appear overly commercial, leading consumers to question the trustworthiness of the influencer and therefore engage less with the content (Michaeloudis 2021).

This research examines whether subtle shifts in language arousal in influencers' posts (e.g., "sensational," "hectic," "shocking") might shape consumer engagement by affecting how trustworthy the influencers seem. Prior research establishes that high arousal language, or the extent to which the source appears energized by the topic being described (Yin, Bond, and Zhang 2017), can increase some forms of engagement (Berger and Milkman 2012; Herhausen et al. 2019). For influencers though, we posit that these effects depend on *who* is talking, namely, a micro (small-scale audience) or macro (massive reach) influencer. In detail, we suggest that high arousal language might increase engagement with micro influencers but decrease engagement with macro influencers, because it makes the micro (macro) influencer seem more (less) trustworthy. Calling some shoes "sensational" rather than "nice," for example, might signify the genuine excitement of a micro influencer, which increases trustworthiness and engagement. But in the case of macro influencers, consumers might perceive it as an overt attempt to persuade, leading to decreased trustworthiness and engagement. Consistent with this suggestion, we also posit that when posts have an informative, rather than commercial, goal, consumers are less likely to suspect a persuasion attempt by macro influencers, so the negative effect of high arousal language could be attenuated.

Six studies, combining automated text, image, video, and audio analyses of thousands of influencers' sponsored posts with controlled experiments, test these possibilities. In turn, we make three theoretical, practical, and methodological contributions. Theoretically, we add to studies of how language arousal in online communications shapes consumer responses (e.g., Berger and Milkman 2012), by explicating how the effect of arousal depends on the source of the message. Drawing on persuasion knowledge theory (Friestad and Wright 1994), we demonstrate that consumers elaborate on both arousal and influencer type (micro vs. macro) to make inferences about influencer trustworthiness and decide whether to engage with the content or not. This analysis of the role of influencer type establishes an important boundary condition for language arousal's effects on engagement, while also contributing to research on influencers (e.g., Cascio Rizzo et al. 2023; Hughes, Swaminathan, and Brooks 2019; Karagür et al. 2022; Wies, Bleier, and Edeling 2023).

For practice, our results also offer clear implications. First, we provide recommendations for influencers regarding how they should compose their sponsored messages. High arousal language can boost micro influencers' effectiveness; conversely, it can backfire for macro influencers unless their posts aim to spread information (vs. purchase), include trustworthiness cues, or exhibit counterbalanced valence. Our field data suggest that increasing arousal by 10% (e.g., from "great" to "superb") is associated with a

5.4% increase in engagement (i.e., likes and comments) for micro influencers but an 8.4% decrease in engagement for macro influencers, on average. Second, we offer brands insights into when to hire micro or macro influencers, depending on the goal of the post. As this goal shifts from commercial to informative, engagement increases by 1.8% for macro influencers, on average.

Methodologically, we demonstrate the use of text analysis as a viable means to gain key consumer insights in social media marketing contexts (Berger et al. 2020). First, we extend previous arousal operationalizations based on words (Berger and Milkman 2012), by incorporating paralanguage (i.e., emojis, capital letters, and punctuation; Luangrath, Xu, and Wang 2023), which provide a deeper conceptual and empirical perspective on language arousal in social media. Second, we advance prior investigations of influencer effectiveness and campaign goals (Hughes, Swaminathan, and Brooks 2019) by establishing dictionaries for informative and commercial goals (Humphreys, Isaac, and Wang 2021), which in turn can predict when micro or macro influencers will be more effective. Third, we use Wordify (Hovy, Melumad, and Inman 2021), an automated text analysis tool, to construct a dictionary of trustworthiness words (e.g., “help,” “learn,” “experience”) and demonstrate how arousal’s negative effect on engagement with macro influencers can be attenuated by increased uses of words that signal their trustworthiness.

Conceptual Background

Language Arousal

Prior arousal research offers an activation account, based in the autonomic nervous system. Low arousal (or deactivation) is manifested as relaxation, whereas high arousal (or activation) is manifested as activity (Berger 2011). In a communication context, the level of arousal reflects the extent to which people are energized by a description that uses a particular wording (Yin, Bond, and Zhang 2017). Saying something like “it’s sensational” rather than "it's nice," for example, conveys higher arousal. It suggests the source of the description is strongly energized by that topic.[1]

Consumer psychology literature also relates language arousal to language intensity, defined as the extent to which a message deviates from neutrality (e.g., "detested" vs. "dislike"; Pogacar, Shrum, and Lowrey 2018). Arousal and linguistic extremity also likely covary, such that the use of stylistic markers can increase the perceived extremity of a message's position (Craig and Blankenship 2011). Studies of paralanguage (Luangrath, Xu, and Wang 2023) further establish that capitalization, exclamation marks, and emojis can emphasize and intensify messages in social media communication. Text communication lacks certain verbal and nonverbal components, so textual paralanguage can alter other linguistic components, such as arousal (Moore and Lafreniere 2020). Thus, "GREAT!" evokes a higher level of arousal than "great." To conceptualize arousal, we thus focus on the level of activation established by the level of arousal in both words and paralanguage.

Language arousal might boost engagement. For example, news articles that evoke high arousal emotions (e.g., awe, anger) go viral more than ones evoking low arousal (e.g., sadness) (Berger and Milkman 2012). Online firestorms in brand communities are more likely to arise from negative customer posts that are high in arousal (e.g., "this is so frustrating!") rather than low in arousal (e.g., "this is disappointing") (Herhausen et al. 2019). However, high arousal content can have null effects, such as when it is posted in the evening (Kanuri, Chen, and Sridhar 2018) or for YouTube content (Tellis et al. 2019). Alternatively, in some cases, heightened arousal may backfire and decrease sharing, such as if the topic discussed does not reflect on the sharer (Weingarten and Berger 2017), if readers regard it as a signal of irrationality (Yin, Bond, and Zhang 2017), or if it makes messages seem overly

[1] Note that terms may have the same valence—both "sensational" and "nice" imply positive perceptions—but evoke varying arousal levels (Berger and Milkman 2012). In detail, psycholinguistic scales indicate that "sensational" and "nice" have similar positive valence (M = .94 vs. .93), but the former scores substantially higher on arousal (M = .83 vs. .44; Mohammad 2018).

commercial and exaggerated (Haan and Berkey 2002). When shared by influencers, heightened arousal could lead consumers to believe that their intent is manipulative, resulting in mistrust (Eisend and Tarrahi 2022) and reduced engagement.

Among this wealth of insights into the effects of language arousal on consumer behavior, we know of no studies of how language arousal shapes engagement with influencer-sponsored content. Furthermore, even if some research has identified boundary conditions of the effects of language arousal (e.g., Kanuri, Chen, and Sridhar 2018; Yin, Bond, and Zhang 2017), the contingent influence of *who* is talking remains unknown.

Influencer Marketing and Consumer Engagement

Reflecting the rise of social media, recent marketing approaches increasingly depend on influencer content produced on social media platforms. Influencer effectiveness largely depends on the level of engagement that influencers attract. Gaining more engagement (e.g., likes, comments) signals that the post has resonated with consumers, which should increase sales (Kumar et al. 2016; Liadeli, Sotgiu, and Verlegh 2023). Many brands even measure the return on their investments in influencer marketing according to achieved engagement (Santora 2022).

Accordingly, research has begun to investigate what boosts consumer engagement with influencer content. When influencers create posts by themselves, engagement increases, because those influencers seem more original (Leung et al. 2022). If they follow fewer others, greater engagement also results, because it signals that the influencer is more autonomous (Valsesia, Proserpio, and Nunes 2020). Other identified determinants of engagement include the extent to which influencers share personal experiences, interact with close others, and display expertise (Cascio Rizzo et al. 2023; Chen, Yan, and Smith 2022; Chung, Ding, and Karla 2023; Hughes, Swaminathan, and Brooks 2019). But few studies address the language that influencers use when posting about a product, beyond general indications that message valence can affect engagement (Gerrath and Usrey 2021; Leung et al. 2022). In addition to being categorized as positive or negative, messages also vary in the degree of arousal they signal. If influencers' general goal is to stimulate consumers to consider or buy a product, they already are unlikely to use negative terms, but they might use language that differs in its level of arousal, to signal excitement about a product. But might the effects of high arousal depend on whether the poster is a micro or a macro influencer? And if so, why?

Micro and Macro Influencers. Micro influencers, with relatively few followers, share their opinions of products and post content about activities they commonly perform. Macro influencers instead have hundreds of thousands of followers, casting a wide net with their messages (Pozharliev, Rossi, and De Angelis 2022). Recent research details their distinct impacts on consumer behavior. For example, Karagür et al. (2022) show that people tend to perceive posts shared by macro (vs. micro) influencers as advertising, which reduces their engagement. Wies, Bleier, and Edeling (2023) find an inverted U-shaped effect, such that engagement increases, then decreases, as influencer follower count rises, due to perceived tie strength. But other studies reveal how more followers can boost engagement (Leung et al. 2022), by providing signals of popularity, status, and reputation.

To explain such mixed results, we propose that the language that micro and macro influencers use is relevant, a notion that resonates with Cascio Rizzo et al.'s (2023) recommendation that macro influencers should use sensory language to increase engagement. Language arousal similarly might affect engagement differently for micro versus macro influencers. That is, high arousal language should increase engagement with micro influencers but decrease engagement with macro influencers. We suggest this possibility based on research on persuasion knowledge and trust.

Persuasion Knowledge and Trust. Trust is a key driver of consumer engagement (Leung, Gu, and Palmatier 2022). Just as trust increases persuasion (Packard, Gershoff, and

Wooten 2016), consumers are more likely to like and comment on posts when they feel like the influencer is more trustworthy (i.e., sincere or motivated to provide accurate information; Pornpitakpan 2004). The persuasion knowledge model emphasizes the central role of trust for social influence (Friestad and Wright 1994), by detailing that when people recognize manipulativeness in a persuasion attempt, they infer low trustworthiness of the source, and then react negatively (Campbell and Kirmani 2000). Disclosing that a post is paid, for example, may alert followers to the message's commercial intent, leading to decreased trust (e.g., Boerman, Willemsen and Van Der Aa 2017). Beyond such straightforward disclosures, which have become even compulsory on some social media platforms, language and follower counts also might activate persuasion knowledge (Hughes, Swaminathan, and Brooks 2019; Karagür et al. 2022). Therefore, when consumers encounter arousing language in a sponsored post (e.g., "what an incredible product!!"), they might react differently if it comes from a micro versus macro influencer, based on their perceptions of the genuineness of the expressed arousal. If they find genuine intent, consumers do not perceive a persuasion attempt (i.e., persuasion knowledge is not activated), so they tend to respond positively (Berger and Milkman 2012). If they believe the post represents advertising, however, their persuasion knowledge gets activated, so they may attribute manipulative intent to the influencer (Hughes, Swaminathan, and Brooks 2019).

We posit that a micro influencer's use of high arousal language could make them seem more trustworthy, because with their relatively small follower base, these influencers do not activate consumer persuasion knowledge. Consumers typically see such influencers as everyday users, whose content seems less like advertising (Hotmart 2022). Thus, if a micro influencer says something like, "this protein shake is AMAZING!," their language arousal suggests genuine excitement (Berger and Milkman 2012) and implies they want to share a great product with others. Believing that someone is genuinely excited about what they are talking about should increase trustworthiness (Pogacar, Shrum, and Lowrey 2018). After all, an influencer actually excited about a product should seem less driven by monetary interests (i.e., whether the company paid them) and more inspired by their sincere beliefs.

In contrast, a macro influencer's use of high arousal language might activate persuasion knowledge, in that consumers assume influencers with lots of followers may be more likely to get paid to say positive things about products (Hatton 2018). Because those influencers appear motivated mainly by economic incentives, consumers regard their posts as forms of advertising (Karagür et al. 2022). High arousal language in advertising sparks persuasion knowledge (Haan and Berkey 2002); such a post appears to be trying too hard to convince consumers to buy (Yin, Bond, and Zhang 2017). Believing that an influencer has such a manipulative intent then should decrease consumers' sense of their trustworthiness. Formally,

> **H_1**: High arousal increases consumer engagement with micro influencers but decreases consumer engagement with macro influencers.
>
> **H_2**: The positive (negative) effect of high arousal language on consumer engagement with micro (macro) influencers is driven by persuasion knowledge and trust.

Influencers' Post Goals. Influencer marketing campaigns typically pursue two main purposes: increasing awareness or encouraging trial (Hughes, Swaminathan, and Brooks 2019). From a language perspective, posts aimed at increasing awareness tend to signal an informative goal, whereas posts encouraging trial indicate commercial goals (e.g., Villarroel Ordenes et al. 2019). These two distinct goals reflect the beginning and the end of the consumer decision journey (Colicev et al. 2018; Humphreys, Isaac, and Wang 2021) and affect the activation of consumer persuasion knowledge differently (Williams, Fitzsimons, and Block 2004). Compared with commercial posts, informative posts do not evoke strong perceptions of advertising motives, so they are less likely to activate persuasion knowledge (Hughes, Swaminathan, and Brooks 2019). A macro influencer who posts with an informative goal thus might mitigate the negative effects of high arousal language. Posting "*Check out* [vs. *Buy*] this amazing shake!" for example indicates to consumers that the goal is just to make people aware of the product's existence. The high arousal language in such a post then should appear less instrumental and lessen the potential negative effect. Formally,

H$_3$: The negative effect of high arousal language on consumer engagement with macro influencers is mitigated if posts have an informative goal.

The Current Research

Taken together, we suggest that using high arousal language increases (decreases) engagement for micro (macro) influencers. Further, this effect is driven by influencer trustworthiness. High arousal language increases beliefs that the micro (macro) influencer's intent is less (more) persuasive, which increases (decreases) influencer trustworthiness and boosts (reduces) engagement. To test these predictions, we adopt a multimethod approach.

Study 1 provides an initial field test on Instagram, examining whether high arousal language increases engagement with micro influencers but decreases it with macro influencers. Furthermore, it examines whether the negative effect of high arousal language on macro influencers can be mitigated by posts that clearly seek to generate information, rather than encourage purchase. With this study, we also offer an initial exploration of the underlying role of influencer trustworthiness. In a follow-up to Study 1, we test the generalizability of the effects to a different platform and communication modality (i.e., spoken language in TikTok videos).

To establish the causal impact of language arousal and the underlying process, we then conduct four preregistered experiments. In Study 2, we manipulate language arousal and influencer type (micro vs. macro) to determine whether high arousal language increases (decreases) engagement with the micro (macro) influencer's content. Study 3a tests the underlying role of trustworthiness through mediation. We determine whether high arousal language increases (decreases) engagement for micro (macro) influencers by causing consumers to believe that the influencer's intent is less (more) persuasive, which increases (decreases) influencer trustworthiness. Both Studies 3b and 4 provide tests of the moderating process for macro influencers. If the effects of high arousal language are driven by trust, as we suggest, they also should be mitigated when macro influencers' trustworthiness is less questionable, such as when their posts feature counterbalanced valence (Study 3b), or the post goal is informative rather than commercial (Study 4).

We also acknowledge that consumers might trust micro influencers more than macro influencers, potentially polarizing the effects of high arousal language on trustworthiness (i.e., high arousal language makes trusted micro influencers more trustworthy and distrusted macro influencers less trustworthy). An exploratory study (Web Appendix D) casts doubt on this explanation, by showing that consumers generally trust micro and macro influencers equally. Instead, people appear to exhibit latent differences in their trust of micro versus macro influencers, which get activated by language arousal.[2]

Study 1: Language Arousal in the Field

Study 1 investigates whether high arousal language increases engagement with micro influencers but decreases engagement with macro influencers (H_1). It also tests whether arousal's negative effect for macro influencers can be mitigated by posts that aim to be informative rather than commercial (H_3). We establish the validity of text-based measures of language arousal and post goals (Table 1) and conduct several robustness checks (see Table 5, subsequently). In addition, with some ancillary analyses, we explore the underlying role of influencer trustworthiness (H_2).

---

[2] Although our study relies on prior research into influencer language and engagement, it differs in critical ways. For example, Cascio Rizzo et al. (2023) examine how sensory language (i.e., words that engage the senses such as "tasty" or "crunchy") leads consumers to believe that macro influencers have actually used the product, which increases engagement. In contrast, we focus on high arousal language (i.e., words that reflect the level of activation such as "amazing" or "exciting") to learn if it might decrease engagement with macro influencers by activating persuasion knowledge.

Data and Measures

We collaborated with a large influencer marketing agency to acquire a sample of 20,923 sponsored posts on Instagram from 1,376 influencers, posted between October 13, 2019, and October 30, 2021. The posts cover both products and services from 18 different industries, such as beauty, food, gaming, and travel. Descriptive statistics and correlations are in Web Appendix B (Tables WB1 and WB2). Table 2 contains a full list of the measures, their operationalizations, and sources, and rationales for the controls.

Engagement. The measure of engagement equals the total number of likes and comments a post receives (e.g., Cascio Rizzo et al. 2023; Herhausen et al. 2019).[3] On average, posts received 3,432 likes (SD = 6,023) and 105 comments (SD = 282).

Arousal. Most measures of arousal focus on words (e.g., Berger and Milkman 2012; Kuperman et al. 2014), but social media content features the active use of paralanguage (e.g., emojis, punctuation, capitalization; Luangrath, Peck, and Barger 2017). For a more accurate operationalization, we thus adopt a two-step approach that combines words and paralanguage, together with a thorough validation. First, we used the Mohammad's (2018) valence, arousal, and dominance (VAD) dictionary, which has been used widely to quantify emotional features in social sciences (e.g., De Deyne et al. 2021; Felbermayr and Nanopoulos 2016). It provides human ratings of valence, arousal, and dominance for more than 20,000 English words, using scores that range from 0 to 1. For our analysis, we used the arousal dimension to measure the level of arousal in influencer posts. Its correlation with Whissel's (2009) Dictionary of Affect in Language (DAL; $r = .69$), provides concurrent validity.

Second, we account for arousal conveyed by emojis, capitalization, and exclamation marks, because of their potential to intensify messages in social media (Luangrath, Xu, and

[3] We use the sum of likes and comments for three main reasons. First, companies consider such composite measures to select influencers (Influencer Marketing Hub 2022). Second, our focus is on the implications of influencers' language on the volume of engagement generated, not the type (Lee, Hosanagar, and Nair 2018). Third, using just likes (wholly positive outcome) as a dependent variable produces the same results (Table 5).

Wang 2023). To determine emoji arousal, we used Kralj Novak et al.'s (2015) Emoji Sentiment Ranking, a sentiment dictionary of the 751 most frequent emojis on social media, annotated by humans with three ordered values of sentiment (i.e., negative, neutral, or positive), each ranging from 0 to 1. Because our focus is on the degree of emotional intensity, not its direction (i.e., valence), we computed emoji arousal as the absolute sum of an emoji's positive and negative scores (Kuppens et al. 2013). Our data include 1,277 unique emojis, of which 781 lack an arousal score, so for those, we undertook manual annotation.[4] The final emoji lexicon encompasses emojis with both higher arousal (e.g., □, 🎉, 🙌□) and lower arousal (e.g., ☺□, ✅, 🙈). Table WB3 in Web Appendix B lists the most frequent emojis (i.e., top 10% of occurrences, with a minimum threshold of 10) and their corresponding arousal scores. Two research assistants (r = .73) rated these emojis on arousal (from 0 to 1) separately. In addition, we account for capital letters and exclamation marks, by using a weighting approach to increase the arousal level of posts that include them (e.g., Villarroel Ordenes et al. 2017). Considering the lack of evidence about the exact weights to assign to capitalization and exclamation marks, we used a sensitivity approach and tried several weights (e.g., Ananthakrishnan, Proserpio and Sharma 2023).

On the basis of both words and paralanguage, we operationalize arousal as:

$$\boldsymbol{Post\ Arousal_i} = \frac{[\sum_1^n AWW_j + \sum_1^n AEW_j + 1.5 * \sum_1^n C\_WA_j] * (1 + 0.2 * EM)}{WC_i + EC_i} \tag{1}$$

where arousal in post i equals (1) the sum of all arousal word weights (AWW) in lowercase, plus (2) the sum of all arousal emoji weights (AEW), plus (3) the sum of all capitalized AWW multiplied by 1.5, times (4) a value of 1.2 if at least one exclamation mark is present (EM), all divided by the sum of word count (WC) and emoji count (EC) of post i. To

[4] We determined that 256 emoji are simply color tone variations of entries in Kralj Novak et al.'s (2015) dictionary (e.g., existing: 🙌, color tone variations: 🙌□, 🙌□), so we assigned them the same arousal score. Of the remaining 525 emoji, only 108 are unique (e.g., □♀□, 🔍), because the rest feature other color tone variations (e.g., □□♀□) or very similar depictions (e.g., 🔎).

illustrate, "GOOD work! ☺" produces GOOD = .368×1.5, work = .596, and ☺ = .764, so it is equal to [.552 + .596 + .764] × 1.2 / 3 = .765. To ensure convergent validity, two research assistants (blinded to the hypotheses) rated a random sample of 2,000 posts on the level of language arousal (1 = not at all, 7 = very much; *r* = .64).[5] Our automated measure correlates with human perceptions of arousal (*r* = .67), confirming its construct validity.

Influencer Type. For influencer type, we follow recent literature (Lee and Junqué de Fortuny 2021) and classify all influencers with 10,000–100,000 followers as micro influencers, but influencers with 100,000–1,000,000 followers as macro influencers. In our data, 54% are micro influencers ($M_{followers}$ = 50,221; SD = 27,186), and the rest are macro influencers ($M_{followers}$ = 377,600; SD = 205,470). We use a dummy variable ("Macro" = 1 if macro, 0 if micro) to reflect this classification. Using different operationalizations and cutoff points produces similar results (see Table 5).

Informative Goal. To measure how informative (vs. commercial) a post goal is, we created two dictionaries, one including 43 informative words (e.g., "check," "discover,"), and another listing 48 commercial words (e.g., "discount," "offer"; see "New Dictionaries for Informative and Commercial Post Goal" in Web Appendix B for the full list of words). Then, we computed the proportion of informative words to the total of informative and commercial words.[6] We name the resulting variable "informative," for which higher values indicate an informative (cf. commercial) post goal. We followed Humphreys and Wang's (2018) guide to develop the dictionaries. First, we drew the 100 most frequent terms from posts in our data set. We added synonyms, excluded homonyms, and ensured context specificity. Second, we computed an informative goal score at the post level (using the previously mentioned

[5] Coders were given a definition of arousal that read, "language that communicates the influencer is energized by what he/she is sponsoring. For example, the word 'sensational' is higher on arousal than 'great,' 'shocked' is higher in arousal than 'surprised,' and '☺' is higher in arousal than '□'."
[6] This procedure is required because posts can be used to stimulate both information and purchase.

formula).[7] Third, we asked two research assistants to rate a random sample of 250 posts on two questions, pertaining to the likelihood that the post had an informative or commercial goal (1 = not at all, 7 = very; $\alpha_{k,\text{ informative}}$ =.77, $\alpha_{k,\text{ commercial}}$ =.70). The automated measure closely matches the human ratings ($r_{\text{informative}}$ = .60; $r_{\text{commercial}}$ =.62), supporting construct validity. The two measures are weakly related (r = .14), confirming discriminant validity. The dictionary of commercial words also shows a positive and significant correlation with Jalali and Papatla's (2019) list of sales promotion words[8] (r = .33), indicating convergent validity. An influencer marketing manager supported the face validity of our two measures. To further ensure the validity of our construct, following Ludwig et al. (2022), we compared the coders' classification with the text-mined version. The results suggest an overall classification accuracy of .82 (Table WB4).

Control Variables. We include multiple control variables, to account for unobserved heterogeneity. First, the results might reflect the influence of the person posting the content, so we control for the influencer's characteristics: whether they have a verified account, the number of posts they have shared, and the product category in which they usually post. We also control for the level of influencer specialization, using word embeddings to compute the degree of semantic variation across all posts by each influencer (i.e., content variation; see "Content Variation" in Web Appendix B). Higher content variation scores indicate less specialization.[9] Second, we control for textual aspects: topics discussed; number of questions, mentions, hashtags, and emojis;[10] wordcount; text complexity, valence, and concreteness; and

[7] Note that 1,908 posts do not include any informative or commercial words, so we assigned them the average score.
[8] The word list includes "chance," "commercial," "free," "gift," "giveaway," "promo," "win" and "sale."
[9] We restricted our analysis to influencers who have posted at least three times, resulting in a final sample of 20,590 posts. Micro influencers arguably are just more specialized than macro influencers, such that their specialized expertise, rather than follower count, drives the effects. But we determined that micro and macro influencers do not differ in content variation (b = –.001; SE = .003; t = –.22; p = .825). We also obtain similar results when we operationalize specialization as the number of industries (1 to 3; data provided by the agency partner) about which an influencer posts.
[10] Table WB5 contains results for the paralanguage detected by PARA (Luangrath, Xu, and Wang 2023).

word familiarity. Third, we control for aspects of the images that appear in the post, including their type (image or video),[11] if it features a human face, color saturation, and dominance. With regard to visible faces, we further control for image emotionality, depending on the emotional state they depict (joy, sorrow, anger, or surprise), using the Google's Cloud Vision API (see "Image Emotionality" in Web Appendix B). Fourth, we include fixed effects for year, month, weekday, and time of the day, and we control for whether more than 24 hours occurred between two consecutive posts shared by the same influencer.

Table 1. Dictionary Validation for Language Arousal and Informative Goal

Type of Validity	**Validation Procedure**
Construct validity: Does the text represent the theoretical concept?	Following Humphreys and Wang (2018), for arousal, we used a top-down approach, combining Mohammad's (2018) VAD dictionary with paralanguage. For informative goal, we used a bottom-up approach, empirically guided by the most frequent words in our data.
Concurrent validity: Does the measurement of the constructs relate to other measurements?	Our measurement indicates concurrence with human ratings for both arousal ($r_{intercoder}$ = .64, r = .67) and informative/commercial goal (average r_{coders} = .79, average r = .61; classification accuracy = .82).
Convergent validity: Do multiple measurements of the construct all converge to the same concept?	The VAD arousal score correlates with the DAL arousal score (r = .69). The commercial word count at a post level correlates with Jalali and Papatla's (2019) list of sale promotional words (r = .33).
Discriminant validity: Does the measurement differentiate from measures of other constructs?	Our arousal measure does not relate to valence (r = .02, ns.). Our informative goal measure does not relate to arousal (r = .02; ns.). The informative and commercial goals measures are weakly related (r = .14)
Causal validity: Is the construct in the data set causally related to other constructs?	We include several controls in the model to rule out alternative explanations (e.g., influencer, text, image, other).
Predictive validity: Does the construct have the expected effects of a meaningful variable?	Across different measurement approaches, we confirm the theoretically derived relationship between arousal and informative goal with engagement for macro and micro influencers in the field.
Face validity: Does the construct measure what it claims to measure?	High arousal: "Sharing TONS of essentials to help you! ✿ this sweater is SO soft! ☺ @liketoknow.it #liketkit #ad @nordstrom." Low arousal: "Just out here making my own pizza dreams come true with this Spicy Broccoli Pizza with a quick homemade pizza crust using @fleischmannsyeast. It's really a wonder #ad." Informative goal: "Have you heard? □ Hi-C is BACK at @mcdonalds□ Wops World Out Now 🎶□🚀 #Ad #McDonalds." Commercial goal: "Holiday Gift from @manscaped 🎁. Use code "TREVOR20" to save 20% #sponsored #ad."
Robustness: Is more than one method used?	We replicate the focal relationships in controlled experimental settings, in which we manipulate arousal (Studies 2–4) and informative goal (Study 4).
Generalizability: Are results based on multiple data sets?	The relationship of arousal, influencer type, and engagement is replicated with two independent samples (i.e., Instagram posts and TikTok videos).

[11] About 11% of the visuals in our data are videos. Following Villarroel Ordenes et al. (2019), we extracted the first screenshot of the video, then dummy coded the post type variable (0 = video, 1 = image).

Table 2. Control Variables Operationalization and Sources

Variable	Operationalization	Source	Related Studies
if Verified	If the influencer has a verified account (dummy coded).	Company data	Verified accounts are associated with public figures and thus may increase engagement (Valsesia, Proserpio, and Nunes 2020).
# of Posts	Number of posts shared by the influencer.	Company data	Posting more frequently can make followers think the influencer provides fresh and up-to-date information, which may increase engagement (Leung et al. 2022).
Content Variation	Degree of semantic variation across all influencer's posts computed using Word2Vec.	Company data	Higher influencer expertise can boost engagement (Hughes, Swaminathan, and Brooks 2019).
Topics	Empath's 194 pre-built topics (Fast, Chen, and Bernstein 2016). Factorized into 64 overarching topics using varimax rotation.	Text mining	Conversation themes may drive some forms of engagement (Packard and Berger 2021).
# of Questions	Number of question marks.	Text mining	Questions increase the interactivity of a post (Villarroel Ordenes et al. 2019).
# of Hashtags	Number of hashtags per post.	Text mining	Hashtags may heighten post impressions and engagement (Stieglitz and Dang-Xuan 2013).
# of Mentions	Number of mentions per post.	Text mining	Mentions can increase views and thus engage users (Leung et al. 2022).
Word Count	Number of words per post.	Text mining	Longer posts may convey more information and thus increase engagement (Berger and Milkman 2012).
# of Emojis	Number of emojis per post.	Text mining	More emojis may lead to improved engagement (Luangrath, Xu, and Wang 2023).
Complexity	The level of text complexity computed using Flesch–Kincaid measure.	Text mining	Easy-to-read posts are more fluently processed and so may increase engagement (Berger and Milkman 2012).
Valence	Degree of text positivity using Mohammad's (2018) VAD lexicon (average across words).	Text mining	Positive content text may increase engagement (Berger and Milkman 2012).
Concreteness	Linguistic concreteness ratings from Paetzold and Specia (2016).	Text mining	More concrete language can suggest direct experience and increase engagement (Packard and Berger 2021).
Familiarity	Word familiarity ratings from Paetzold and Specia (2016).	Text mining	Familiar language is more fluent, which may increase engagement (Pancer et al. 2019).
if Image (vs. Video)	If the post features an image or a video (dummy coded).	Company data	Video content has a greater tendency to go viral than images (Borah et al. 2020).
if Face Present	If the image features a human face (dummy coded).	Image Mining	Faces may receive more attention and induce higher engagement (Li and Xie 2020).
Image Emotionality	Maximum of emotions scores per post.	Image Mining	Higher image emotionality may affect engagement (Li and Xie 2020).
Color Dominance	Sum of pixel percentage of top-three colors.	Image Mining	Higher color dominance may enhance attention and engagement (Li and Xie 2020).
Color Saturation	Level of saturation cross pixel of the image.	Image Mining	Higher saturation may affect engagement (Li and Xie 2020).
Time difference	Whether more than 24 hours occurred between two consecutive posts (dummy coded).	Company data	Lesser temporal distance between two consecutive posts may reduce engagement (Villarroel Ordenes et al. 2019).

Method

Finally, we examined the joint effect of language arousal and macro (vs. micro) influencers on engagement. The dependent variable is a count variable and overdispersed (p < .001, likelihood ratio test), so we use negative binomial regression. Because the variables rely on divergent scales, we standardize all the continuous variables.

Results

Arousal. As predicted, we find a significant interaction effect of arousal × macro (vs. micro) influencer on engagement (incident rate ratio [IRR] = .906; SE = .011; t = –8.03; p < .001; Table 3, column 1). Even after accounting for the control variables, we continue to find a significant effect of arousal × macro (vs. micro) influencer (IRR = .913; SE = .011; t = –7.70; p < .001; Table 3, column 2), in support of H_1. High arousal language increases engagement with micro influencers' content (IRR = 1.036; SE = .008; t = 4.61; p < .001; Table 4, column 1). In turn, a 10% increase in arousal is associated with a 5.4% increase in engagement, implying 49 additional likes or comments, on average. Conversely, it decreases engagement with macro influencers' content (IRR = .944; SE = .009; t = –6.30; p < .001; Table 4, column 3), such that a 10% increase in arousal decreases engagement by 8.4%, meaning 346 fewer likes or comments, on average (see Figure 1 for marginal effects).

Informative Goal. First, including the macro (vs. micro) × informative interaction in the full model yields a significant outcome (IRR = 1.025; SE = .012; t = 2.13; p = .033, Table 3, column 3),[12] such that when the post is more informative (cf. commercial), macro influencers attract marginally more engagement (IRR = 1.018; SE = .010; t = 1.79; p = .073, Table 4, column 4). Second, consistent with our theorizing, the arousal × informative interaction reveals a significant positive effect for macro influencers (IRR = 1.027; SE = .009; t = 2.95; p

[12] Using informative words as the focal variable while controlling for commercial words produces similar results (IRR = 1.014; SE = .007; t = 2.10; p = .036). The results also reveal a marginally significant arousal × macro (vs. micro) × informative interaction (IRR = 1.022, SE = .012, t = 1.86, p = .063).

= .003; Table 4, column 4), suggesting that arousal's negative effect becomes attenuated when the goal of the post is informative (see Figure 2), in support of H_3.[13] Notably, our results corroborate some insights from prior research, as well as highlighting pertinent differences (see "Confirmation of Prior Findings" in Web Appendix B).

Table 3. Study 1 Results, Full Sample

	(1)	(2)	(3)
IV			
Arousal	1.036** (.009)	1.034** (.008)	1.034** (.008)
Macro (vs. Micro)	5.296** (.066)	4.522** (.065)	4.521** (.065)
Arousal × Macro	.906** (.011)	.913** (.011)	.912** (.011)
Informative Goal			1.002 (.008)
Macro × Informative Goal			1.025* (.012)
Controls			
Influencer			
if Verified		1.247** (.020)	1.247** (.020)
# of Posts		.905** (.006)	.904** (.006)
Content Variation		.981** (.006)	.982** (.006)
Category Fixed Effect		Included	Included
Text			
Topics		Included	Included
# of Question Marks		.992 (.006)	.992 (.006)
# of Hashtags		.991 (.007)	.990 (.007)
# of Mentions		1.002 (.006)	1.002 (.006)
Word Count		.997 (.007)	.994 (.007)
# of Emojis		1.011† (.006)	1.011† (.006)
Complexity		1.013 (.008)	1.013 (.008)
Valence		1.010 (.007)	1.010 (.007)
Concreteness		.979** (.007)	.980** (.007)
Familiarity		.981* (.007)	.981** (.007)
Image			
if Image (vs. Video)		1.353** (.030)	1.353** (.030)
if Face Present		.984 (.013)	.985 (.013)
Image Emotionality		1.051** (.007)	1.051** (.007)
Color Dominance		1.004 (.006)	1.003 (.006)
Color Saturation		1.012† (.006)	1.012† (.006)
Additional			
Time Difference		.989 (.006)	.989 (.006)
Time Fixed Effect		Included	Included
N	20,923	20,590	20,590
Log-likelihood	–184,781	–180,110	–180,106

† $p < .10$, * $p < .05$, ** $p < .01$.
Notes: Standard errors are in parentheses. We do not report coefficients for the fixed effects and topics, for parsimony.

[13] The effects could be driven by how frequently companies hire micro versus macro influencers to achieve specific post goals, but they do not differ in informative goal (.39 vs. .39).

Table 4. Study 1 Results, Micro and Macro Influencers

	Micro Influencers		Macro Influencers	
	(1)	(2)	(3)	(4)
Arousal	1.036** (.008)	1.036** (.008)	.944** (.009)	.946** (.009)
Informative Goal		.998 (.008)		1.018† (.010)
Arousal × Informative Goal		1.011 (.008)		1.027** (.009)
Controls	Included	Included	Included	Included
N	11,153	11,153	9,437	9,437
Log likelihood	–88,687	–88,686	–90,766	–90,760

† $p < .10$, * $p < .05$, ** $p < .01$.
Notes: Standard errors are in parentheses. We do not report coefficients for the controls, for parsimony.

Figure 1: Effects of Macro (vs. Micro) Influencers on Arousal

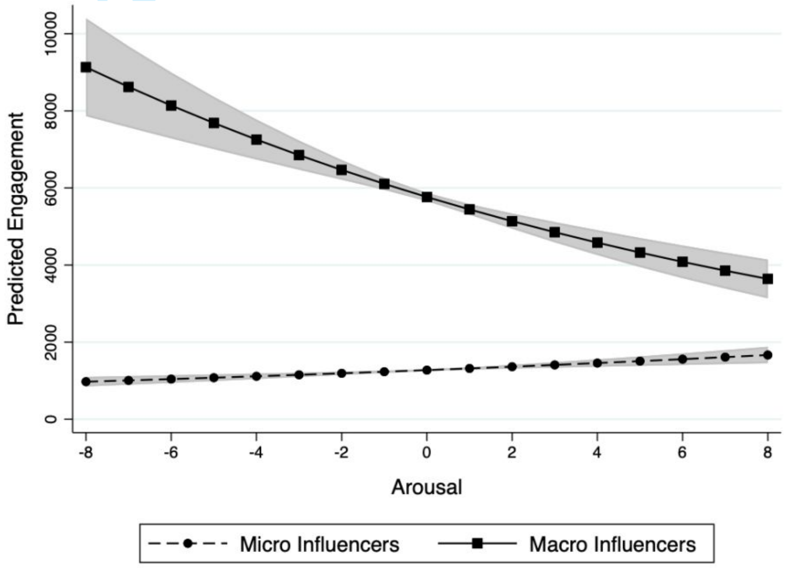
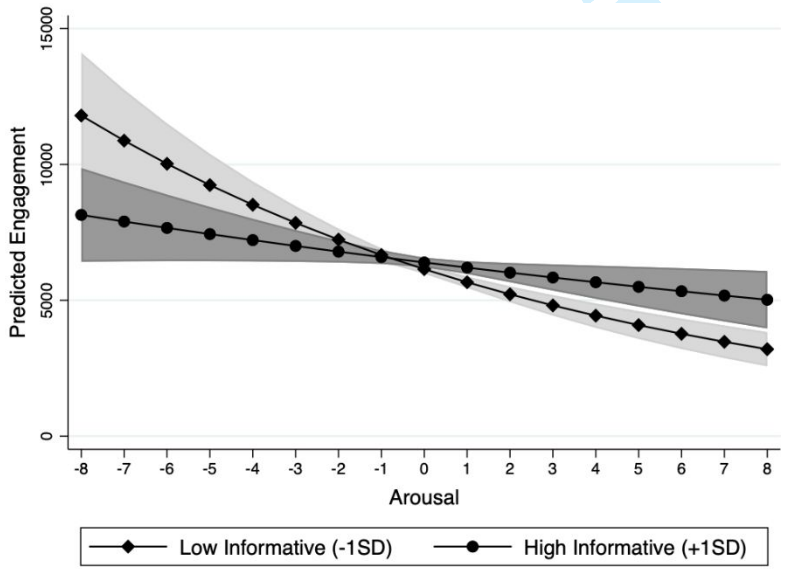


Figure 2: Effects of Informative Goal on Arousal for Macro Influencers



Robustness

With additional analyses, we test the robustness of our model specification and measurements. Table 5 includes an overview of what we tested and how; Table WB6 in Web Appendix B contains detailed results. First, the amount of arousal that an influencer includes in a post and the decision to use specific informative goal–related words might not be random. Variables that influence both these factors also might be unobserved or not available in the data set, which would lead to endogeneity (see "Addressing Endogeneity" in Web Appendix B). For arousal, we apply a control function approach (Petrin and Train 2010), using two-stage regression to derive a proxy variable that is conditional on the part of the observed endogenous regressor that depends on the error term. For the informative post goal, due to a lack of instruments that pertain to the marketing campaign, we use a Gaussian copula (Becker, Proksch, and Ringle 2022) to address the endogeneity that arises from strategic behaviors. Furthermore, different influencers might be better or worse at garnering engagement, so we account for this possibility using influencer fixed effects.

Second, the results might be driven by the particular sample of micro and macro influencers used (i.e., selection bias), so with propensity score matching, we match influencers on all the dimensions from the main model, allowing only arousal to vary (see "Selection Bias" in Web Appendix B).[14]

Third, we test whether the results hold when we adopt (1) an ordinary least squares (OLS) regression with a log-transformed dependent variable as the modelling approach and (2) the number of likes (wholly positive outcome) as the dependent variable.

Fourth, the results might reflect the measure of arousal that we use. To see if we can replicate the results, we (1) assess the arousal evoked by simple words and paralanguage separately, (2) use a different paralanguage measurement (PARA's effect size on sentiment

[14] We thank an anonymous reviewer for this suggestion.

intensity; Luangrath, Xu, and Wang 2023), and (3) replace our arousal measure with the LIWC words classified as high in activation by Villarroel Ordenes et al. (2017).[15]

Fifth, the results could be driven by the exact weights we used in the arousal operationalization in Equation 1, so we conducted several separate sensitivity analyses with different weights.

Sixth, in the original analysis, we compared micro and macro influencers using a dummy-coded classifier based on follower count. We test whether the results hold if we measure this variable continuously. The cutoff point we assign to categorize micro and macro influencers (100,000 followers) also might be influential, so we test the robustness of our results using alternative thresholds.

Finally, one might wonder whether high arousal language is also more sensory (Cascio Rizzo et al. 2023), and maybe that drove the effects. Yet the two constructs are not related (r = –.02), and when sensory language is included as a control, the arousal × macro (vs. micro) interaction produces the same results (IRR = .913; SE = .011; t = –7.69; p < .001).

Table 5. Overview of Robustness Checks

	What We Test	How We Test It
Modeling		
Endogeneity, Arousal	Might arousal decisions be driven by unobserved factors?	Control function approach (IRR = .912; SE = .011; t = –7.55; p < .001; Table WB6, col. 1).
Endogeneity, Informative Goal	Can informative-related post goals be explained by strategic behaviors?	Gaussian copula (not significant copula term: φ = –.004, p =. 11).
Endogeneity, Influencer	Do the effects depend on influencer heterogeneity?	Influencer fixed effects with cluster-robust standard errors (IRR = .950; SE = .013; t = –4.50; p < .001).
Selection Bias	Are the effects driven by the particular sample of micro and macro influencers used?	Propensity score matching (IRR = .910; SE = .015; t = –5.75; p < .001; Table WB6, col. 2).
Alternative approach	Do data ranges make the use of count distributions inappropriate?	OLS with log-transformed DV (b = –.089; SE = .015; t = –6.10; p < .001; Table WB6, col. 3).
Alternative DV	Do the results hold for likes only?	Likes as DV (IRR = .910; SE = .011; t = –7.90; p < .001).
Measurement		

[15] Note that the valence × macro (vs. micro) interaction is not significant (IRR = .991; SE = .012; t = –.75; p = .452), casting doubt on the possibility that valence drives the effects.

Alternative measure of arousal	Do simple words and paralanguage arousal separately replicate the effects?	Assessed arousal for simple words (IRR = .926; SE = .012; t = –6.02; p < .001; Table WB6, col. 4) and paralanguage (IRR = .917; SE = .032; t = –2.50; p = .012; Table WB6, col. 5), separately.
	Does a different paralanguage measure replicate the effects?	Applied the effect size (γ = .36) of PARA on sentiment intensity (IRR = .938; SE = .009; t = –6.29; p < .001; Table WB6, col. 6).[a]
	Do measurements related to arousal all converge to the same concept?	Used high arousal LIWC words as focal variable (IRR = .908; SE = .007; t = –1.66; p = .098; Table WB6, col. 7).
Sensitivity analysis for arousal	Are the results driven by the exact weights used in the arousal operationalization in Equation 1?	For capital letters, the results hold when the multiplier changes from 1.5 to 1.4 (IRR = .912; SE = .011; t = –7.81; p < .001) or to 1.6 (IRR = .914; SE = .011; t = –7.58; p < .001). For exclamation marks, the results hold when the multiplier changes from 1.2 to 1.1 (IRR = .903; SE = .011; t = –8.65; p < .001) or to 1.3 (IRR = .922; SE = .011; t = –6.80; p < .001). Results remain the same even setting both multipliers at 1 (IRR = .930; SE = .016; t = –5.88; p < .001).
Alternative measure of Macro	Do results hold without dummy-coding macro?	Macro measured continuously (IRR = .976; SE = .006; t = –3.90; p < .001; Table WB6, col. 8).
	Do results hold for different macro classification thresholds?	Three alternative follower count thresholds: 75,000 (IRR= .938; SE = .012; t = –5.14; p < .001), 150,000 (IRR= .913; SE = .011; t = –7.72; p < .001), and 200,000 (IRR= .915; SE = .011; t = –7.36; p < .001).

[a]In their Study 2, Luangrath, Xu, and Wang (2023) identify an effect size of .36 for the presence of text paralanguage (TPL) on sentiment intensity. We use their PARA tool to detect TPL in posts and dummy-code for its presence. Every time TPL appears within a post, we use .36 as a multiplier of simple words' arousal.

Exploring the Hypothesized Process

Testing Trust in Comments. We test the hypothesized mechanism in more detail in Study 3a, but with a preliminary test, we seek insights into the relationship between arousal × macro (vs. micro) and trust. If high arousal language increases (decreases) engagement for micro (macro) influencers because it makes the consumer believe the influencer is more (less) trustworthy, as we suggest, then we should expect more (less) trust in followers' comments when posts include grater language arousal. To test this possibility, we use the trust scores established by Mohammad's (2017) NRC Affect Intensity Lexicon to measure expressions of trust in 457,132 followers' comments (averaging scores across comments at the post level). The dependent variable is truncated in the interval [0, 1], so we adopt censored Tobit model.[16] Consistent with our theorizing, the arousal × macro (vs. micro) interaction is significant (b = –.013; SE = .002; t = –5.56; p < .001; Table WB7). Greater language arousal increases trust in micro influencers (b = .009; SE = .002; t = 5.16; p < .001) but decreases trust in macro influencers (b = –.004; SE = .001; t = –2.67; p = .008).

Moderating Role of Trustworthiness Cues. Another test offers preliminary insights into the underlying process for macro influencers, through moderation. If high arousal language decreases engagement because it makes macro influencers seem less trustworthy, as we suggest, then their use of trustworthiness cues should mitigate the negative effect of high arousal language. To gauge the degree of trustworthiness, as signaled by cues in influencer posts, we rely on Wordify (Hovy, Melumad, and Inman 2021) and create a dictionary of 22 trustworthiness words (e.g., "help," "learn," "experience") and 8 non-trustworthiness words (e.g., "gifted," "sponsor"; see "New Dictionary for Language Trustworthiness" in Web Appendix B for the full list of words). Consistent with our predictions, the arousal × trustworthiness interaction is significant (IRR = 1.021; SE = .009; t = 2.41; p = .016), such that more trustworthiness cues reduce the negative effect of high arousal language on engagement with macro influencers.[17]

Discussion

Study 1 provides preliminary support for our theorizing. First, with an analysis of more than 20,000 influencers' sponsored posts, we determine that greater language arousal (i.e., 10% increase) increases engagement with micro influencers by 5.4% but decreases engagement with macro influencers by 8.4%. The results are robust to various controls and model specifications.

[16] Our data include the first 24 comments per post. When we apply OLS, the same results emerge.
[17] The arousal × trustworthiness interaction is not significant for micro influencers (IRR = 1.009; SE = .008; t = 1.16; p = .248).

Second, with two ancillary analyses, we begin to explore the underlying role of trust and persuasion knowledge. Followers express more (less) trust when micro (macro) influencers use high arousal language, and the negative effect of high arousal language for macro influencers is mitigated when the post has a more informative than commercial goal or if they include trustworthiness cues in their posts.

Follow-Up to Study 1

Although the preceding results are consistent with our theorizing, they arguably could reflect the specific platform used or a reliance on written communication. In a follow-up study, we therefore examine TikTok, the video-sharing social network that aims to drive engagement with audio-visual content. Rather than focusing on posts' text (e.g., title, hashtags), we consider influencers' speech.

With a sample of 654 influencers' sponsored videos on TikTok (from Cascio Rizzo et al. 2023; see Tables WC1 and WC2 in Web Appendix C), we use automated audio analysis to measure the level of arousal in influencers' voices. Arousal can be gauged by the pitch of such speech; when people are in a high arousal state (e.g., happy), they tend to speak in a higher pitched voice (Bänziger and Scherer 2005; Laukka et al. 2016; Mauss and Robinson 2009). After extracting the audio speech from videos, we used the YIN frequency estimator algorithm (De Cheveigné and Kawahara 2002) to measure the level of each influencer's pitch. Because the data include influencers with more than 100,000 followers, we measure influencer type continuously (i.e., follower count) and account for controls similar to those in Study 1. To examine the features of the verbatim speech, we hired professionals from Upwork to transcribe the videos, from which we also identified other vocal factors (loudness, intonation, brightness, articulation rate, and speech duration) that might affect the results. Thus, we test the relationship among pitch, follower count, and engagement.

Consistent with the Study 1 results, we find a negative, significant pitch × follower count interaction effect (IRR = .776; SE = .047; t = –4.19; p < .001) on engagement. Even after accounting for the controls, this significant effect of pitch × follower count persists (IRR = .785; SE = .040; t = –4.78; p < .001; see Table WC3 for results and Figure WC1 for marginal effects). Thus, as the follower count increases, higher pitch (i.e., higher arousal) exerts a negative effect on engagement (see Web Appendix C). We also test the effect of arousal in the audio verbatims. Consistent with Study 1, the verbatim arousal × follower count interaction exhibits consistent results (IRR = .704; SE = .091; t = –2.71; p < .001). By specifying similar effects across distinct social media platforms, which prioritize spoken versus written language, we affirm the robustness and generalizability of the effects.

Study 2: Manipulating Arousal

Although the Study 1 results are consistent with our theorizing and cast doubt on various alternative explanations, they do not establish whether the relationships among arousal, influencer type, and engagement are causal. In Study 2, we manipulate language arousal and influencer type (micro vs. macro) to examine whether high arousal language increases engagement with micro influencers while decreasing it with macro influencers (H_1).

Method

Participants (N = 279, Prolific) were randomly assigned to a 2 (language arousal: high vs. low) × 2 (influencer type: micro vs. macro) between-subjects design. Web Appendix D contains the preregistrations, exclusions, demographics, stimuli, and manipulation checks for all experiments. Briefly though, all participants saw a fictitious influencer's Instagram post, sponsoring a granola product. Conditions varied in terms of the arousal exhibited by the language in the post. In the high [*low*] arousal condition, the post read, "#adv Choose granola by @sobbis for the BEST [*great*] snack!!![.] Love [*like*] its ginger snap in my smoothies. Delish [*Good*] also as topping or eaten by itself![.]" (arousal score using Study 1 measurement = .59 [*.43*][18]). We also varied the influencer type (micro: 20,000 followers, macro: 660,000 followers).[19] A pretest confirmed the effectiveness of the manipulation, indicating that the macro influencer appeared able to reach more people than the micro influencer (M = 6.10 vs. 5.17; $F(1, 78) = 14.93$, $p < .001$, $\eta^2 = .161$).

Then participants were asked how likely they would be to engage with the post (i.e., like or comment on it; 1 = not at all, 7 = very; Cascio Rizzo et al. 2023; Valsesia, Nunes, and Proserpio 2020). Finally, respondents completed manipulation checks, an attention check, and demographic items.

Results

Engagement. A 2 × 2 analysis of variance (ANOVA) indicates the predicted arousal × influencer type interaction ($F(1, 275) = 9.10$; $p = .003$, $\eta p^2 = .032$). Consistent with Study 1, high (vs. low) arousal language increases engagement with micro influencers ($M_{high} = 2.60$; $M_{low} = 1.94$; $F(1, 275) = 4.91$, $p = .028$, $\eta p^2 = .017$) but decreases engagement with macro influencers ($M_{high} = 1.90$; $M_{low} = 2.49$; $F(1, 275) = 4.20$, $p = .041$, $\eta p^2 = .015$).

Discussion

Study 2 provides direct causal evidence that high arousal language increases engagement with micro influencers but decreases it with macro influencers, in support of H_1. Participants engage more when exposed to a micro influencer's post featuring high rather than low arousal language, but they are less likely to engage with a macro influencer's post featuring high rather than low arousal language.

---

[18] We designed our experimental stimuli to vary in arousal levels, based on our field data (M = .50, SD = .07), but not in valence (.69 vs .69). A pretest in which participants had to assess the level of arousal (7-point scales, "very passive–very active," "very mellow–very fired up," "very low energy–very high energy"; $\alpha = .92$; Berger and Milkman 2012) confirmed that the language in the high arousal condition was perceived as more arousing than the language in the low arousal condition (M = 5.06 vs. 3.2, $F(1, 78) = 44.10$, $p < .001$, $\eta^2 = .361$) but similar in valence (M = 5.23 vs. 4.84, $F(1, 78) = 2.00$, $p = .161$, $\eta^2 = .025$; 7-point scales, "very bad–very good," "very unfavorable–very favorable," "very unpleasant–very pleasant"; $\alpha = .91$; Berger and Milkman 2012).
[19] The number of followers for the micro influencer falls between the 5th and 10th percentiles, while the number for the macro influencer falls between the 90th and 95th percentiles, according to our Study 1 data (Valsesia, Proserpio, and Nunes 2020). We informed participants about the number of followers each influencer had.

Study 3a: Testing the Process

Study 3a has three main goals. First, we test the hypothesized underlying process. We suggest that high arousal language increases (decreases) engagement for micro (macro) influencers because it makes people believe the influencer's intent is less (more) persuasive, which increases (decreases) influencer trustworthiness (H_2). Second, Study 1 already spanned a broad range of product categories, but to confirm generalizability in this experimental context, in Study 3a, we use a different product category and different wording. Third, we test alternative explanations based on valence, likeability, effort, helpfulness, and expertise.

Method

Participants (N = 271, Prolific) were randomly assigned to a 2 (language arousal: high vs. low) × 2 (influencer type: micro vs. macro) between-subjects design. Everyone was shown a fictitious influencer's Instagram post sponsoring a restaurant. To vary language arousal, we used a method similar to that in Study 2, such that in the high [*low*] arousal condition, the post read, "#ad When my schedule is busy from start to finish, I LOVE [*like*] heading to @fishdancer at lunchtime. Great [*Good*] food in the ultimate [*right*] place!!![.] Try it!![.] 😁 [☺]" (arousal score = .59 [*.42*]).[20] The influencer type manipulation was the same as in Study 2.

Next, we collected process measures. After viewing the post, participants rated the influencer's intent using three persuasion knowledge items (7-point scale, "good–bad," "not pushy–pushy," and "not aggressive–aggressive"; α = .80; Ahluwalia and Burnkrant 2004), and they indicated their perceptions of the influencer's trustworthiness using five items (7-point scale, "untrustworthy–trustworthy," "insincere–sincere," "undependable–dependable," "dishonest–honest," "unreliable–reliable"; α = .95; Ohanian 1990). The measure of

[20] Conditions do not vary in valence (.68 vs .66). The pretest confirmed that the language in the high arousal condition appeared more arousing than that in the low arousal condition (M = 4.86 vs. 3.88, $F(1, 78) = 10.87$, $p = .001$, $\eta^2 = .122$), but they were similar in valence (M = 5.60 vs. 5.28, $F(1, 78) = 1.28$, $p = .262$, $\eta^2 = .016$).

engagement matches that from Study 2. Finally, participants completed ancillary measures to test for alternative explanations (i.e., valence, likeability, effort, helpfulness, and expertise), as well as the same manipulation and attention checks and demographic items from Study 2.

Results

Engagement. The 2 × 2 ANOVA reveals the predicted arousal × influencer type interaction ($F(1, 267) = 10.33$; $p = .001$, $\eta p^2 = .037$). That is, high (vs. low) arousal language again increases engagement with the micro influencer ($M_{high} = 2.80$; $M_{low} = 2.03$; $F(1, 267) = 6.94$, $p = .009$, $\eta p^2 = .024$) and decreases engagement with the macro influencer ($M_{high} = 1.84$; $M_{low} = 2.41$; $F(1, 267) = 3.66$, $p = .047$, $\eta p^2 = .013$).

Persuasion Knowledge. A 2 × 2 ANOVA reveals the predicted arousal × influencer type interaction ($F(1, 267) = 18.07$, $p < .001$, $\eta p^2 = .063$). High (vs. low) arousal language leads people to perceive the micro influencers' intent as less persuasive ($M_{high} = 2.93$; $M_{low} = 3.38$; $F(1, 267) = 4.87$, $p = .028$, $\eta p^2 = .018$) but the macro influencers' intent as more persuasive ($M_{high} = 3.60$; $M_{low} = 2.82$; $F(1, 267) = 14.46$, $p < .001$, $\eta p^2 = .051$).

Trustworthiness. A 2 × 2 ANOVA reveals the predicted arousal × influencer type interaction ($F(1, 267) = 13.09$, $p < .001$, $\eta p^2 = .047$). High (vs. low) arousal language increases micro influencers' trustworthiness ($M_{high} = 4.83$; $M_{low} = 4.31$; $F(1, 267) = 6.91$, $p = .015$, $\eta p^2 = .025$); it decreases macro influencers' trustworthiness ($M_{high} = 4.11$; $M_{low} = 4.68$; $F(1, 267) = 7.10$, $p = .008$, $\eta p^2 = .026$; see Figure 3).

Serial Moderated Mediation. In a moderated serial mediation analysis (PROCESS model 83; Hayes 2018), with influencer type as a moderator of language arousal's effects on persuasion knowledge and trustworthiness, we find significant moderated serial mediation on engagement ($b = –.66$; 95% confidence interval [CI] = –1.06; –.33). As predicted, in the micro influencer condition, the effect of language arousal on engagement is driven sequentially by persuasion knowledge and trustworthiness ($b = .24$; 95% CI = .01; .49).

Using high arousal language makes people believe that the intent is less persuasive (b = –.45, SE = .20, t = –2.21, p = .028), which makes the influencer seem more trustworthy (b = .66, SE = .05, t = 13.74, p < .001), which enhances engagement (b = .81, SE = .09, t = 9.27, p < .001). In the macro influencer condition, the effect of language arousal also is sequentially driven by persuasion knowledge and trustworthiness (b = –.42; 95% CI = –.68; –.20), such that high arousal language increases feelings that the influencer's intent is persuasive (b = .78, SE = .21, t = 3.80, p < .001), which decreases trustworthiness and engagement.[21]

Figure 3: Moderation by Influencer Type (on Trustworthiness)

Alternative Explanations. With ancillary assessments, we can discount several alternative explanations. First, rather than being driven by arousal, the results might reflect post valence. Both Study 1 and a priori tests with experimental stimuli dismiss this possibility, but to explore it further, we collect measures of valence (α = .95). Because

[21] A moderated serial mediation with the positions of persuasion knowledge and trustworthiness switched does not hold (b = –.02; 95% CI = –.14; .11).

valence does not vary by condition (F(1, 269) = 2.48, p = .116, η^2 = .009), the results cast doubt on this explanation.

Second, maybe the high arousal language somehow made the influencer seem more (or less) likeable, which drove the effect. To test this possibility, we used a two-item measure of source likeability ("unlikeable–likeable," "unfriendly–friendly"; r = .80; Schwartz, Luce, and Ariely 2011). The results reveal a non-significant arousal × influencer interaction on likeability though (F(1, 267) = .45; p = .504, ηp^2 = .002).

Third, perhaps high arousal language informs people about the level of effort that influencers put into composing the post, which could drive the effect. To test this possibility, we used a one-item measure adapted from Yin, Bond, and Zhang (2017, "In your opinion, how much effort did the influencer put into writing the post?"; 1 = not at all, 7 = very much). Results reveal a non-significant arousal × influencer type interaction on effort (F(1, 267) = .49; p = .487, ηp^2 = .002), casting doubt on this alternative.

Fourth, maybe high arousal language makes the post seem more helpful. To test this possibility, we asked participant how helpful they found the post (1 = not at all, 7 = very much). These results also reveal a non-significant arousal × influencer type interaction on helpfulness (F(1, 267) = 2.06; p = .152, ηp^2 = .008).

Fifth, maybe people elaborate on arousal and influencer type to make inferences about influencer expertise, which drives the effects. Using Packard and Berger's (2021) two-item measure of expertise ("not expert–expert," "not intelligent–intelligent"; r = .80), we find another non-significant arousal × influencer interaction on expertise (F(1, 267) = 1.21; p = .273, ηp^2 = .004). Thus, none of the alternative explanations appear reasonable.

Discussion

Study 3a provides further evidence of the effects of language arousal while also illustrating why they occur. First, as in our prior studies, the presence of high arousal language leaves participants more (less) willing to like and comment on the micro (macro) influencer's post (H_1). Second, consistent with our theorizing, these effects are sequentially driven by persuasion knowledge and trust (H_2). Using high arousal language makes people feel like the micro (macro) influencer's intent is less (more) persuasive, which causes the influencer to seem more (less) trustworthy, which increases (decreases) engagement. Third, we cast doubt on the possibility that valence, likeability, effort, helpfulness, or expertise drive the effects.

Study 3b: Process by Moderation (Language Valence)

For another test of the process, in Study 3b, we examine the role of language valence.[22] If our theorizing about the underlying role of persuasion knowledge is correct, counterbalancing the valence of the message should mitigate the negative effect of high arousal language for macro influencers. That is, an influencer who acknowledges some concerns or negative aspects of the promoted product may appear less likely to be driven by manipulative intent (Uribe, Buzeta, and Velásquez 2016). Therefore, high arousal language should have a weaker (negative) influence. To test this prediction, we use the design from Study 3a but add a sentence that signals negative valence to the conditions. We also include a different measure of persuasion knowledge (Campbell 1995) and rule out an alternative explanation based on warmth.

Method

Participants (N = 290, Prolific) were randomly assigned to a 2 (language arousal: high vs. low) × 2 (influencer type: micro vs. macro) between-subject design.

Most of the materials were the same as those in Study 3a, except that we added a sentence signaling negative valence to the language arousal conditions (i.e., "SUPER expensive" for high arousal, "very expensive" for low arousal one). The engagement and

[22] We thank an anonymous reviewer for this suggestion.

trust (α = .92) measures were the same as in Study 3a. To test robustness to different persuasion knowledge measures, we used six items adapted from Campbell's (1995) persuasion knowledge scale (i.e., "The influencer tried to manipulate the audience in ways that I do not like," "I was annoyed by this post because the influencer seemed to be trying to inappropriately manage or control the consumer audience," "The influencer was excessively manipulative," "The influencer was unfair in what she said," "I think that this influencer was unfair," "The way the influencer tries to persuade seems unacceptable to me," 1 = strongly disagree, 7 = strongly agree; α = .94).

Because high arousal language also might shape perceptions of influencers' warmth, which could drive the identified effects, we test this possibility with an adapted version of Wang et al.'s (2017) three-items measure of warmth ("warm," "kind," "friendly"; α = .93). Participants completed this measure, along with an attention check and the demographic items, as the last step.

Results

Engagement. The 2 × 2 ANOVA reveals the predicted arousal × influencer type interaction ($F(1, 286) = 5.36$; $p = .021$, $\eta p^2 = .018$). Consistent with our prior experiments, high (vs. low) arousal language increases engagement with the micro influencer (M_{high} = 1.84; M_{low} = 1.35; $F(1, 286) = 6.52$, $p = .011$, $\eta p^2 = .024$). Consistent with our theorizing, adding a sentence with negative valence eliminates this negative effect of high arousal for macro influencers (M_{high} = 1.62; M_{low} = 1.75; $F(1, 286) = .49$, $p = .483$, $\eta p^2 = .004$).

Persuasion Knowledge. A 2 × 2 ANOVA reveals the predicted arousal × influencer type interaction ($F(1, 286) = 4.14$, $p = .043$, $\eta p^2 = .014$). High (vs. low) arousal language leads people to perceive the micro influencers' intent as less persuasive (M_{high} = 2.68; M_{low} = 3.19; $F(1, 286) = 4.69$, $p = .031$, $\eta p^2 = .015$). As expected, arousal has no effect for macro influencers (M_{high} = 3.26; M_{low} = 3.10; $F(1, 286) = .48$, $p = .489$, $\eta p^2 = .002$).

Trustworthiness. Another 2 × 2 ANOVA reveals the predicted arousal × influencer type interaction ($F(1, 286) = 6.51$, $p = .011$, $\eta p^2 = .022$). High (vs. low) arousal language increases micro influencers' trustworthiness ($M_{high} = 4.49$; $M_{low} = 3.89$; $F(1, 286) = 8.51$, $p = .004$, $\eta p^2 = .028$) but has no effect on macro influencers' trustworthiness ($M_{high} = 3.93$; $M_{low} = 4.06$; $F(1, 286) = .45$, $p = .504$, $\eta p^2 = .003$).

Serial Moderated Mediation. In a moderated serial mediation analysis (PROCESS model 83; Hayes 2018), with influencer type as a moderator of language arousal's effects on persuasion knowledge and trustworthiness, we find significant moderated serial mediation on engagement ($b = .12$; 95% CI = .01; .26). As in Study 3a, in the micro influencer condition, the effect of language arousal on engagement is driven sequentially by persuasion knowledge and trustworthiness ($b = .09$; 95% CI = .01; .20). Using high arousal language makes people believe that the intent is less persuasive ($b = –.50$, SE = .23, $t = –2.17$, $p = .031$), which increases trustworthiness ($b = .48$, SE = .04, $t = 10.80$, $p < .001$), which enhances engagement ($b = .39$, SE = .06, $t = 6.36$, $p < .001$). In the macro influencer condition, language arousal no longer affects persuasion knowledge ($b = .16$; SE = .23, $t = .70$, $p = .487$), and the serial mediation is not significant ($b = –.03$, 95% CI = –.12, .05).

Alternative Explanation. Although we test warmth as an alternative explanation for the identified effects, the results indicate a non-significant arousal × influencer type interaction ($F(1, 286) = .95$; $p = .331$, $\eta p^2 = .003$). These findings cast doubt on the relevance of this alternative pathway.

Discussion

Study 3b provides further evidence of the underlying role of persuasion knowledge. Consistent with the prior studies, using high arousal language leads consumers to believe the micro influencer's intent is less manipulative and thus that the influencer is more trustworthy, which increases their engagement. Consistent with the role of persuasion knowledge, a sentence with negative valence lowers perceptions of manipulative intent, such that the negative effect of high arousal language is mitigated for macro influencers.

Furthermore, the findings reveal a boundary condition of the effect of language arousal. Prior research (Berger and Milkman 2012) indicates that positive or negative arousal boosts certain forms of engagement; we find that only positive arousal shapes engagement in relation to influencer content and persuasion knowledge. Indeed, unlike news content (Berger and Milkman 2012), when it comes to influencers' recommendation, consumers may tend to scrutinize the use of positive arousal to infer the poster's true motives.

Study 4: Process by Moderation (Post Goal)

Study 4 has two main goals. First, it tests the hypothesized process for macro influencers through both mediation and moderation. If high arousal language decreases engagement by making it seem like macro influencers have a persuasive intent, as we suggest, then the effect should be mitigated when posts focus less on persuasion. To test this possibility, in addition to manipulating arousal, we manipulate the post goal (informative vs. commercial; similar to the analysis in Study 1). If our theorizing is correct, high arousal posts should have less of an effect when their goal is informative rather than commercial (H_3). Second, we examine an additional dependent variable. Engagement is correlated with sales (Kumar et al. 2016), so we test if the observed effects also extend to intentions to choose the endorsed product.

Method

Participants (N = 279, Prolific) were randomly assigned to a 2 (language arousal: high vs. low) × 2 (post goal: baseline [commercial] vs. informative) between-subjects design. The baseline condition mimicked that for Study 3a. To signal that the post goal was informative, we replaced the phrase "Try it" with "Learn more about it" (reflecting the dictionary and measure in Study 1). We also described the influencer as a macro influencer, with 660,000 followers.

The process and engagement measures also were the same as in Study 3a. To explore whether language arousal affects consumer choice, we asked participants how likely they were to choose the endorsed restaurant in the future (1 = not at all, 7 = very; Packard, Gershoff, and Wooten 2016). Participants then completed two manipulation checks, the tests of alternative explanations from Study 3a, an attention check, and demographic items.

Results

Engagement. A 2 × 2 ANOVA reveals the predicted arousal × post goal interaction ($F(1, 275) = 8.17$; $p = .005$, $\eta p^2 = .029$). Consistent with our prior studies, high (vs. low) arousal language decreases engagement with the macro influencer in the commercial condition ($M_{high} = 1.87$; $M_{low} = 2.70$; $F(1, 275) = 8.25$, $p = .004$, $\eta p^2 = .029$). Consistent with the hypothesized underlying role of trust, however, when the post has an informative goal, this difference disappears ($M_{high} = 2.74$; $M_{low} = 2.43$, $F(1, 275) = 1.28$, $p = .259$, $\eta p^2 = .005$).

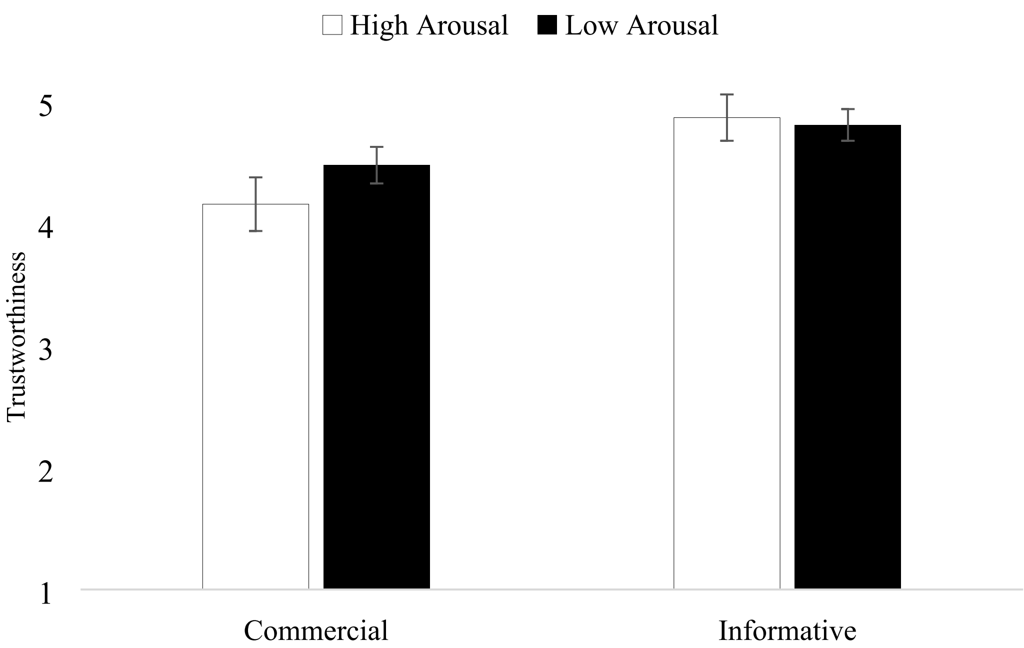
Choice Likelihood. We observe similar effects for choice likelihood. A 2 × 2 ANOVA reveals the predicted arousal × post goal interaction ($F(1, 275) = 5.76$; $p = .017$, $\eta p^2 = .020$). Consistent with our theorizing, high (vs. low) arousal language makes people less likely to choose the endorsed restaurant in the commercial condition ($M_{high} = 3.02$; $M_{low} = 3.61$; $F(1, 275) = 4.97$, $p = .026$, $\eta p^2 = .018$), but this effect is mitigated when the post indicates an informative goal ($M_{high} = 3.79$; $M_{low} = 3.50$, $F(1, 275) = 1.30$, $p = .255$, $\eta p^2 = .005$).

Persuasion Knowledge. The main effect of the post goal ($F(1, 275) = 7.70$, $p = .006$, $\eta p^2 = .027$) is qualified by the predicted arousal × post goal interaction ($F(1, 275) = 4.04$, $p = .045$, $\eta p^2 = .014$). Consistent with Study 3a, high (vs. low) arousal language leads people to perceive the influencers' intent as more persuasive in the commercial condition ($M_{high} = 3.49$;

M_{low} = 2.89; $F(1, 275) = 10.38$, $p = .001$, $\eta p^2 = .036$), but this effect is mitigated when the post has an informative goal (M_{high} = 2.87; M_{low}= 2.79, $F(1, 275) = .21$, $p = .644$, $\eta p^2 = .001$).

Trustworthiness. The main effect of the post goal ($F(1, 275) = 7.96$, $p = .005$, ηp^2 = .028) is qualified by the predicted arousal × post goal interaction ($F(1, 275) = 3.96$, $p = .048$, $\eta p^2 = .014$). Consistent with Study 3a, high (vs. low) arousal language decreases influencers' trustworthiness in the commercial condition (M_{high} = 4.16; M_{low} = 4.68; $F(1, 275) = 6.08$, p = .014, $\eta p^2 = .022$), but this effect is mitigated when the post has an informative goal (M_{high} = 4.87; M_{low} = 4.81, $F(1, 275) = .09$, $p = .766$, $\eta p^2 = .001$; see Figure 4).

Figure 4: Moderation by Post Goal for Macro Influencers (on Trustworthiness)



Error bars: +/– 1 SE

Serial Moderated Mediation. A moderated serial mediation analysis (PROCESS model 83; Hayes 2018), which incorporates the post goal as a moderator of language arousal's effects on persuasion knowledge and trustworthiness, reveals significant moderated serial mediation on both engagement (b = –.31; 95% CI = –.06; –.01) and choice (b = –.28; 95% CI = –.59; –.01). As predicted, in the commercial condition, the effect of arousal is driven sequentially by persuasion knowledge and trustworthiness (engagement: b = –.35; 95% CI =

–.63; –.11; choice: b = –.32; 95% CI = –.58; –.10). Using high arousal language causes consumers to believe the influencer's intent is more persuasive (b = .60, SE = .19, t = 2.22, p = .001), which makes the influencer seem less trustworthy (b = –.67, SE = .05, t = –12.36, p < .001), which decreases both engagement (b = –.87, SE = .08, t = –10.85, p < .001) and choice (b = –.78, SE = .07, t = –10.72, p < .001). When the post goal is informative though, language arousal no longer induces persuasion knowledge (b = .08, SE = .18, t = .46, p = .644), and the mediation is not significant (engagement: b = –.05, 95% CI = –.23, .15; choice: b = –.04, 95% CI = –.22, .12).

Alternative Explanations. In these ancillary analyses, we use the measures from Study 3a to test the alternative explanations. We only expected language arousal to exert effects in the commercial condition though, so we focus the analyses on that condition. We do not find any variations with regard to valence ($F(1, 277) = .01$, p = .995, $\eta^2 = .001$), likeability ($F(1, 131) = .02$; p = .894, $\eta p^2 = .001$), effort ($F(1, 131) = .10$; p = .751, $\eta p^2 = .001$), helpfulness ($F(1, 131) = .21$; p = .648, $\eta p^2 = .002$), or expertise ($F(1, 131) = 1.14$; p = .288, $\eta p^2 = .009$) and thus can rule out these alternative explanations again.

Discussion

Study 4 underscores the hypothesized underlying role of trust (and persuasion knowledge) for macro influencers, through both mediation and moderation. First, consistent with our prior studies, high arousal language decreases engagement with macro influencers because it leads people to believe that the influencer's intent is more persuasive, which makes the influencer appear less trustworthy. That said, consistent with the notion that these effects are driven by trust, if the campaign goal is more informative than commercial, the effects of arousal disappear, in support of H_3. Second, these effects extend to likelihood of choosing the endorsed product. Consumers are less willing to choose a restaurant that the macro influencer has endorsed if that influencer uses high arousal language.

General Discussion

Influencers represent a popular marketing strategy for brands. Consumers watch movies influencers suggest, buy apparel influencers wear, and visit places influencers recommend. But though influencers have the potential to spread awareness and boost sales, not all posts are equally effective. Whereas some posts attract lots of engagement, others do not, in part because people question the motivations that lead influencers to sponsor brands. This multimethod study explores how subtle shifts in language arousal can shape influencer trustworthiness and engagement, depending on the type of influencer they are.

Contributions

This research makes several contributions. First, we advance extant knowledge on the role of arousal in marketing communication, moving beyond its effects in traditional communication channels (Berger and Milkman 2012), to address source characteristics that inform its impact on online consumer responses. As we show, language arousal exerts differential impacts on consumers' engagement with sponsored content, depending on the influencer who posts the content. In particular, whereas high arousal language triggers positive responses with micro influencers, it backfires with macro influencers. In this sense, our work is in line with and advances prior research (e.g., Kanuri, Chen, and Sridhar 2018; Weingarten and Berger 2017; Yin, Bond, and Zhang 2017) that identifies boundary conditions for the positive behavioral outcomes of language arousal. Also, finding similar effects across distinct social media platforms and communication modalities (i.e., Instagram text data and TikTok audio data) offers evidence of the robustness and generalizability of the effects (Boegerhausen et al. 2022; Packard and Berger 2024).

Second, we shed light on the drivers of influencer trustworthiness. Although research has begun to investigate *who* (micro or macro influencer) seems more trustworthy (Karagür et al. 2022), less attention centers on *how* the type of language that micro (vs. macro) influencers use might inform perceptions of their trustworthiness. As we demonstrate, simple shifts in language arousal (depending on influencer type) can signal trustworthiness, which provides a crucial asset that influencers can use to market themselves to brands. Our ancillary analyses in Study 1 offer evidence of this underlying role of trustworthiness. High arousal language leads followers to express more (less) trust in comments when the poster is a micro (macro) influencer. We use Wordify to construct a dictionary of words indicative of trustworthiness (e.g., "help," "learn," "experience"); using such words to signal trustworthiness can attenuate the negative effect of language arousal for macro influencers.

Third, we extend research on persuasion knowledge (Karagür al. 2022) and campaign goals (Hughes, Swaminathan, and Brooks 2019) in influencer marketing. We demonstrate that macro influencers attract more engagement when their posts have an informative, rather a commercial, goal. These results are in line with practical advice (Weber 2022); our research provides the first empirical evidence in support of such recommendations. We also establish that the negative effect of arousal on engagement with macro influencers is mitigated when the post has an informative purpose, because informative-oriented posts focus less on persuasion, so they are less likely to activate persuasion knowledge. Furthermore, we demonstrate that counterbalancing the valence of the message (e.g., noting negative aspects of the endorsed product) has a similar effect, such that our findings appear limited to positive language with high arousal.[23] Unlike news content (Berger and Milkman 2012), in the context of influencers, people are more likely to elaborate on the use of positive language arousal (e.g., "fantastic product!") to make inferences about the poster's true motives.

Fourth, our findings also have implications for text-based social media measures. As we show, by integrating insights about language arousal (Berger and Milkman 2012) and

[23] The valence range in Study 1 is positive (M = .67), with a low standard deviation (SD = .06), which aligns with the notion that the effects hold only for positive arousal.

paralanguage (Luangrath, Peck, and Barger 2017), it is possible to move beyond a sole focus on the effects of arousal communicated through words (e.g., Herhausen et al. 2019, 2023; Kuperman et al. 2014). Dictionaries such as the VAD do not account for paralanguage features, but such features clearly can communicate arousal. For relevant text analysis efforts in marketing research (Berger, Moe, and Schweidel 2023; Berger et al. 2020; Humphreys and Wang 2018; Packard, Moore, and McFerran 2018), we suggest a better option for gauging arousal and the psychological functions of language. We also provide two dictionaries that can be used to measure the goals of posts (informative or commercial).

Practical Implications

Organizations leverage influencer content, with the assumption that it appears more trustworthy and boosts engagement. However, as we show, that is not always the case. In our observations, about 85% of posts use at least one word with an arousal value above the 75th percentile, and 80% include high arousal paralanguage (emojis, capitalization, exclamation marks), which may undermine their engagement effects. On the basis of field data, involving more than 1,000 Instagram influencers in 18 different industries (e.g., economics, design, travel, food), we offer practical and generalizable findings for influencer marketing.

Arousal for Macro Influencers. Our findings raise concerns about macro influencers who use high arousal language in posts. When macro influencers increase arousal by 10%, it reduces consumer engagement by 8.4%, implying 346 fewer likes and comments, on average. Yet macro influencers can use high arousal language. As we show, informative posts tend to be beneficial; a 34% stronger signal of an informative goal is associated with a 1.8% increase in engagement, equivalent to 108 additional likes and comments, on average. Counterbalancing the valence of the message (e.g., highlighting a negative side of the product discussed) similarly helps macro influencers seem more genuine in their endorsement, thus increasing engagement. Specifically, two-sided messages can mitigate the arousal penalty that macro influencers incur when it comes to engagement. Finally, macro influencers can use high arousal language if they also include words that signal trustworthiness (e.g., "learn"). Accordingly, brands and influencers should collaborate to make sure posts include phrases like "that's what I *learned* about this incredible product" rather than "that's how to *use* this incredible product."

Arousal for Micro Influencers. Micro influencers tend to prompt greater engagement (Weber 2022). We recommend that they use high arousal language to increase it even more. A 10% increase in arousal is associated with a 5.4% increase in engagement, on average, so for example, posting a recommendation such as "that's superb [vs. great]" would attract 49 additional likes or comments. In addition to our core findings, we offer insights for content composition. Table 6 provides a summary view of how the key drivers tend to affect engagement with various influencers.

Finally, we offer brands evidence that language arousal also plays a role on TikTok. Many brands hire influencers to spread information and influence among younger consumers who embrace TikTok, but both research and practice suffer from a limited understanding of what features make TikTok content more impactful (Haenlein et al. 2020).

Table 6: Summary of Effects on Engagement

	Micro	Macro
Influencer Characteristics		
1. Verified Account	+10.5%	+27.2%
2. Number of Posts Shared	–9.4%*	–14.9%*
3. Content Variation	ns	–3.7%*
Post Characteristics		
4. Arousal	+3.6%*	–5.6%*
5. Informative (vs. Commercial) Goal	ns	+1.8%*
6. Linguistic Concreteness	–4.2%*	ns
7. Word Familiarity	–3.2%*	ns
8. Image Emotionality	+3.4%*	+6.7%*

* Change in engagement corresponding to a one-standard deviation increase.
Notes: ns = non-significant effect.

Limitations and Further Research

Some limitations of this study might help stimulate additional research. Writing tends to allow people to express more emotional attitudes (Rocklage and Fazio 2015), but speaking offers additional ways to express emotionality, including auditory aspects. Although we examined the effect of influencers' speech pitch in TikTok videos, continued research could account for other vocal cues that potentially express arousal (e.g., tone of voice, tempo).

Including trustworthiness cues in posts, counterbalancing the valence of the message, and posting with an informative goal all can mitigate the negative effects of high arousal language for macro influencers. Continued work might identify other viable solutions. As we note, Cascio Rizzo et al. (2023) determine that sensory language increases engagement with macro influencers' content, because it suggests influencers have actually used the product. Arguably, macro influencers might use high arousal language effectively if they also incorporate sensory words.

Researchers also might examine if influencers' use of high arousal language stimulates prosocial behaviors. If influencers seem sincere, they might encourage and inspire followers. For brands seeking to enhance sustainable offerings, influencers who are more trustworthy but also use high arousal language might prompt increased consumption of sustainable products. Public institutions and nonprofit organizations also might leverage influencer effects to encourage responsible actions. The growing impacts and concerns surrounding global challenges (natural and humanitarian crises, technological disruptions, social movements) highlight the relevance of such considerations.

We focus on influencers, but high arousal language might determine the effectiveness of other spokespersons too. People often doubt politicians, question scientists during public emergencies, distrust education systems, or express wariness of brand claims (e.g., Kreps and Kriner 2020; Swaminathan et al. 2020). We call for further research into whether and how the effects of arousal extend to other contexts in which source characteristics inform people's judgments and decisions.

Our observational data cannot control for potential selection effects. Companies strategically choose certain influencers, on the basis of their characteristics or fit with the brand, so a selection bias appears likely. Our field data do not include data at the campaign level. The field analysis using propensity score matching, as well as the controlled lab experiments, suggest that selection effects do not drive the research outcomes, but we caution that self-selection could play a role in the observational data. We also call for caution regarding the effects sizes observed in the experiments. Our studies feature robust arousal manipulations, and we primed participants about the influencers' follower count. In reality, followers might pay little attention to an influencer's follower count, so the actual effect sizes could be more modest than what we observe.

Continued marketing research might examine the degree of arousal evoked by images too. We attempted such an analysis in Study 1, in which we extracted the emotional state exhibited by human faces, but this effort is clearly limited. The Google API we used retrieves only four emotions (anger, joy, sorrow, surprise). Images without human faces also can convey arousal, such as a picture of an appetizing meal. We encourage researchers to apply image mining in efforts to investigate how best to integrate text and visual content.

References

Ahluwalia, Rohini and Robert. E. Burnkrant (2004), "Answering questions about questions: A persuasion knowledge perspective for understanding the effects of rhetorical questions," *Journal of Consumer Research*, 31 (1), 26-42.

Ananthakrishnan, Uttara, Davide Proserpio, and Siddhartha Sharma (2023), "I hear you: Does quality improve with customer voice?," *Marketing Science* (published online April 6), https://doi.org/10.1287/mksc.2023.1437.

Bänziger, Tanja and Klaus S. Scherer (2005), "The role of intonation in emotional expressions," *Speech communication*, 46 (3-4), 252-267.

Becker, Jan-Michael, Dorian Proksch, and Christian M. Ringle (2022), "Revisiting Gaussian copulas to handle endogenous regressors," *Journal of the Academy of Marketing Science*, 50 (1), 46-66.

Berger, Jonah (2011), "Arousal increases social transmission of information," *Psychological Science*, 22 (7), 891-893.

Berger, Jonah, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel (2020), "Uniting the tribes: Using text for marketing insight," *Journal of Marketing*, 84 (1), 1-25.

Berger, Jonah, and Katherine L. Milkman (2012), "What makes online content viral?" *Journal of Marketing Research*, 49 (2), 192-205.

Berger, Jonah, Wendy W. Moe, and David Schweidel (2023), "What Holds Attention? Linguistic Drivers of Engagement," *Journal of Marketing*, 87 (5), 793-809.

Boegerhausen, Johannes, Hannes Datta, Abhishek Borah, and Andrew T. Stephen (2022), "Fields of gold: Scraping web data for marketing insights," *Journal of Marketing*, 86 (5), 1-20.

Boerman, Sophie C., Lotte M. Willemsen, and Eva P. Van Der Aa (2017), "This post is sponsored" effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook," *Journal of Interactive Marketing,* 38 (1), 82-92.

Borah, Abhishek, Sourindra Banerjee, Yu-Ting Lin, Apurv Jain, and Andreas B. Eisingerich (2020), "Improvised marketing interventions in social media," *Journal of Marketing*, 84 (2), 69-91.

Campbell, Margaret C., and Amna Kirmani (2000), "Consumers' use of persuasion knowledge: The effects of accessibility and cognitive capacity on perceptions of an influence agent," *Journal of Consumer Research*, 27 (1), 69-83.

Campbell, Margaret C. (1995), "When attention-getting advertising tactics elicit consumer inferences of manipulative intent: The importance of balancing benefits and investments," *Journal of Consumer Psychology*, 4 (3), 225-254.

Cascio Rizzo, Giovanni Luca, Jonah Berger, Matteo De Angelis, and Rumen Pozharliev (2023), "How Sensory Language Shapes Influencer's Impact," *Journal of Consumer Research*, (published online March 8), https://doi.org/10.1093/jcr/ucad017.

Chen, Li, Yajie Yan, and Andrew N. Smith (2022), "What drives digital engagement with sponsored videos? An investigation of video influencers' authenticity management strategies," *Journal of the Academy of Marketing Science*, 51 (1), 198-221.

Chung, Jaeyeon, Yu Ding, and Ajay Kalra (2023), "I Really Know You: How Influencers Can Increase Audience Engagement by Referencing Their Close Social Ties," *Journal of Consumer Research*, (published online March 14), https://doi.org/10.1093/jcr/ucad019.

Colicev, Anatoli, Ashwin Malshe, Koen Pauwels, and Peter O'Connor (2018), "Improving consumer mindset metrics and shareholder value through social media: The different roles of owned and earned media," *Journal of Marketing*, 82 (1), 37-56.

Craig, Tracy Y., and Kevin L. Blankenship (2011), "Language and persuasion: Linguistic extremity influences message processing and behavioral intentions," *Journal of Language and Social Psychology*, 30 (3), 290-310.

De Cheveigné, Alain, and Hideky Kawahara (2002), "YIN, a fundamental frequency estimator for speech and music," *The Journal of the Acoustical Society of America*, 111 (4), 1917-1930.

De Deyne, Simon, Danielle J. Navarro, Gullielm Collell, and Andrew Perfors (2021), "Visual and affective multimodal models of word meaning in language and mind," *Cognitive Science*, 45 (1), e12922.

Eisend, Martin, and Farid Tarrahi (2022), "Persuasion knowledge in the marketplace: A meta-analysis," *Journal of Consumer Psychology* 32 (1), 3-22.

Fast, Ethan, Binbin Chen, and Michael S. Bernstein (2016, May), "Empath: Understanding topic signals in large-scale text", In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 4647-4657).

Felbermayr, Armin and Alexandros Nanopoulos (2016), "The role of emotions for the perceived usefulness in online customer reviews," *Journal of Interactive Marketing*, 36 (1), 60-76.

Friestad, Marian, and Peter Wright (1994), "The persuasion knowledge model: How people cope with persuasion attempts," *Journal of Consumer Research*, 21(1), 1-31.

Gerrath, Maximilian H. E. E., and Bryan Usrey (2021), "The impact of influencer motives and commonness perceptions on follower reactions toward incentivized reviews," *International Journal of Research in Marketing*, 38 (3), 531-548.

Haan, Perry, and Cal Berkey (2002), "A study of the believability of the forms of puffery," *Journal of Marketing Communications,* 8 (4), 243-256.

Haenlein, Michael, Ertan Anadol, Tyler Farnsworth, Harry Hugo, Jess Hunichen, and Diana Welte (2020), "Navigating the new era of influencer marketing: How to be successful on Instagram, TikTok, & Co.," *California Management Review*, 63 (1), 5-25.

Hatton, Georgia (2018), "Micro influencers vs macro influencers," (accessed January 13, 2023), https://www.socialmediatoday.com/news/micro-influencers-vs-macro-influencers/516896/.

Hayes, Andrew F. (2018), "Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation," *Communication Monographs*, 85 (1), 4-40.

Herhausen, Dennis, Stephan Ludwig, Dhruv Grewal, Joechen Wulf, and Marcus Schoegel (2019), "Detecting, preventing, and mitigating online firestorms in brand communities," *Journal of Marketing*, 83 (3), 1-21.

Herhausen, Dennis, Lauren Grewal, Krista Hill Cummings, Anne L. Roggeveen, Francisco Villarroel Ordenes, and Dhruv Grewal (2023), "Complaint De-Escalation Strategies on Social Media," *Journal of Marketing*, 87 (2), 210-231.

Hotmart (2022), "What are micro influencers, and how can they help your business grow," (accessed May 23, 2023), https://hotmart.com/en/blog/micro-influencers.

Hovy, Dirk, Shiri Melumad, and J. Jeffrey Inman (2021), "Wordify: a tool for discovering and differentiating consumer vocabularies," *Journal of Consumer Research*, 48 (3), 394-414.

Hughes, Christian, Vanitha Swaminathan, and Gillian Brooks (2019), "Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns," *Journal of Marketing*, 83 (5), 78-96.

Humphreys, Ashlee, and Rebecca Jen-Hui Wang (2018), "Automated text analysis for consumer research," *Journal of Consumer Research*, 44 (6), 1274-1306.

Humphreys, Ashlee, Mathew Isaac, and Rebecca Jen-Hui Wang (2021), "Construal matching in online search: applying text analysis to illuminate the consumer decision journey," *Journal of Marketing Research*, 58 (6), 1101-1119.

Influencer Marketing Hub (2021), "9 factors to consider to pick the right Instagram influencer for your brand," (accessed April 11, 2022), https://influencermarketinghub.com/pick-the-right-instagram-influencer/.

Jalali, Nima Y, and Purushottam Papatla (2019), "Composing tweets to increase retweets," *International Journal of Research in Marketing*, 36 (4), 647-668.

Kanuri, Vamsi K., Yixing Chen, and Shrihari Sridhar (2018), "Scheduling content on social media: Theory, evidence, and application," *Journal of Marketing*, 82 (6), 89-108.

Karagür, Zeynep, Jan-Michael Becker, Kristina Klein, and Alexander Edeling (2022), "How, why, and when disclosure type matters for influencer marketing," *International Journal of Research in Marketing*, 39 (2), 313-335.

Kralj Novak, Petra, Jasmina Smailović, Borut Sluban, and Igor Mozetič (2015), "Sentiment of emojis," *PloS One*, 10 (12), e0144296.

Kreps, Sarah E., and Douglas L. Kriner (2020), "Model uncertainty, political contestation, and public trust in science: Evidence from the COVID-19 pandemic," *Science Advances*, 6 (43), eabd4563.

Kumar, Ashish, Ram Bezawada, Rishika Rishika, Ramkumar Janakiraman, and P. K. Kannan (2016), "From social to sale: The effects of firm-generated content in social media on customer behavior," *Journal of Marketing*, 80 (1), 7-25.

Kuperman, Victor, Zachary Estes, Marc Brysbaert, and Amy Beth Warriner (2014), "Emotion and language: valence and arousal affect word recognition," *Journal of Experimental Psychology: General*, 143 (3), 1065.

Kupfer, Ann-Kristin, Nora Pähler Vor der Holte, Raoul V. Kübler, and Thorsten Hennig-Thurau (2018), "The role of the partner brand's social media power in brand alliances," *Journal of Marketing*, 82 (3), 25-44.

Kuppens, Peter, Francis Tuerlinckx, James A. Russell, and Lisa Feldman Barrett (2013), "The relation between valence and arousal in subjective experience," *Psychological Bulletin*, 139 (4), 917.

Laukka, Petri, Hillary Anger Elfenbein, Nutankumar S. Thingujam, Thomas Rockstuhl, Frederick K. Iraki, Wanda Chui, and Jean Althoff (2016), "The expression and recognition of emotions in the voice across five nations: A lens model analysis based on acoustic features," *Journal of Personality and Social Psychology*, 111 (5), 686.

Lee, Jeffrey K., and Enric Junqué De Fortuny (2021), "Influencer-generated reference groups," *Journal of Consumer Research*, 49 (1), 25-45.

Lee, Dokyoun, Kartik Hosanagar, and Harikesh S. Nair (2018), "Advertising content and consumer engagement on social media: Evidence from Facebook," *Management Science*, 64 (11), 5105-5131.

Leung, Fine F., Flora F. Gu, and Robert W. Palmatier (2022), "Online influencer marketing," *Journal of the Academy of Marketing Science*, 50 (2), 226-251.

Leung, Fine F., Flora F. Gu, Yiwei Li, Jonathan Z. Zhang, and Robert W. Palmatier (2022), "Influencer Marketing Effectiveness," *Journal of Marketing*, 86 (6), 93-115.

Li, Yiyi, and Ying Xie (2020), "Is a picture worth a thousand words? An empirical study of image content and social media engagement," *Journal of Marketing Research*, 57 (1), 1-19.

Liadeli, Georgia, Francesca Sotgiu, and Peetr W. Verlegh (2023), "A meta-analysis of the effects of brands' owned social media on social media engagement and sales," *Journal of Marketing*, 87 (3), 406-427.

Luangrath, Andrea W., Joann Peck, and Victor A. Barger (2017), "Textual paralanguage and its implications for marketing communications," *Journal of Consumer Psychology*, 27 (1), 98-107.

Luangrath, Andrea W., Xu, Yixiang, and Wang, Tong (2023), "Paralanguage classifier (PARA): An algorithm for automatic coding of paralinguistic nonverbal parts of speech in text," *Journal of Marketing Research*, 60 (2), 388-408.

Ludwig, Stephan, Dennis Herhausen, Dhruv Grewal, Liliana Bove, Sabine Benoit, Ko De Ruyter, and Peter Urwin (2022), "Communication in the gig economy: Buying and selling in online freelance marketplaces," *Journal of Marketing*, 86 (4), 141-161.

Mauss, Iris B. and Michael D. Robinson (2009), "Measures of emotion: A review," *Cognition and emotion*, 23 (2), 209-237.

Michaeloudis, Lucy (2021), "Our top influencer marketing trends for 2021," (accessed January 22, 2022), https://www.upfluence.com/influencer-marketing/kick-start-your-campaigns-with-our-top-influencer-marketing-trends-for-2021.

Mohammad, Saif (2018, July), "Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words," In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 174-184).

Mohammad, Saif (2017), "Word affect intensities," *arXiv preprint arXiv:1704.08798*.

Moore, Sarah G., and Katherine C. Lafreniere (2020), "How online word-of-mouth impacts receivers," *Consumer Psychology Review*, 3 (1), 34-59.

Ohanian, Roobina (1990), "Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness," *Journal of Advertising*, 19 (3), 39-52.

Packard, Grant, and Jonah Berger (2021), "How concrete language shapes customer satisfaction," *Journal of Consumer Research*, 47 (5), 787-806.

Packard, Grant, Andrew D. Gershoff, and David B. Wooten (2016), "When boastful word of mouth helps versus hurts social perceptions and persuasion," *Journal of Consumer Research*, 43 (1), 26-43.

Packard, Grant, Sarah G. Moore, and Brent McFerran (2018), "(I'm) happy to help (you): The impact of personal pronoun use in customer–firm interactions," *Journal of Marketing Research*, 55 (4), 541-555.

Packard, Grant and Jonah Berger (2024), "The Emergence and Evolution of Consumer Language Research," *Journal of Consumer Research*, (published online February 17), https://doi.org/10.1093/jcr/ucad013.

Paetzold, Gustavo, and Lucia Specia (2016, June), "Inferring psycholinguistic properties of words," *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies* (pp. 435-440).

Pancer, Ethan, Vincent Chandler, Maxwell Poole, and Theodore Noseworthy (2019), "How readability shapes social media engagement," *Journal of Consumer Psychology*, 29 (2), 262-270.

Petrin, Amil, and Kenneth Train (2010), "A control function approach to endogeneity in consumer choice models," *Journal of Marketing Research*, 47 (1), 3-13.

Pogacar, Ruth, L. J. Shrum, and Tina M. Lowrey (2018), "The effects of linguistic devices on consumer information processing and persuasion: a language complexity× processing mode framework," *Journal of Consumer Psychology*, 28 (4), 689-711.

Pornpitakpan, Chanthika (2004), "The persuasiveness of source credibility: A critical review of five decades' evidence," *Journal of Applied Social Psychology*, 34 (2), 243-281.

Pozharliev, Rumen, Dario Rossi, and Matteo De Angelis (2022), "Consumers' self-reported and brain responses to advertising post on Instagram: the effect of number of followers and argument quality," *European Journal of Marketing*, 56 (3), 922-948

Rocklage, Matthew D., and Russell H. Fazio (2015), “The Evaluative Lexicon: Adjective use as a means of assessing and distinguishing attitude valence, extremity, and emotionality,” *Journal of Experimental Social Psychology*, 56, 214-227.

Santora, Jacinda (2022), “Key influencer marketing statistics you need to know for 2022,” (accessed February 14, 2022), https://influencermarketinghub.com/influencer-marketing-statistics/.

Schwartz, Janet, Mary Frances Luce, and Dan Ariely (2011), “Are consumers too trusting? The effects of relationships with expert advisers,” *Journal of Marketing Research*, 48 (SPL), S163-S174.

Swaminathan, Vanitha, Alina Sorescu, Jan-Benedict E.M. Steenkamp, Thomas Clayton Gibson O’Guinn, and Bernd Schmitt (2020), “Branding in a hyperconnected world: Refocusing theories and rethinking boundaries,” *Journal of Marketing*, 84 (2), 24-46.

Tellis, Gerard J., Deborah J. MacInnis, Seshadri Tirunillai, and Yanwei Zhang (2019), “What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence,” *Journal of Marketing*, 83 (4), 1-20.

Uribe, Rodrigo, Buzeta, Christina, and Velásquez, Milenka (2016), “Sidedness, commercial intent and expertise in blog advertising,” *Journal of Business Research*, 69 (10), 4403-4410.

Valsesia, Francesca, Davide Proserpio, and Joseph Nunes (2020), “The positive effect of not following others on social media,” *Journal of Marketing Research*, 57 (6), 1152-1168.

Villarroel Ordenes, Francisco, Dhruv Grewal, Stephan Ludwig, Ko De Ruyter, Dominik Mahr, and Martin Wetzels (2019), “Cutting through content clutter: How speech and image acts drive consumer sharing of social media brand messages,” *Journal of Consumer Research*, 45 (5), 988-1012.

Villarroel Ordenes, Francisco, Stephan Ludwig, Ko De Ruyter, Dhruv Grewal, and Martin Wetzels (2017), “Unveiling what is written in the stars: Analyzing explicit, implicit, and discourse patterns of sentiment in social media,” *Journal of Consumer Research*, 43 (6), 875-894.

Weber, Natalie (2022), “How to make – or break – your brand with micro influencer marketing,” (accessed June 10, 2022), https://nealschaffer.com/micro-influencer-marketing/.

Weingarten, Evan, and Jonah Berger (2017), “Fired up for the future: How time shapes sharing,” *Journal of Consumer Research*, 44 (2), 432-447.

Whissell, Cynthia (2009), “Using the revised dictionary of affect in language to quantify the emotional undertones of samples of natural language,” *Psychological Reports*, 105 (2), 509-521.

Wies, Simone, Alexander Bleier, and Alexander Edeling (2023), “Finding Goldilocks influencers: How follower count drives social media engagement,” *Journal of Marketing,* 87 (3), 383-405.

Williams, Patti, Gavan J. Fitzsimons, and Lauren G. Block (2004), "When consumers do not recognize “benign” intention questions as persuasion attempts," *Journal of Consumer Research,* 31 (3), 540-550.

Yin, Dezhi, Samuel D. Bond, and Han Zhang (2017), “Keep your cool or let it out: Nonlinear effects of expressed arousal on perceptions of consumer reviews,” *Journal of Marketing Research*, 54 (3), 447-463.

How High Arousal Language Shapes Micro versus Macro Influencers' Impact

Author Note

Giovanni Luca Cascio Rizzo is a doctoral student in marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (glcasciorizzo@luiss.it). Francisco Villarroel Ordenes is an assistant professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (fvillarroel@luiss.it). Rumen Pozharliev is assistant professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (rpozharliev@luiss.it). Matteo De Angelis is a professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (mdeangelis@luiss.it). Michele Costabile is a professor of marketing at the LUISS Guido Carli University, viale Romania 32, 00198, Rome, Italy (mcostabile@luiss.it). Please address correspondence to Giovanni Luca Cascio Rizzo. This article is based on the lead author's dissertation.

Acknowledgements

The authors acknowledge their collaboration on this research with Buzzoole S.p.a.. They also thank Ann Kronrod, Andrea Luangrath, Stephan Ludwig, and Stefano Puntoni for their valuable feedback, and participants at the EMAC 2023, Rotterdam School of Management lunchclub, and SIM doctoral consortium 2022 for helpful comments.

Data collection information

The field data for Study 1 and its follow-up study were collected in winter 2021 by the first author. The data for Studies 2, 3a, and 3b were collected by the fourth author in fall 2022 and spring 2023. The exploratory study was collected by the third author in spring 2023. Study 4 was collected by the second author in winter 2022. All experiments used Prolific participants located in the United States and were designed by the first author. The field data measures of Study 1 were developed by the first and second authors. Analysis for all studies was performed by the first author. The data are currently stored in a project directory on the Open Science Framework.

WEB APPENDIX

How High Arousal Language Shapes Micro versus Macro Influencers' Impact

Table of Contents

These materials have been supplied by the authors to aid in the understanding of their paper.

The AMA is sharing these materials at the request of the authors.

Web Appendix A. Literature on Influencer Marketing

Author(s)	Objective	Main Findings
Ansari, Asim, Floran Stahl, Mark Heitmann, and Lucas Bremer (2018), "Building a Social Network for Success," *Journal of Marketing Research*, 55 (3), 321-338.	Model how musical artists can enhance their social networking presence and stimulate relationships between fans to achieve long-term benefits in terms of music plays on a European online social networking site.	Artists can influence the structure of their ego network (a central actor, the friends of the actor, and all of their friends) and drive song plays over the long run by actively sending friend requests or comments to fans.
Cascio Rizzo, Giovanni Luca, Jonah Berger, Matteo De Angelis, and Rumen Pozharliev (2023), "How Sensory Language Shapes Influencer's Impact," *Journal of Consumers Research*, forthcoming.	Examine how sensory language affect purchase and engagement with influencer-sponsored content.	Sensory language makes people believe that influencers have actually used the product, which increases perceived authenticity, which in tun boosts purchase and engagement.
Chen, Li, Yajie Yan, and Andrew N. Smith (2022), "What drives digital engagement with sponsored videos? An investigation of video influencers' authenticity management strategies," *Journal of the Academy of Marketing Science*, 1-24.	Conceptualize and test a framework, involving passion and transparency-based strategies as well as platforms and brand factors, to determine how influencers can manage their authenticity.	Disclosing brand sponsorship, early brand appearance, high video customization, and sharing personal experiences or opinions about the sponsored product, all affect engagement.
Chen, Xi, Ralf Van Der Lans, and Tuan Q. Phan (2017), "Unncovering the importance of relationship characteristics in social networks: Implications for seeding strategies," *Journal of Marketing Research*, 54 (2), 187-201.	Identify influential network members and considers the relative influence of different relationship characteristics on product diffusion.	Development of a multinetwork approach for activating influencers by inferring network connection weights based on features like recency and interaction intensity, as well as dissemination process. The relationship duration and private message exchanges generate a multinetwork extending beyond connections alone.
Chung, Jaeyeon, Yu Ding, and Ajay Kalra (2023), "I Really Know You: How Influencers Can Increase Audience Engagement by Referencing Their Close Social Ties," *Journal of Consumer Research*, forthcoming.	Examine how influencer posting photos with or about people whom influencers share close ties with boost engagement.	Sharing stories with close people make influencers seem more authentic, similar, and warm. This, in turn, increases consumer engagement (i.e., likes).
Goldenberg, Jacob, Gal Oestreicher-Singer, and Shachar Reichman (2012), "The Quest for Content: How User-Generated Links Can Facilitate Online Exploration," *Journal of Marketing Research*, 49 (4), 452-468.	Investigate the role of dual-network structure and specifically of user-generated links in facilitating content exploration.	Exposure to the dual network results in a more efficient (time to desirable outcome) and more effective (average product rating, overall satisfaction) exploration process.
Hinz, Oliver, Bernd Skiera, Christina Barrot, and Jan U. Becker (2011), "Seeding Strategies for Viral Marketing: An Empirical Comparison," *Journal of Marketing*, 75 (6), 55-71.	Compare four seeding strategies: those targeting "hubs," people with a high number of connections; "fringes," people poorly connected; "bridges," people who connect two otherwise unconnected parts of the network; and random people.	The best strategies (i.e., those that achieve the highest number of referrals) target the message to hubs (high-degree seeding) or bridges (high-betweenness seeding).
Hughes, Christian, Vanitha Swaminathan, and Gillian Brooks (2019), "Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns," *Journal of Marketing*, 83 (5), 78-96.	Examine the factors in sponsored blogging (source and content characteristics) that drive success of online brand engagement at different stages of the consumer purchase funnel.	When a sponsored post occurs on a blog, high blogger expertise is more effective when the ad intent is to raise awareness (vs. trial), whereas fails to drive engagement on Facebook. On Facebook, posts high in hedonic content are more effective when the ad intent is to increase trial (vs. awareness). Campaign incentives increase (decrease) engagement on blogs (Facebook).

Katona, Zsold, Peter Pal Zubcsek, and Miklos Savary (2011), "Network Effects and Personal Influences: The Diffusion of an Online Social Network," *Journal of Marketing Research*, 48 (3), 425-443.	Uncover the effects of differences in individuals' connection patterns within a social network on the diffusion process (network adoption). The authors look at the network structure, individual characteristics of adopted neighbors (influencers) and characteristics of potential adopters.	The number and interconnectedness of already adopted friends has a positive effect on the probability of an individual's adoption. People with many friends have a lower average influence than those with fewer friends. Some demographic characteristics also play a role.
Kumar Viswanathan, Vikram Bhaskaran, Rohan Mirchandani and Milap Shah (2013), "Practice Prize Winner—Creating a Measurable Social Media Marketing Strategy: Increasing the Value and ROI of Intangibles and Tangibles for Hokey Pokey," *Marketing Science*, 32 (2), 191-363.	Creation of a unique metric to measure the net influence wielded by a user in a social network, customer influence effect (CIE), and predicting the user's ability to generate the spread of viral information. Creation of a second metric, customer influence value (CIV) to link WOM to actual sales.	Development and validation of CIE (an extension of extend Hubbell's influence measure based on tracking the spread of a message) and CIV (calculated by iteratively summing the CLV of all the people influenced by the Individual) metrics.
Lanz, Andrea, Jacob Goldenberg, Daniel Shapira, and Florian Stahl (2019), "Climb or Jump: Status-Based Seeding in User-Generated Content Networks," *Journal of Marketing Research*, 56 (3), 361-378.	Investigate how music creators can increases exposure to their content by expanding the follower base through direct outbound activities.	Unknown music creators should gradually build their status by targeting low-status users rather than attempt to "jump" by targeting high-status ones.
Lee, Jeffrey K., and Enric Junqué De Fortuny (2021), "Inluencer-Generated Reference Groups," *Journal of Consumer Research,* 49 (1), 25-45.	Explore how consumer influencers can shape reference group meanings in social media.	The typicality of the influencer (relative to a brand's stereotypical consumer) can shape ideas about the perceived homogeneity of the brand's consumers, which ultimately influences the strength and tightness of brand associations.
Leung, Fine F., Flora F. Gu, Yiwei Li, Jonathan Z. Zhang, and Robert W. Palmatier (2022), "Influencer Marketing Effectiveness," *Journal of Marketing*, 86 (6), 93-115.	Examine how factors related to the influencer, influencer's followers, and influencer's posts determine influencer marketing effectiveness.	Influencer originality, follower size, and sponsor salience enhance engagement. Influencer activity, follower-brand fit, and post positivity exert inverted U-shaped moderating effect on engagement.
Pei, Amy, and Dina Mayzlin (2021), "Influencing Social Media Influencers Through Affiliation," *Marketing Science*.	Investigate what is the optimal level of affiliation with influencers from the firm's perspective, and what is the impact of affiliation on consumer welfare.	When the consumer's prior belief is low, the firm needs to affiliate less closely or not at all to preserve influencer persuasiveness. In contrast, when the consumer's prior belief is high, the firm fully affiliates with the influencer to both maximize awareness and prevent a negative review.
Trusov, Michael, Anand V. Bodapati, and Randolph E. Bucklin (2010), "Determining Influential Users in Internet Social Networks," *Journal of Marketing Research*, 47 (4), 643-658.	Develop an approach to determine which users have significant effects on the activities of using the longitudinal records of members' log-in activity	Develop a nonstandard Bayesian shrinkage approach to calculating influence scores. The method extracts, with limited data, the strong links from a large overt network that has mostly weak links.
Valsesia, Francesca, Davide Proserpio, and Joseph Nunes (2020), "The positive effect of not following others on social media," *Journal of marketing research*, 57 (6), 1152-1168.	Investigate whether a visual cue, like an influencer's number of followings, helps to distinguish more vs. less effective influencers on social media.	Following fewer others, conditional on having a substantial number of followers, conveys greater autonomy, a signal of influence which make consumers engage more with the post.
Wies, Simone, Bleier Alexander, and Edeling, Alexander (2023), "Finding Goldilocks Influencers: How Follower Count Drives Social Media Engagement," *Journal of Marketing*, 87 (3), 383-405.	Examine how influencers' follower count shape consumer engagement with sponsored content.	Engagement increases, then decreases, as influencer follower count rises (inverted U-shaped relationship). This effect is driven by perceptions of tie strength. Higher content customization and lower brand familiarity flatten the relationship.

Web Appendix B. Study 1: Language Arousal in the Field

Table WB1. Sample Description

Industry	Infl. Type	Influencers (#)	Posts (#)	Avg. Post Likes (SD)	Avg. Post Comments (SD)
Architecture & Design	Micro	42	805	1,143 (880)	108 (111)
	Macro	46	898	4,904 (4,054)	197 (363)
Art & Culture	Micro	44	558	917 (2,431)	38 (44)
	Macro	39	517	3,118 (6,777)	97 (292)
Beauty	Micro	39	805	909 (1,206)	71 (80)
	Macro	37	504	7,501 (8,178)	136 (177)
Economics	Micro	38	577	814 (631)	38 (45)
	Macro	32	168	3,069 (3,491)	66 (76)
Environment & Ecology	Micro	25	231	760 (890)	50 (54)
	Macro	11	114	4,235 (2,225)	58 (137)
Family & Parenting	Micro	39	869	952 (1,328)	55 (49)
	Macro	36	674	8,873 (12,957)	231 (846)
Fashion	Micro	62	1,042	1,242 (1,065)	90 (59)
	Macro	63	1,467	4,928 (4,702)	169 (254)
Food & Drinks	Micro	60	1,005	801 (864)	58 (72)
	Macro	61	1,278	4,296 (5,069)	188 (509)
Gaming	Micro	22	198	992 (1,753)	18 (24)
	Macro	28	183	4,548 (12,244)	68 (71)
Health & Wellness	Micro	51	971	910 (1,761)	49 (75)
	Macro	43	541	7,537 (7,298)	154 (503)
Hobbies & Interests	Micro	32	329	1,040 (1,360)	39 (46)
	Macro	40	372	12,806 (12,611)	127 (213)
Home & Gardening	Micro	27	419	978 (1,138)	31 (49)
	Macro	16	277	4,178 (9,283)	94 (278)
Lifestyle	Micro	64	1,195	1,955 (1,640)	73 (84)
	Macro	59	933	6,961 (6,571)	189 (446)
Media & Entertainment	Micro	36	418	1,812 (1,407)	43 (53)
	Macro	35	294	12,192 (12,883)	139 (289)
News & Society	Micro	36	328	630 (851)	41 (77)
	Macro	26	164	2,981 (3,960)	98 (216)
Sport & Fitness	Micro	44	578	998 (1,160)	34 (58)
	Macro	35	264	8,133 (7,715)	143 (333)
Tech & Science	Micro	39	630	1,125 (1,565)	70 (75)
	Macro	27	247	3,872 (4,261)	65 (89)
Travel & Tourism	Micro	17	367	1,164 (1,121)	48 (52)
	Macro	49	755	7,018 (8,105)	150 (191)

Table WB2. Descriptive Statistics and Correlations

Variable	M	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
(1) Engagement	3,538	6,167	1.000																					
(2) Arousal	.50	.07	-.020*	1.000																				
(3) Macro	.46	.50	.416*	.019*	1.000																			
(4) If Trial	2.61	.64	-.007	.020*	.008	1.000																		
(5) if Verified	.29	.45	.266*	.010	.470*	-.012	1.000																	
(6) # of Posts	41.9	43.27	-.099*	.003	-.146*	-.025*	-.138*	1.000																
(7) Content Variation	.23	.01	-.030*	-.003	-.002	-.052*	-.062*	.308*	1.000															
(8) # of Questions	.53	.90	-.002	.013	-.002	.039*	.021*	-.017*	-.068*	1.000														
(9) # of Hashtags	7.60	8.42	-.036*	.017*	-.072*	-.020*	-.060*	.132*	.080*	.072*	1.000													
(10) # of Mentions	1.82	1.60	.004	-.015*	-.002	.038*	-.006	.022*	.041*	.058*	.070*	1.000												
(11) Word Count	113.1	71.1	-.003	.016*	-.017*	.170*	.016*	-.028*	-.128*	.291*	.141*	.199*	1.000											
(12) # of Emojis	1.90	2.75	-.008	.007	-.026*	.081*	-.024*	-.088*	.000	.127*	.107*	.130*	.229*	1.000										
(13) Complexity	11.17	6.37	-.028*	.008	-.057*	-.172*	-.050*	.162*	.183*	-.147*	.488*	.064*	-.100*	.032*	1.000									
(14) Valence	.67	.06	.010	.016*	.012	.144*	-.010	.051*	.015*	-.028*	-.006	.024*	-.091*	-.005	.006	1.000								
(15) Concreteness	351.5	21.3	-.017*	-.007	-.020*	-.062*	-.010	.077*	.093*	-.064*	.099*	.048*	.058*	.040*	.212*	-.099*	1.000							
(16) Familiarity	576.5	15.9	.008	-.021*	.021*	.067*	.007	-.095*	-.061*	.019*	-.101*	-.071*	-.228*	-.068*	-.289*	.137*	-.460*	1.000						
(17) if Image	.91	.29	.008	-.042*	-.103*	.008	-.142*	.112*	.051*	.000	.033*	.010	.010	-.010	.033*	.012	.015*	-.007	1.000					
(18) if Face Present	.67	.47	.007	.031*	.008	.013	.001	-.021*	.032*	.016*	-.140*	-.009	-.051*	.009	-.092*	.063*	-.159*	.128*	.015*	1.000				
(19) Image Emotionality	4.4	3.38	.070*	.091*	-.010	.004	.018*	.001	-.002	.014*	-.027*	.004	-.004	.000	-.023*	.017*	-.037*	.025*	.252*	.152*	1.000			
(20) Color Dominance	.72	.21	-.010	.001	-.003	.026*	-.011	.027*	-.005	.015*	.039*	-.008	.046*	.031*	.021*	.019*	.058*	-.033*	.019*	-.109*	-.022*	1.000		
(21) Color Saturation	.21	.20	.019*	.004	.011	-.002	-.002	-.011	-.031*	.004	-.027*	.006	.008	.018*	-.015*	.005	-.022*	-.006	-.011	.007	.012	-.327*	1.000	
(22) Time Difference	.07	.26	-.044*	-.019*	-.044*	-.009	-.039*	.268*	.072*	.004	.036*	.011	-.009	-.014*	.045*	.009	.022*	-.017*	.023*	-.006	-.034*	-.002	-.003	1.000

* $p < .05$

Notes: Fixed effects and topics are not included.

Table WB3. Most Frequent Emojis and Corresponding Arousal Score

Emoji	Frequency	Arousal
✨	2603	0,46
❤️	1060	0,83
😍	1051	0,78
😂	824	0,72
💕	695	0,72
🥰	645	0,76
🙌🏻	628	0,76
🙌🏼	579	0,76
😉	535	0,66
☀️	462	0,52
🎄	428	0,66
👏🏻	419	0,73
🔥	406	0,39
📸	399	0,65
🎉	395	0,82
💗	366	0,76
🤩	365	0,78
🤗	334	0,76
🤣	322	0,72
😋	318	0,72
🌿	306	0,58
💛	306	0,8
💙	302	0,8
⭐	288	0,65
🙌	288	0,76
☺️	282	0,55
💫	267	0,65
😆	267	0,72
🥳	264	0,81
🎁	263	0,84
🌱	254	0,58
😁	248	0,7
💚	244	0,8
🖤	237	0,8
😅	236	0,76
🌟	233	0,45
1️⃣	232	0,37
2️⃣	231	0,37
❄️	219	0,7
💜	217	0,8
☕	212	0,57
✅	205	0,56
👏🏼	200	0,73

New Dictionaries for Informative and Commercial Post Goal

Informative Words*: announcement, aware*, benefits, check*, detail, explor*, hear*, inform*, ingredients, knew, know*, learn, listen, post, question*, reason, reasons, stories, story, blog*, live, remember, watch, tell, launch, launched, discover, produced, designed, content, show, showing, showed, details, detailed, deets, search, searching, searched, announce*, read, browse, find, contain.*

Commercial Words*: ad, ads, advertis*, blackfriday, bought, buy*, click, coupon, deal*, discount*, giftcard*, item, items, offer, order, paid, pay*, price, product, products, promo, promote, promotion*, purch*, sale, save, shop*, try, visit, prices, sales, code, off, offs, checkout, gift card, gift cards, get yours, subscribe, subscription, subscribed, deals, black friday, cybermonday, cyber monday, givaway, givaways, give away, give away*

Validity in a Classification task. We followed a process similar to Ludwig et al. (2022, p. 149) to further ensure the validity of our informative goal measure. Specifically, we assessed the ability of our automated measure to solve a binary classification problem (informative vs. commercial goal). Our automated measure of informative goal assigns to each post a score ranging from 0 to 1; values higher than 0.5 show a greater informative goal, and values below 0.5 show a high commercial goal. To mimic a classification problem, we assigned values higher than 0.5 to an informative category, and values below 0.5 to a commercial category. Then, we followed a similar process to convert coders' ratings into a binary variable. Results find an overall classification accuracy of .82, confirming a good predictive ability (see Table WB4).

Table WB4. Results from Binary Classification Problem

Class	Recall	Precision	F-measure	Accuracy
Informative	.76	.71	.74	
Commercial	.85	.88	.86	
Overall				.82

Content Variation

More semantic variation across an influencer's posts is an indicative of diversity in the content that is being posted, which can be associated with generalists. Less semantic variation is an indicative of concentration or specialization in the type of content that is being posted. In line with recent NLP, we used word embeddings (Word2Vec) to represent influencer posts as semantic vectors, because they take into account the meaning of words in context (e.g., chocolate and candies; Berger et al. 2022). Before using embeddings, we used standard preprocessing steps such as: removed stop words, punctuation, numbers, words with less than two characters, URLs, and emojis. We used Word2Vec with the following parameters:

- A learning rate of .025
- Minimum learning rate of .004
- Layer size of 10
- Number of Epochs 5
- Number of Training Iterations 1
- Context Window of words 5
- Minimum word frequency 4
- Negative Sampling Rate 5.0
- CBOW algorithm

After extracting the word embeddings for each influencer, we used cosine distance (validated with Euclidian distance) to measure the semantic distance across his/her post. Finally, we computed the standard deviation across the word embeddings of each influencer post, which provides an indicative of variation (higher values) or concentration (lesser values) in the semantic content across posts.

Image Emotionality

Following Li and Xie (2020), we used the function “face detection” offered by Google Cloud Vision API to detect the emotional state of a human face when present within the image. The face detection service aims at mapping a human face to four emotional states (joy, sorrow, anger, and surprise), and scaling it on likelihood of values (“very unlikely”, “unlikely”, “possible”, “likely”, “very unlikely”) depending on the confidence percentage. If no specific emotion is detected, the “very unlikely” label is used. We also noted that such ratings indicate how strongly a particular emotional state appears in the image. For instance, a score of 4 on the joy scale portrays a greater happiness of a 3. Thus, we made the sum of joy, sorrow, anger, and surprise scores to account for image emotionality. We make two specifications. First, the face detection service localizes multiple faces and emotion estimates are returned for each detected face. When an image featured more faces, we averaged emotion scores and then we made the sum. Second, when a post featured multiple images we computed the image emotionality for each image, and we considered the maximum emotion score across all images to get a measure of image emotionality at post level.

Paralanguage Features from PARA

Influencers' posts can feature multiple types of paralanguages (e.g., vocal aspects conveying tempo: "amazingggg"; alphanumeric letters and symbols: "*high-five*"; tactile emojis: the hug emoji). So, we used the paralanguage classifier (PARA; Luangrath, Xu, and Wang 2023) to detect nonverbal communication cues in influencers' posts and included them as controls in the full model.

Table WB5. PARA Results

	DV: ENGAGEMENT
IV	
Arousal	1.035** (.008)
Macro (vs. Micro)	4.515** (.065)
Arousal × Macro	.913** (.011)
Controls	Included
Pitch	1.001 (.006)
Rhythm	1.006 (.006)
Stress	.974** (.006)
Emphasis	1.013* (.007)
Tempo	1.010 (.006)
Volume	.993 (.006)
Censorship	.991 (.005)
Spelling	.987* (.006)
Alternant	1.006 (.006)
Differentiatior	.990 (.006)
Alphahaptics	1.006 (.007)
Alphakinesics	1.006 (.006)
Formatting	.997 (.006)
Tactile	.996 (.006)
Bodily	.987* (.006)
Nonbodily	1.033** (.007)
N	20,590
Log likelihood	–180,084

* $p < .05$, ** $p < .01$.
Notes: Standard errors are in parentheses. All controls from the full model are included, but not reported for parsimony.

Addressing Endogeneity

Arousal

Influencers can adjust their content based on exogenous factors or posting characteristics in a preceding post. Thus, to accommodate such potential source of endogeneity, we adopt a control function (CF) approach (Petrin and Train 2010) which has been already used in marketing research (Kumar, Choi, and Greene 2017; George, Kumar, and Grewal 2013). The correlation between the endogenous variable and unobserved (omitted) variables is the cause for endogeneity. Thus, the idea behind the CF approach is to derive the part of the endogenous variable that depends on the unobserved variables in the first stage regression, and then include fitted residuals into the main response function in the second stage. In doing so, the fitted residuals capture the omitted variables that make our focal variable arousal endogenous. By including this term in the main response function, we can control for endogeneity, and obtain correct(ed) estimates of the coefficients (Imbens and Wooldridge 2007).

We applied the control function sequentially. In the first stage we regressed arousal on an unobserved variable, that is lagged arousal, and two exogenous instruments, that are "Holiday" and "Announcement". The rationale for including these two instruments is the following. Prior research (Nguyen et al. 2012) and field data observation suggest that influencers use to increase their arousal when the post is related to an incoming holiday event. Thus, we dummy coded "Holiday" (= 1 if a holiday is mentioned; = 0 otherwise) if the post mentions major holidays such as Christmas, New Year, Easter, Halloween, Thanksgiving, St. Patrick's Day, and Valentine's Day. Second, field observation suggests that influencers use to share announcements about new brand partnerships and achievements with higher levels of arousal. We accounted for and assessed "Announcement" via the corresponding dictionaries related to affiliation and achievement words from the Pennebaker et al.'s (2015) Linguistic Inquiry and Word Count (i.e., LIWC). Specifically, the standard output includes the percentage of words in the text pertaining to this variable. Note that, consistent with instruments validity criteria (Angrist and Pischke 2010), both instruments relate to arousal (holiday: $r = .016$; $p = .017$; announcement: $r = .168$; $p < .001$), but not to engagement (holiday: $r = .013$; $p = .151$; announcement: $r = .003$; $p = .575$).

Thus, we express Arousal as a function of lagged Arousal and instruments as follows:

$$\textit{Arousal} = \beta_0 + \beta_1 \text{ Arousal}_{(t-1)} + \beta_2 \text{ Holiday} + \beta_3 \text{ Announcement} + \varepsilon. \quad (1)$$

After estimating the first stage regression with OLS in Equation 1, we computed fitted residuals τ, and in the second stage we included them in the main response function in Equation 2:

$$\textit{Engagement} = \beta_1 \text{ Arousal} + \beta_2 \text{ Macro} + \beta_3 \text{ Arousal} \times \text{Macro} + \mathbf{X'}\,\gamma + \tau + \varepsilon, \quad (2)$$

where the dependent variable is the Engagement a post generated at the time it was published; Arousal and Macro are the focal variables; Arousal × Macro is the interaction term; **X'** includes all the controls; τ indicates the endogeneity correction, and ε is the error term (see Table WB6, column 1 for the results of the second stage).

Results from first stage-regression

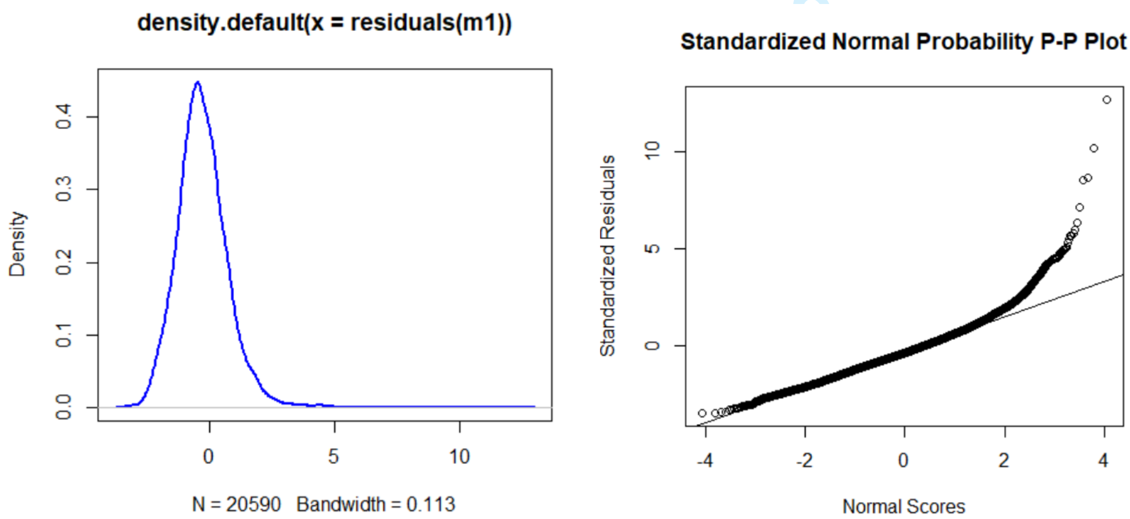
	DV: AROUSAL
Lagged Arousal	.248** (.007)
if Holiday	.075** (.026)
Announcement	.157** (.007)
N	19,547
R-square	.091

* $p < .05$, ** $p < .01$.
Notes: Standard errors are in parentheses.

Informative Goal

To account for influencers' strategic use of "informative goal" in sponsored posts, we introduce one copula term into the regression equation; it can account for the endogenous regressor (Park and Gupta 2012). We start by carefully checking the theoretical and empirical evidence demanded by a Gaussian copula approach (Becker, Proksch, and Ringle 2022): a large enough sample size, the endogenous regressor's sufficient nonnormality, and the error term distribution's normality. First, our sample size of 20,923 observations is sufficient to inspect the nonnormality of the continuous endogenous regressor. Second, informative goal fulfills the nonnormality criterion for our sample size because: 1) despite that skewness is lower than .77 (in our case .21) our sample is higher than 2,000 observations, 2) the Cramer-von Mises is higher than 2.682 (1,542.5, $p < .001$). These criteria suggest a copula term with power of 80% and higher. Third, a Kernel density plot and standardized normal probability (P-P) plot both suggest that the regression residual in the estimation without copula terms is normally distributed. The values at the end of the distribution are more extreme due to the characteristic of the negative binomial distribution.

Table WB1. Kernel density plot and standardized normal probability



Thus, we add a Gaussian Copula to our regression model to account for the correlation between the informative goal in influencer posts and the error term such that:

$$GC_{ij} = \Phi^{-1}[H(informative\ goal_{ji})],$$

where $\Phi^{-1}$is the inverse of the normal distribution function, and $H(information\ goal_{ji})$ represents the empirical distribution function of informative goal. Note that in line with Papies, Ebbes and Van Heerde (2018), we use only a Gaussian Copula for the potential endogenous regressor and not for the interaction term, and we used bootstrapped standard errors for the estimation. The copula term is non-significant ($GC\_informative\ goal_{ij}$= .004, $p > .1$), so these finding do not support including the copula term in our model (Wlömert and Papies 2019).

Selection Bias

While our results show that high arousal language boosts engagement for micro influencers while it decreases engagement for macro influencer, one could wonder whether the relationship is driven by the particular sample used (i.e., selection).

To address this possibility, we rely on propensity score matching (PSM, Rosenbaum and Rubin 1983). PSM assumes that there are control variables capable of identifying the selection into treatment and control groups, and uses these controls to estimate a score such that the distribution of all the observed variables and behaviors among the treated units is similar to that among the control units (Imbens and Rubin 2015). In other words, the PSM "adjusts" for the differences in the treatment and control group which may bias the inferences about the treatment effect. When the propensity scores for two observations are close enough to each other, the treatment is considered random. Thus, the biases in the comparisons between treated and control units are eliminated (i.e., "quasi-experiment"; Goldfarb, Tucker, and Wang 2022).

In our case, the propensity score is the predicted probability that a unit receives the treatment (i.e., the poster is a macro influencer) conditional on the value of covariates. To create a matching sample, all influencer characteristics (e.g., post count, if verified), text features of a post (e.g., topics, number of mentions, hashtags, emojis, concreteness), aspects of the image (e.g., color dominance, face presence, saturation), and posting time (e.g., time fixed effects) were included in the matching model, letting only the arousal vary.

To estimate p_{kt}, the probability of being a macro influencer as a function of the covariates, a logistic regression model was used as follows:

$$px_{kt} = P(T_{kt} = 1 \mid X_{kt}) = \exp(T_{kt}\beta) \;/\; [1 + \exp(X_{kt}\beta)],$$

where T_k is the treatment status which indicates whether the influencer who posted content k at time t was a macro influencer, and X_{kt} includes all the covariates.

To calculate the propensity score for each post in our sample, following prior work (Li and Xie 2020), we adopted a 1:1 nearest-neighbor matching algorithm without replacement and a caliper of .01 to match a post shared by a macro influencer with a post shared by a micro influencer, but with the closest propensity score. The resulting matched sample contains 10,554 posts, half from macro influencers and half from micro influencers.

Table WB6: Robustness Checks

Column 1: Addressing endogeneity with Control Function approach, **Column 2**: Addressing selection bias with PSM, **Column 3**: OLS with log-transformed DV, **Column 4:** Simple words arousal**, Column 5:** Paralanguage arousal, **Column 6:** Simple words and PARA arousal, **Column 7:** High-arousal words from Villarroel Ordenes et al. (2017), **Column 8:** Macro measured continuously (follower count).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IV								
Arousal	1.077** (.023)	1.031** (.012)	.062** (.010)	1.002 (.007)	1.026 (.024)	1.006 (.006)	1.018** (.005)	1.007 (.006)
Macro	4.448** (.066)	4.332** (.070)	1.426** (.015)	4.532** (.065)	5.289** (.344)	4.534** (.065)	4.578** (.074)	2.310** (.017)
Arousal × Macro	.912** (.011)	.910** (.015)	–.089** (.015)	.926** (.012)	.917* (.032)	.938** (.009)	.988† (.007)	.976** (.006)
Controls								
Influencer								
if Verified	1.246** (.020)	1.183** (.029)	.254** (.017)	1.251** (.020)	1.263** (.041)	1.251** (.020)	1.255** (.020)	1.083** (.016)
# of Posts	.903** (.006)	.890** (.011)	–.045** (.007)	.905** (.006)	.896** (.014)	.906** (.006)	.904** (.006)	.876** (.006)
Content Variation	.982** (.006)	.963** (.009)	–.018* (.008)	.982** (.006)	.978 (.014)	.982** (.006)	.983** (.006)	1.006 (.006)
Category FE	Included	Included	Included	Included	Included	Included	Included	Included
Text								
Topics	Included	Included	Included	Included	Included	Included	Included	Included
# of Question Marks	.997 (.006)	.992 (.008)	–.019** (.007)	.993 (.006)	.984 (.013)	.993 (.006)	.994 (.006)	.987* (.006)
# of Hashtags	.993 (.007)	.991 (.010)	–.010 (.008)	.991 (.007)	.984 (.015)	.991 (.007)	.991 (.007)	1.002 (.007)
# of Mentions	1.001 (.006)	.995 (.008)	.009 (.006)	1.001 (.006)	1.008 (.013)	1.002 (.006)	1.001 (.006)	.993 (.006)
Word Count	.999 (.008)	.975 (.010)	–.002 (.008)	.997 (.007)	.982 (.015)	.997 (.007)	.990 (.008)	.991 (.007)
# of Emojis	1.008 (.007)	1.031** (.009)	.011 (.007)	1.010 (.006)	.997 (.013)	1.010 (.006)	1.011† (.006)	1.020** (.006)
Complexity	1.012 (.008)	.987 (.011)	.015† (.008)	1.013† (.008)	1.045* (.018)	1.013† (.008)	1.014† (.008)	1.004 (.007)
Valence	1.012† (.007)	.997 (.009)	.006 (.007)	1.011 (.007)	.976 (.015)	1.011 (.007)	1.009 (.007)	1.016* (.006)
Concreteness	.979** (.007)	.983 (.007)	–.022** (.008)	.979** (.007)	.963* (.015)	.979** (.007)	.980** (.007)	.974** (.007)
Familiarity	.975** (.007)	.994** (.010)	–.021** (.008)	.980** (.007)	.990 (.016)	.980** (.007)	.981* (.007)	.977** (.007)
Image								
if Image (vs. Video)	1.331** (.031)	1.411** (.044)	.442** (.028)	1.344** (.030)	1.402** (.061)	1.344** (.030)	1.363** (.030)	1.554** (.032)
if Face Present	.979 (.013)	.980 (.018)	–.037* (.015)	.985 (.013)	.991 (.028)	.985 (.013)	.985 (.013)	.961** (.012)
Image Emotionality	1.052** (.007)	1.052** (.010)	.075** (.007)	1.052** (.007)	1.063** (.015)	1.052** (.007)	1.050** (.007)	1.022** (.006)
Color Dominance	1.004 (.007)	1.001 (.009)	.005 (.007)	1.003 (.006)	1.007 (.013)	1.003 (.006)	1.003 (.006)	1.010 (.006)
Color Saturation	1.008 (.006)	1.008 (.008)	.023** (.007)	1.012† (.006)	1.008 (.014)	1.012† (.006)	1.011† (.006)	1.015* (.006)
Additional								
Time Difference	.991 (.006)	1.007 (.009)	–.013† (.006)	.990† (.006)	.968* (.013)	.989† (.006)	.991 (.006)	.981** (.006)
Time FE	Included	Included	Included	Included	Included	Included	Included	Included
Residuals	.959** (.020)							
N	19,429	10,554	20,590	20,590	4,920	20,590	20,590	20,590
Log likelihood	–169,769	–92,867	-	–180,116	–43,030	–180,115	–180,135	–178,611

†$p < .10$, * $p < .05$, ** $p < .01$. Notes: Standard errors are in parentheses. We do not report coefficients for fixed effects and topics, for parsimony.

Confirmation of Prior Findings

Our findings also corroborate several insights from prior research. Verified influencers are more likely to receive likes or comments (Valsesia, Proserpio, and Nunes 2020); expertise boosts engagement (Hughes, Swaminathan, and Brooks 2019); and image emotionality can affect engagement too (Li and Xie 2020). Then we note some pertinent differences. For example, Stieglitz and Dang-Xuan (2013) indicate that hashtags increase post visibility and engagement, but we cannot replicate this effect. Influencers use popular hashtags (e.g., “love,” “picoftheday”), many of which are inconsistent with the content posted, resulting in irrelevant content for consumers’ searches. Prior work also implies that videos go more viral than images (Borah et al. 2020), but our findings suggest the opposite. This result might reflect the relatively fewer videos in our data (around 11%), but it also might signal consumers’ lack of patience to keep watching long video advertisements (Tellis et al. 2019). People tend to scroll through their social media feeds quickly, making it unlikely that they will watch the whole video, elaborate on it, and like or comment on it.

Table WB7. The Effects on Trust

	DV: TRUST
IV	
Arousal	.009** (.001)
Macro (vs. Micro)	.017** (.003)
Arousal × Macro	–.013** (.002)
Controls	
Influencer	
if Verified	–.001 (.003)
# of Posts	.003* (.001)
Content Variation	.002 (.001)
Category FE	Included
Text	
Topics	Included
# of Question Marks	–.001 (.001)
# of Hashtags	.002 (.001)
# of Mentions	–.004** (.001)
Word Count	.002 (.001)
# of Emojis	.001 (.001)
Complexity	.001 (.001)
Valence	–.001 (.001)
Concreteness	–.003 (.001)
Familiarity	–.002 (.001)
Image	
if Image (vs. Video)	.009* (.004)
if Face Present	–.001 (.002)
Image Emotionality	.005** (.001)
Color Dominance	.001 (.001)
Color Saturation	.002 (.001)
Additional	
Time Difference	–.005** (.001)
Time FE	Included
N	20,590
Log likelihood	6,088

* $p < .05$, ** $p < .01$.

Notes: Standard errors are in parentheses. We do not report coefficients for fixed effects and topics, for parsimony.

New Dictionary for Language Trustworthiness

Two research assistants (blinded to hypotheses) received a definition of trustworthiness ("source's sincerity and motivation to provide accurate information"; Pornpitakpan 2004). The assistants also received a random, industry-stratified sample of 2,000 posts (10% of all data) and had to annotate each post, according to how trustworthy the language was (1 = "not at all," 7 = "very much"). Their ratings were highly correlated (r = .83), so we averaged their ratings to get a unique measure of language trustworthiness at the post level. Posts with a score above the mean plus one standard deviation were classified as trustworthy, and those with a score below mean minus one standard deviation were classified as nontrustworthy. Then, we used Wordify (Hovy, Melumad, and Inman 2021) to find which n-grams (i.e., words and concatenations of words) in our sample data are most indicative of each of variable class (trustworthy and nontrustworthy). For each n-gram, Wordify returns a correlation score value, which is positive if the word is more likely to belong to the trustworthy class and negative otherwise. Wordify returned a list of 28 words, including 22 words signaling trustworthiness (e.g., help, r = .324; learn r = .272) and 6 words signaling nontrustworthy (e.g., gifted, r = -.33; advertising, r = -0.26). We operationalized the construct as: (trustworthiness words' score – nontrustworthiness words' score).

Trustworthiness words (correlation): *year (.378), know (.374), go (.37), experience (.338), love (.332), add (.326), help (.324), week (.322), way (.316), body (.31), like (.306), want (.306), day (.304), small (.28), learn (.272), partner (.272), time (.264), thing (.26), recipe (.258), special (.258), come (.254), space (.254).*

Nontrustworthiness words (correlation): *ad (-.392), sponsor (-.38), gifted (-.33), sponsored (-.274), gift (-.266), advertising (-.26).*

Web Appendix C. Follow-Up to Study 1: Language Arousal on TikTok

The agency partner selected all influencers they work with who had published at least one sponsored post in the last 2 years. Data include 654 TikTok posts from 172 influencers between January 23, 2020, and Oct 30, 2021. The posts cover five industries (Table W6).

Table WC1. Sample Description

Category	Number of influencers	Number of Posts	Avg Post Likes (SD)	Avg Post Comments (SD)
Beauty	68	183	86,626 (215,086)	782 (5,894)
Fashion & Lifestyle	139	333	144,505 (488,272)	1,970 (12,553)
Food & Drinks	48	101	196,201 (759,171)	6,162 (47,373)
Gaming	12	31	203,719 (462,749)	3,043 (7,350)
Travel & Tourism	5	6	42,316 (70,919)	168 (191)

Table WC2. Descriptive Statistics and Correlations

Variables	M	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)
(1) Engagement	137,603	483,608	1.000																						
(2) Pitch	44.43	8.45	.074	1.000																					
(3) Follower Count	5,725,121	1,470,007	.878*	.091	1.000																				
(4) Loudness	.12	.08	-.107*	-.126*	-.125*	1.000																			
(5) Intonation	.18	.04	.093	.343*	.073	.381*	1.000																		
(6) Brightness	2,077	643.36	.049	.503*	.080	.079	.410*	1.000																	
(7) Articulation Rate	4.13	1.26	-.022	.522*	.047	.109*	.195*	.431*	1.000																
(8) Clarity	.96	.15	.029	.808*	.039	.198*	.597*	.529*	.519*	1.000															
(9) Duration	51.28	34.39	-.093	-.021	-.085	.368*	.224*	.072	-.118*	.217*	1.000														
(10) Arousal	.43	.07	-.106*	.143*	-.033	.055	.117*	.086	.171*	.157*	-.028	1.000													
(11) # of Posts	10.13	9.27	-.072	.033	-.099	.058	.017	.057	.044	.058	.173*	-.040	1.000												
(12) # of Questions	1.15	1.89	.009	-.058	-.044	-.049	.068	-.053	-.094	-.037	-.058	.038	-.076	1.000											
(13) # of Mentions	.91	.29	-.001	.104	.016	.129*	.004	.018	.120*	.072	.182*	.009	.067	-.166*	1.000										
(14) Word Count	110.8	62.01	-.070	-.010	-.050	-.046	-.203*	-.063	-.063	-.111*	.341*	-.051	.200*	-.169*	.372*	1.000									
(15) Complexity	13.01	6.38	-.115*	-.009	-.044	.038	-.144*	-.044	.207*	.015	.052	.105*	.038	-.226*	.210*	.030	1.000								
(16) Valence	.64	.09	-.147*	.132*	-.072	.075	.040	.052	.094	.117*	.112*	.533*	.022	.076	.125*	.055	.097	1.000							
(17) Concreteness	341.63	27.26	-.107*	.236*	-.127*	.064	.225*	.129*	.167*	.328*	.021	.152*	.024	.106*	.006	-.156*	.038	.211*	1.000						
(18) Familiarity	591.53	35.04	.053	.238*	.043	.078	.182*	.171*	.136*	.263*	.091	.350*	-.020	.056	.174*	.184*	-.158*	.411*	.481*	1.000					
(19) if Face Present	.54	.50	-.052	-.007	-.050	-.067	.029	.017	-.033	-.002	-.107*	.098	.033	.159*	-.123*	-.094	-.164*	.067	.032	-.016	1.000				
(20) Image Emotionality	.34	.16	.140*	-.021	.111*	.077	.106*	.045	.047	-.007	-.053	-.002	-.227*	.059	.002	-.167*	.016	.020	.013	.053	.017	1.000			
(21) Color Dominance	.69	.09	.048	.105*	.048	-.060	-.021	.057	-.032	.024	.061	-.107*	.212*	.047	-.040	.123*	-.060	-.060	-.038	.024	-.193*	-.113*	1.000		
(22) Color Saturation	.57	.06	.018	-.087	.004	.014	-.026	-.024	.076	-.077	-.052	-.052	-.028	-.050	.123*	.024	-.012	-.088	.026	.032	-.021	.020	-.007	1.000	
(23) Time Difference	.37	.48	.107*	.039	.090	-.009	-.002	-.063	-.028	.026	-.010	.024	.094	-.029	.068	-.062	.033	.034	.006	.042	.055	-.167*	.006	.019	1.000

* p <. 05. Notes: Fixed effects and topics are not included.

Method

Engagement. As in Study 1, engagement was operationalized as the sum of likes and comments. On average, posts received 137,603 likes (SD = 483,608, ranging from 31 to 5,700,000) and 2,319 comments (SD = 20,953, ranging from 0 to 461,600; see Table WB8 for descriptive statistics and correlations).

Pitch. Arousal was operationalized as the level of pitch of influencers' voice. We measured the pitch using the YIN frequency estimator algorithm. This algorithm estimates the fundamental frequency given the frame of an audio signal, and is based on autocorrelation methods (please consult de Cheveigné and Kawahara 2002 for details).

Influencer type. In Study 1, we classified micro and macro influencers based on a follower count threshold (i.e., 100,000 followers). Our TikTok data, however, include influencers with more than 100,000 followers. Given the definition of a cut-off point might vary based on idiosyncratic characteristics of the social media platform (e.g., total audience size, prevalence of influencers), we measured this variable continuously (i.e., follower count).

Controls. We included similar controls to Study 1 (see Table WB9 for full list). All influencers in our data set were verified, so this variable was not included, and speech does not include hashtags and emojis, so these were not included either. Approximately 60 percent of videos featured a speech. To account for the difference between a video with a speech and a video without, we dummy coded the speech presence variable (0 = no speech, 1 = speech). Given that only 285 videos with speech were posted by influencers who have shared at least three sponsored posts, we did not include the "content variation" variable. We extracted video features using an open-source video mining tool from Schwenzow et al. (2021). Finally, *librosa* Python package was used to measure various acoustic features (McFee et al. 2015), such as voice loudness (computed as the mean of the frame's root-mean-square), intonation (computed as standard deviation of the pitch), voice brightness (computed as the mean of the audio signal's spectral centroid), articulation rate (i.e., computed as the number of syllables per speech duration using the spectral flux), and speech duration.

We used the approach from Study 1 to test the relationship between pitch, follower count, and engagement.

Results

Consistent with Study 1, results show a significant pitch × follower count interaction (IRR = .776; SE = .047; t = –4.19; p < .001; Table WB9, column 1) on engagement. Even after accounting for all the controls, we find a significant effect of pitch × follower count (IRR = .785; SE = .040; t = –4.78; p < .001; Table WB9, column 2).[1] The results suggest that as the follower count grows, high arousal (i.e., higher pitch) has an increasingly negative effect on engagement (see Figure WC1).

Discussion

The results of this follow-up to Study 1 underscore the relationship between arousal, influencer type, and engagement in the field. TikTok influencer posts that used higher pitch (i.e., higher arousal) received less engagement as the influencers' follower count grows. This effect persisted controlling for a range of alternative explanations. Finding the same effect using a different social media platform, and spoken (rather than written) language, speaks to the robustness and generalizability of the effect.

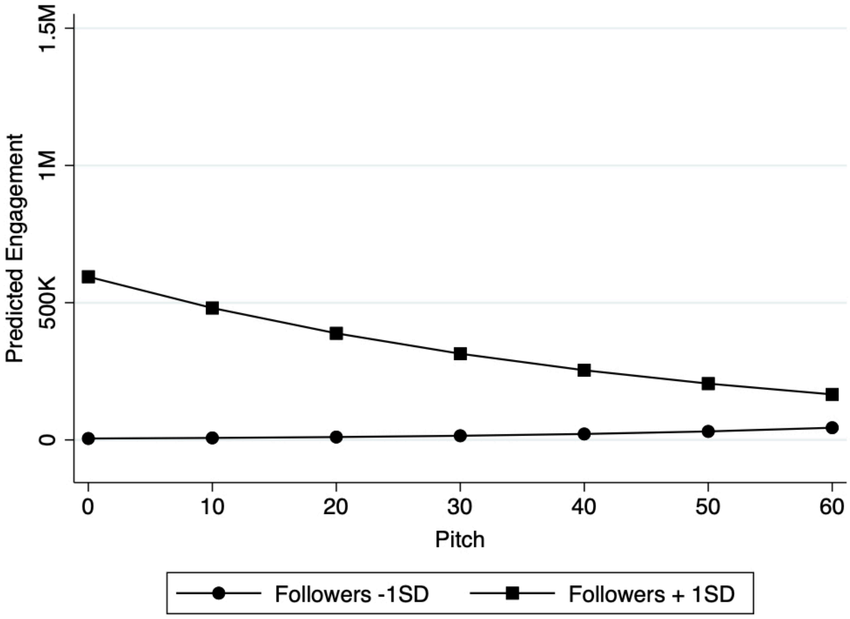
[1] Note that results remain the same even including influencer fixed effects with cluster-robust standard errors.

Table WC3. Results

	DV: ENGAGEMENT	
	(1)	**(2)**
IV		
Pitch	1.065** (.021)	1.064** (.190)
Follower Count	3.728** (.666)	3.022** (.395)
Pitch × Follower Count	.776** (.047)	.785** (.040)
Controls		
Influencer		
# of Posts		.991 (.008)
Category FE		Included
Audio		
Loudness		1.232 (.139)
Intonation		.997 (.115)
Brightness		1.255* (.134)
Articulation Rate		.766* (.084)
Clarity		1.269 (.266)
Duration		.745* (.090)
Text		
Topics		Included
Arousal		1.033 (.387)
if Speech Present		1.407 (.817)
# of Questions		1.300** (.095)
# of Mentions		.934 (.311)
Word Count		1.592** (.241)
Complexity		.883 (.122)
Valence		1.935 (.884)
Concreteness		.806 (.604)
Familiarity		.995 (.208)
Video		
if Face Present		1.093 (.129)
Image Emotionality		.917 (.056)
Color Dominance		.992 (.054)
Color Saturation		.949 (.048)
Additional		
Time Difference		.823 (.106)
Time FE		Included
N	654	654
Log likelihood	–7,752	–7,664

* $p < .05$, ** $p < .01$. Notes: SE are in parentheses. We do not report coefficients for fixed effects and topics, for parsimony.

Figure WC1. The Effects of Follower Count on Pitch

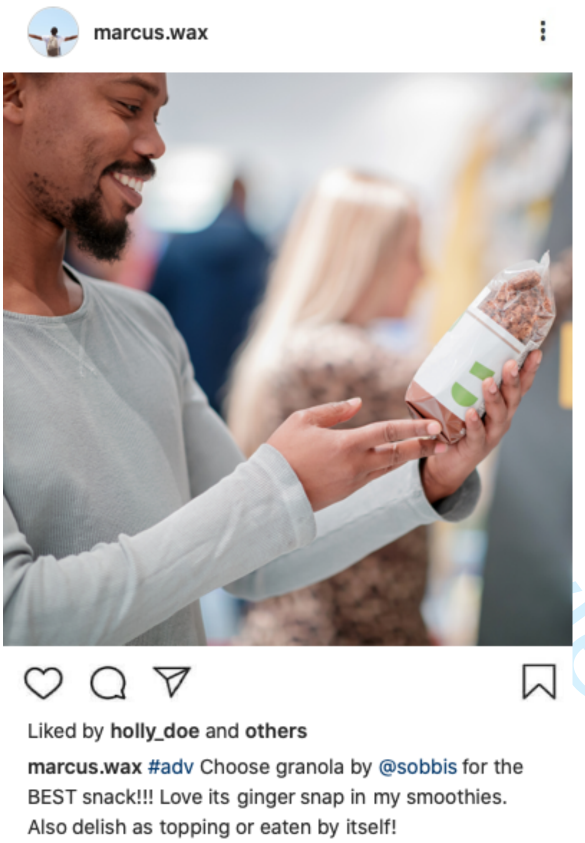


Web Appendix D. Experimental Studies, Details

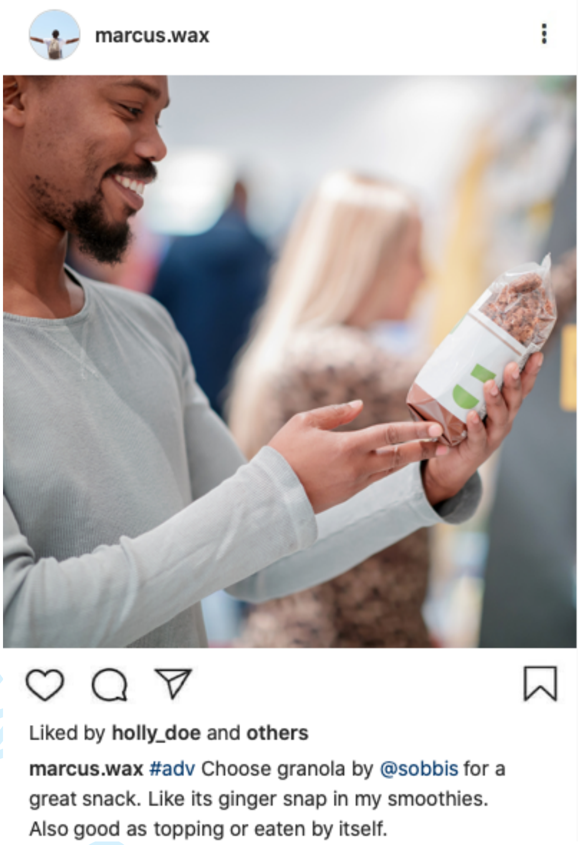
Study 2: Manipulating Language Arousal

Stimuli

High Arousal



Low Arousal



Exclusion and Demographic Information. Three hundred US Instagram users were recruited from Prolific. Following the preregistration (https://aspredicted.org/blind.php?x=L82_4SM), participants (n = 21) were excluded if they failed an attention check asking them “how many followers did the influencer have?: less than 100,000; more than 100,000”. The final sample consisted of 279 participants (60.6% female; mean age = 31.8 years).

Manipulation checks. Participants in the high arousal condition perceived the language to be higher in arousal than in the low arousal condition (M = 5.28 vs. 3.28, $F(1, 277) = 185.52$, $p < .001$, $\eta^2 = .401$). Participants also rated the macro influencer as more able to reach a higher number of people compared to the micro influencer (M = 5.31 vs. 4.26, $F(1, 277) = 43.08$, $p < .001$, $\eta^2 = .134$).

Exploratory Study

This exploratory study has two main goals. First, it tests whether, consistent with our theorizing, people trust micro versus macro influencers equally at the onset (i.e., before the language happens). Second, if macro influencers' use of high arousal activates persuasion knowledge, as we suggest, then this happens because people assume that macro (vs. micro) influencers are more likely to get paid to sponsor products. The exploratory study also tests this.

Method

Participants (N = 120, 60.8% female; mean age = 31.8 years, Prolific) were randomly assigned to a condition in a 2 (influencer type: micro vs. macro) between-subjects design.

Everyone was asked to image coming across an Instagram post sponsored by an influencer. The only difference between condition was the influencer type noted (i.e., micro: 20,000 followers vs. macro: 660,000 followers, as in Study 2).

Next, we collected the measures of trust. Participants rated their perceptions of the influencer's trustworthiness using five items (7-point scale, "untrustworthy–trustworthy," "insincere–sincere," "undependable–dependable," "dishonest–honest," "unreliable–reliable"; $\alpha = .93$; Ohanian 1990). In addition, they indicated the extent to which they thought the influencer usually gets paid to endorse products (1 = not at all, 7 = very).

Results and Discussion

Consistent with our theorizing, people equally trusted the micro versus macro influencer (M_{micro} = 4.58; M_{macro} = 4.37; $F(1, 118) = 1.54$, $p = .217$, $\eta^2 = .013$).

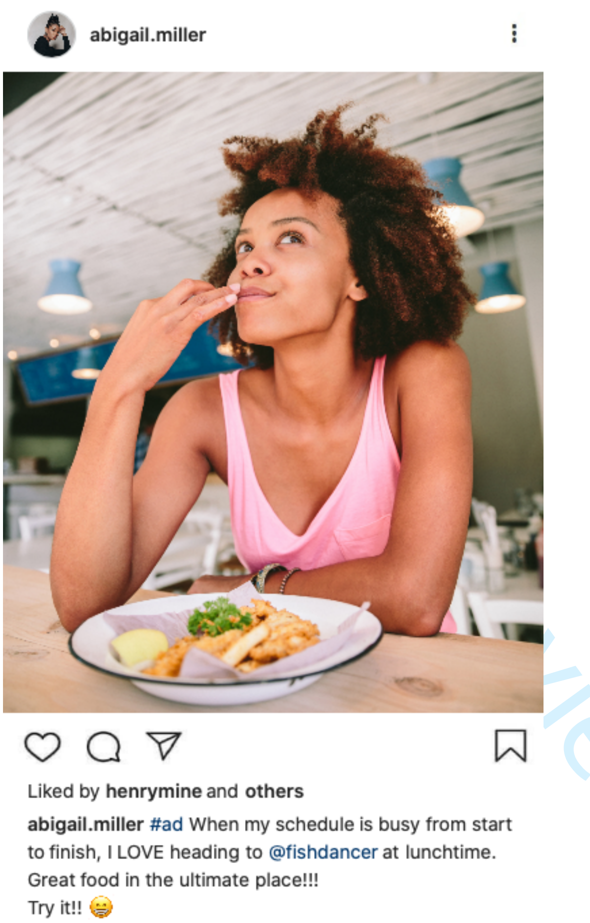
In addition, compared to micro influencers, people assume macro influencers are more likely to get paid to endorse products (M_{micro} = 5.25; M_{macro} = 6.05; $F(1, 118) = 14.22$, $p < .001$, $\eta^2 = .108$).

This exploratory study rules out the possibility that micro influencers are more trusted than macro at the onset (before the language happens), which would make the effects of language arousal on trustworthiness simply polarized (i.e., high arousal language makes trusted micro influencers more trustworthy while less-trusted macro influencers less trustworthy). It also finds that people are more likely to assume that, compared to micro, macro influencers are more likely to get paid to endorse products. Taken together, these findings suggest that people may have some latent differences in trusting micro versus macro influencers, that are then activated by the language arousal used in the post, which causes consumers to trust micro more than macro influencers.

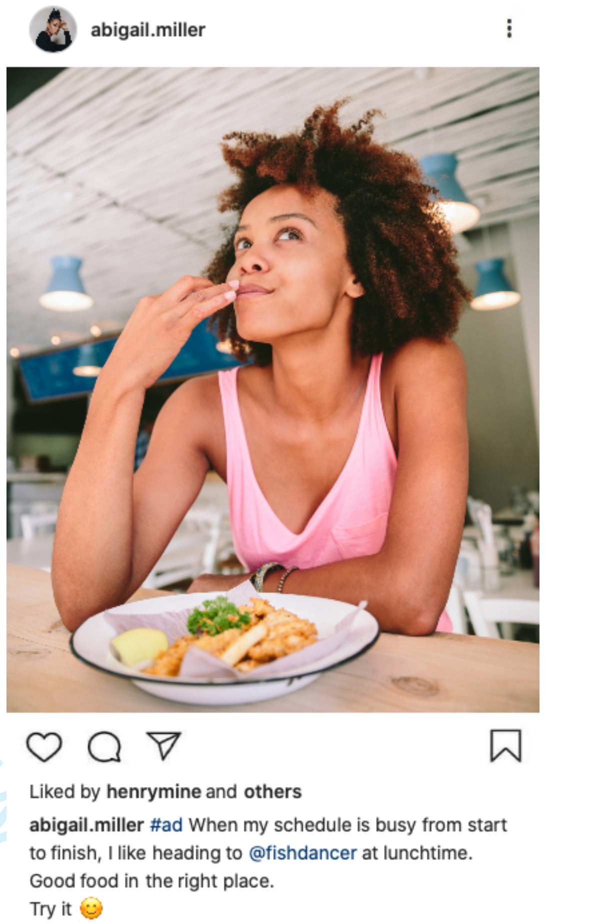
Study 3a: Testing the Process

Stimuli

High Arousal



Low Arousal



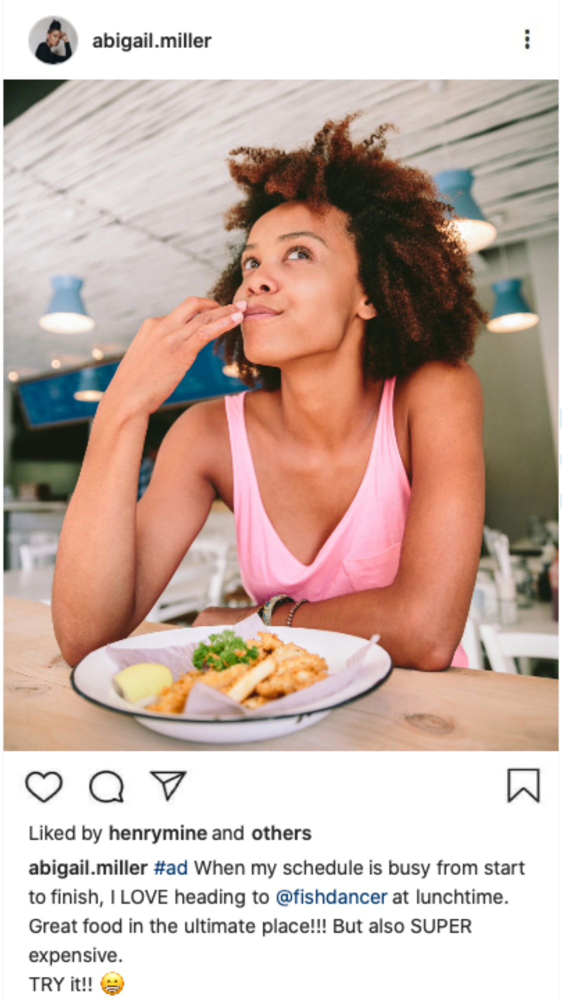
Exclusion and Demographic Information. Three hundred U.S. Instagram users were recruited from Prolific. Following the preregistration (https://aspredicted.org/VZX_HN3), participants (n = 29) were excluded if they failed an attention check asking them "how many followers did the influencer have?: less than 100,000; more than 100,000". The final sample consisted of 271 participants (56.8% female; mean age = 31.5 years).

Manipulation Checks. Participants in the high arousal condition perceived the language to be higher in arousal than in the low arousal condition (M = 5.04 vs. 4.15, F(1, 269) = 40.12, $p <$ .001, η^2 = .130). Participants also rated the macro influencer as more able to reach a higher number of people compared to the micro influencer (M = 5.47 vs. 4.36, $F(1, 269) = 51.26$, $p <$ .001, $\eta^2 = .160$).

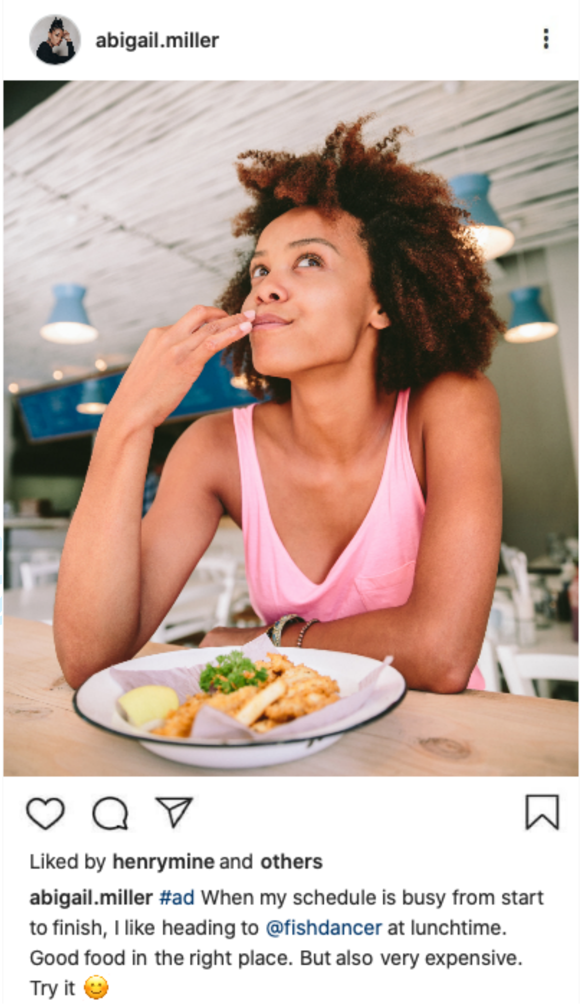
Study 3b: Process by Moderation (Language Valence)

Stimuli

High Arousal



Low Arousal



Exclusion and Demographic Information. Three hundred twenty U.S. Instagram were recruited through Prolific. Participants (n = 30) who failed the attention check (asking, "how many followers did the influencer have?: less than 100,000; more than 100,000") were excluded. The final sample consists of 290 people (70.7% female; mean age = 33.2 years).

Study 4: Process by Moderation

Exclusion and Demographic Information. Three hundred US Instagram users were recruited from Prolific. Following the preregistration (https://aspredicted.org/7ZJ_GTV), participants (n = 21) were excluded if they failed an attention check asking them "how many followers did the influencer have?: less than 100,000; more than 100,000". The final sample consisted of 279 participants (50.2% female; mean age = 34.6 years).

Manipulation checks. Participants in the high arousal condition perceived the language to be higher in arousal than in the low arousal condition (M = 5.05 vs. 4.15, $F(1, 277) = 36.78$, $p < .001$, $\eta^2 = .117$). Participants were also asked what they thought was the intent of the post (1 = informative and 7 = commercial). They were provided a definition that read, "An informative intent is expressed by words aimed at increasing knowledge about a product (e.g., discover, read) while a commercial intent is expressed by words aimed at encouraging consumer actions (e.g., buy, choose)". Participants rated the condition with the sentence "try it" as more commercial-oriented compared to the one including "learn more" (M = 5.07 vs. 4.04, $F(1, 277) = 19.36$, $p < .001$, $\eta^2 = .065$).

References (not in the main text)

Angrist, Joshua D., and Jorn-Steffen Pischke (2010), "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics," *Journal of economic perspectives*, 24 (2), 3-30.

Berger, Jonah, Grant Packard, Reihane Boghrati, Ming Hsu, Ashlee Humphreys, Andrea Luangrath, Sarah Moore, Gideon Nave, Cristopher Olivola, and Matthew Rocklage (2022), "Wisdom from words: marketing insights from text," *Marketing Letters*, 1-13.

George, Morris, Viswanathan Kumar, and Dhruv Grewal (2013), "Maximizing profits for a multi-category catalog retailer," *Journal of Retailing*, 89 (4), 374-396.

Goldfarb, Avi, Catherine Tucker, and Yanwen Wang (2022), "Conducting research in marketing with quasi-experiments," *Journal of Marketing*, 86 (3), 1-20.

Imbens, Guido W., and Jeffrey M. Wooldridge (2009), "Recent developments in the econometrics of program evaluation," *Journal of economic literature*, 47 (1), 5-86.

Imbens, Guido W., and Donald B. Rubin (2015), "Causal inference in statistics, social, and biomedical sciences," Cambridge University Press.

Krippendorff, Klaus (2013), *Content Analysis: An Introduction to Its Methodology*, Thousand Oaks, CA: Sage.

Kumar, Viswanathan, JeeWon Brianna Choi, and Mallik Greene (2017), "Synergistic effects of social media and traditional marketing on brand sales: Capturing the time-varying effects," *Journal of the Academy of marketing Science*, 45 (2), 268-288.

Nguyen, Thin, Dinh Phung, Brett Adams, and Svetha Venkatesh (2013), "Event extraction using behaviors of sentiment signals and burst structure in social media," *Knowledge and information systems*, 37 (2), 279-304.

Papies, Dominik, Peter Ebbes, and Harald J. van Heerde (2017), "Addressing endogeneity in marketing models." In *Advanced methods for modeling markets* (pp. 581-627). Springer, Cham.

Park, Sungho, and Sachin Gupta (2012), "Handling endogenous regressors by joint estimation using copulas," *Marketing Science*, 31 (4), 567-586.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn (2015). *The development and psychometric properties of LIWC2015*.

Rosenbaum, Paul R., and Donald B. Rubin (1983), "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70 (1), 41-55.

Schwenzow, Jasper, Jochen Hartmann, Amos Schikowsky, and Mark Heitmann (2021), "Understanding videos at scale: How to extract insights for business research," *Journal of Business Research*, 123, 367-379.

Stieglitz, Stefan and Lihn Dang-Xuan (2013), "Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior," *Journal of Management Information Systems*, 29 (4), 217-248.

Wang, Ze, Huifang, Yexin Jessica Li, and Fan Liu (2017), "Smile big or not? Effects of smile intensity on perceptions of warmth and competence," *Journal of Consumer Research*, 43 (5), 787-805.

Wlömert, Nils, and Dominik Papies (2019), "International heterogeneity in the associations of new business models and broadband internet with music revenue and piracy," *International Journal of Research in Marketing*, 36 (3), 400-419.