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

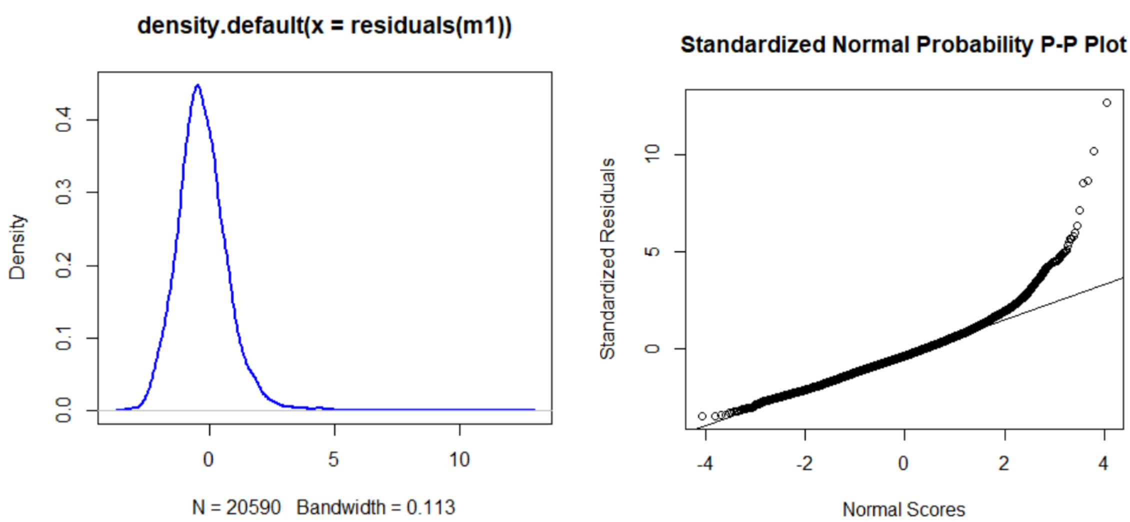
* $p < .05$, ** $p < .01$.

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

$^{\dagger}p < .10$, * $p < .05$, ** $p < .01$. 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.

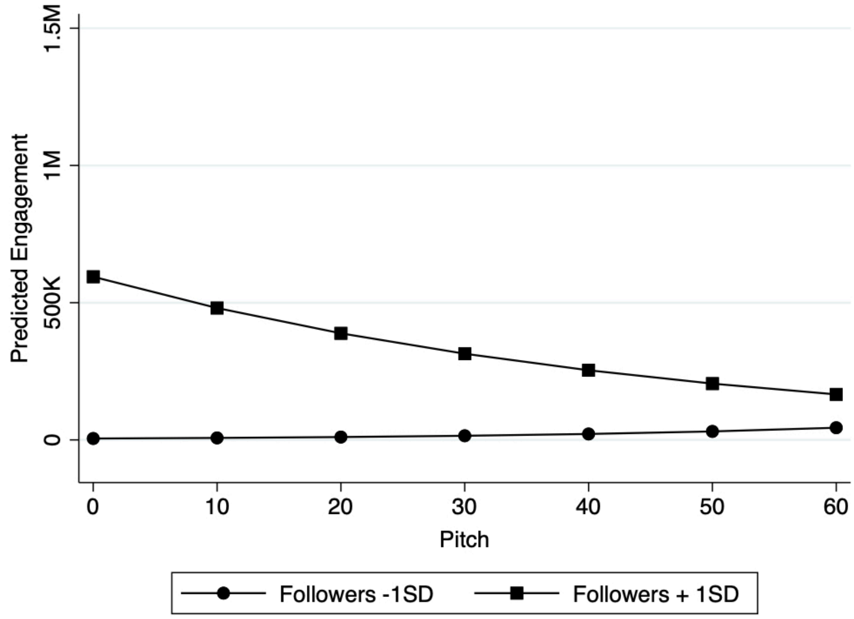
[1] Note that results remain the same even including influencer fixed effects with cluster-robust standard errors.

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

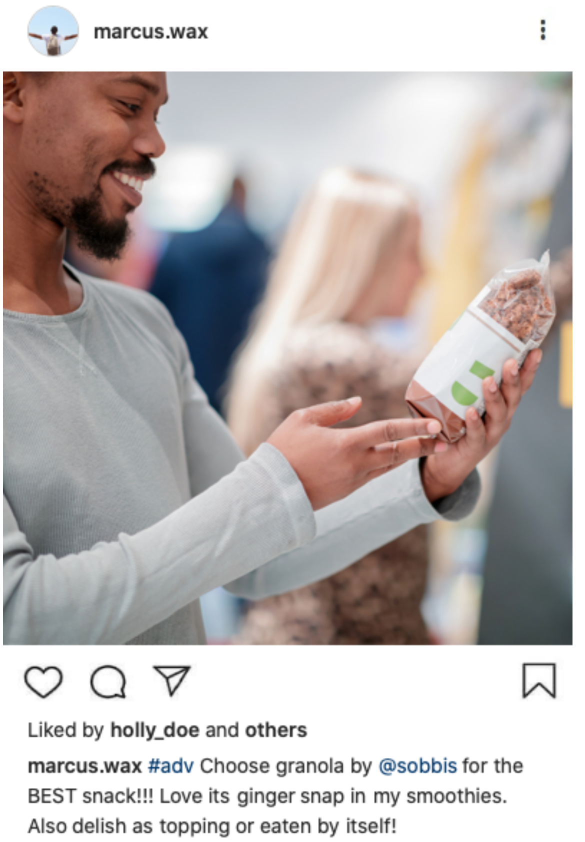


Web Appendix D. Experimental Studies, Details

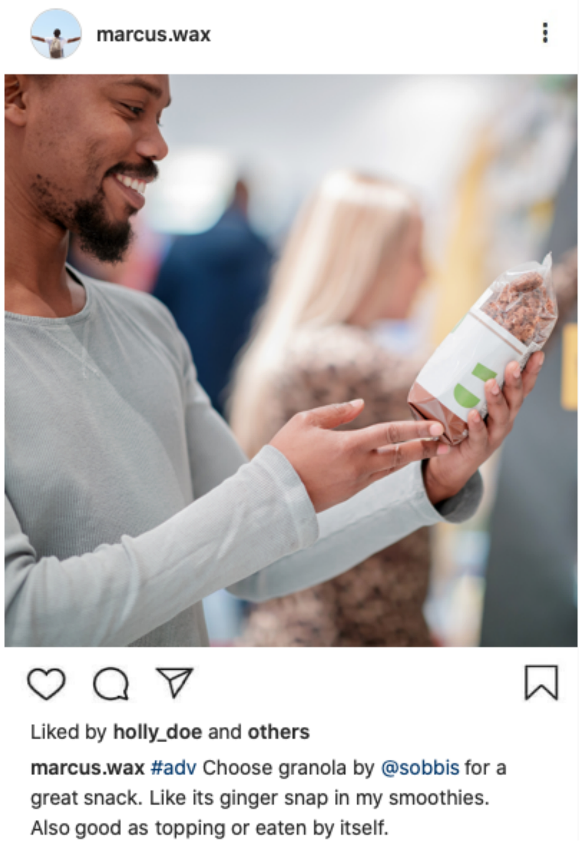
Study 2: Manipulating Language Arousal

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, η^2 = .013).

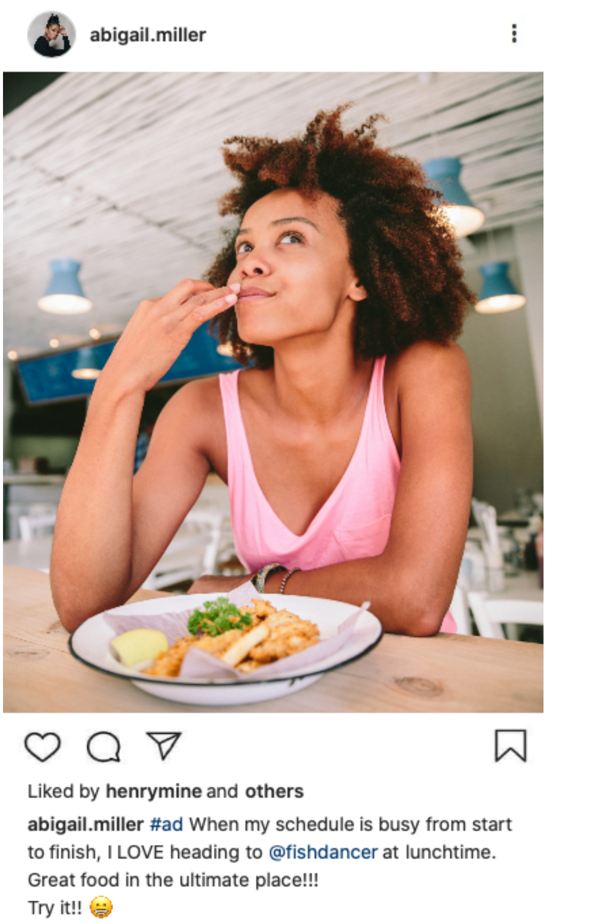
In addition, compared to micro influencers, people assume macro influencers are more likely to get paid to endorse products (M_{micro} = 5.25; M_{macro} = 6.05; F(1, 118) = 14.22, p < .001, η^2 = .108).

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.

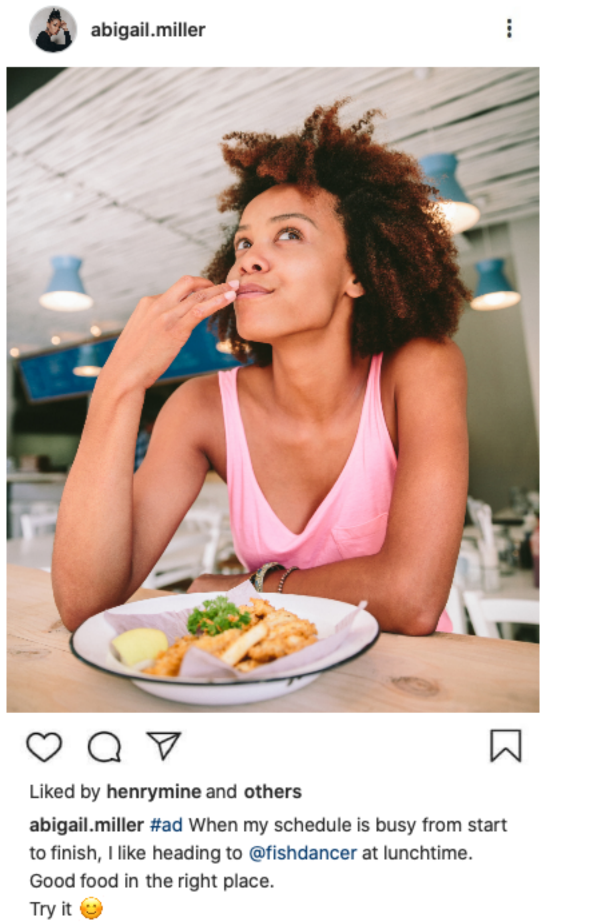
Study 3a: Testing the Process

Stimuli

High Arousal



Low Arousal



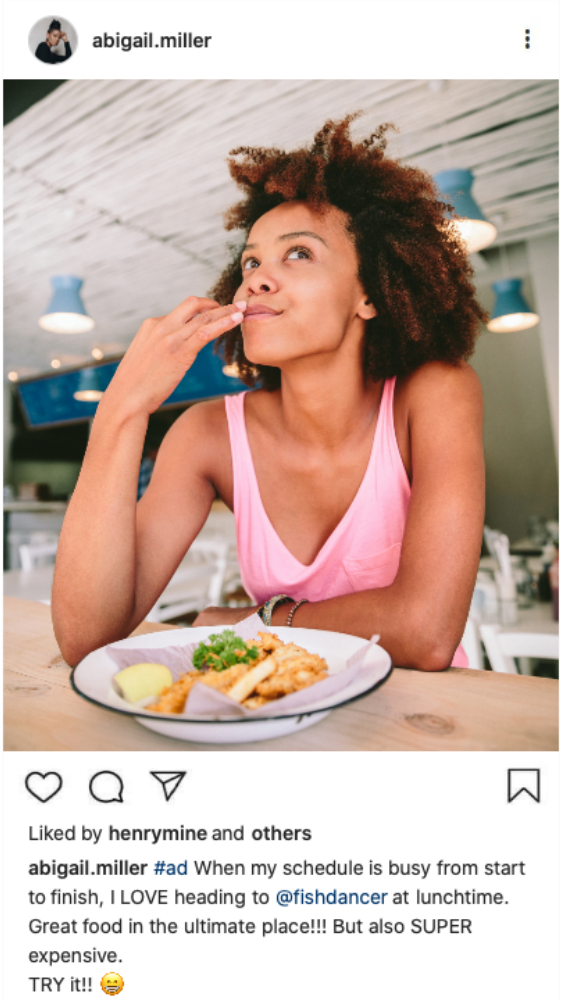
Exclusion and Demographic Information. Three hundred U.S. Instagram users were recruited from Prolific. Following the preregistration (https://aspredicted.org/VZX_HN3), participants (n = 29) were excluded if they failed an attention check asking them "how many followers did the influencer have?: less 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, $F(1, 269) = 51.26$, $p < .001$, $\eta^2 = .160$).

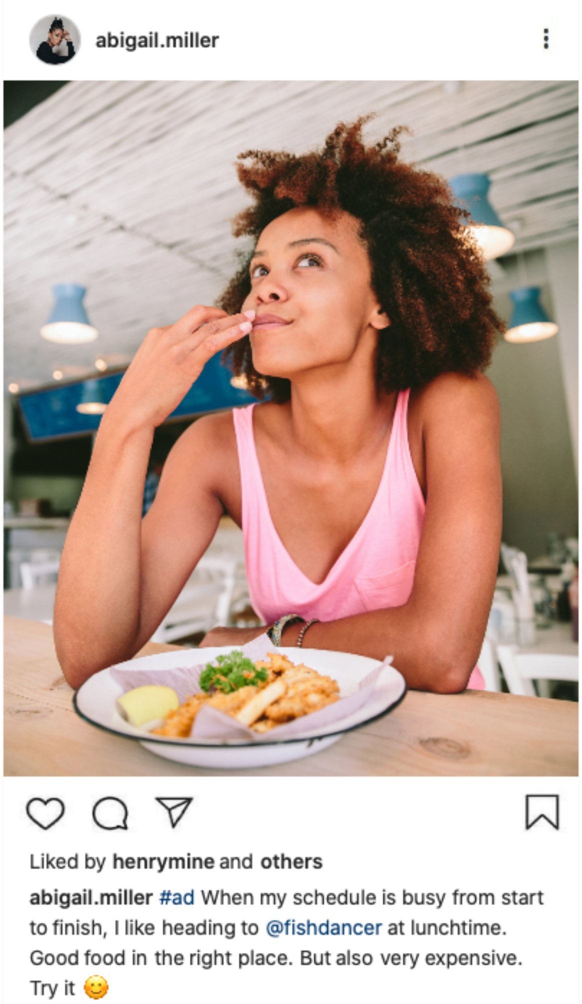
Study 3b: Process by Moderation (Language Valence)

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, η^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, η^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.