Can You Sell Millions of Lipsticks in 5 Minutes? A Multi-method Analysis of Winning Strategies of Livestreaming Influencers

By

Ruiqi Guan

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Marketing

University of Manitoba

Winnipeg

Copyright ©2024 by Ruiqi Guan

Contents

Abstract 4
Introduction 5
Research on Live Streaming 10
Early research on livestreaming. *10*
Multi-modality Communication. *11*
Study 1: Exploratory study 13
Sampling of Influencers. *13*
Sampling of influencers' livestream videos *14*
Analytic plan. *14*
Content analysis to detect rhetoric styles. *14*
Audio analysis. *17*
Facial expression analysis. *19*
Discussion of study 1 (exploratory study) and preparation for study 2 (experiment). *21*
Study 2: Experiment 22
Hypothesis development *22*
Preparation for Influencer Videos. *27*
Power analysis.. ..... *28*
Method. *29*
Results *29*
Discussion *34*
Study 3: Big Data Study of Douyin Livestream Videos 36
Clarified jargon *36*
Audio analysis *37*
Facial expression analysis *39*
Sampling. *42*
Methods. *44*
Results. *45*
Full model. 45
Discussion. *49*
General Discussion 51
Theoretical Contributions and Implications. *51*

Managerial Implications. *55*
Limitation and future research. *55*
Reference 56
Appendix A *69*
Tables and figures. 70

Abstract

Live-streaming is an innovative marketing strategy that has enabled top influencers to sell millions of products in a very short space of time. This study analyzes how rhetoric, voice, and facial expression lead to winning strategies for these influencers. In this work, we employ text, audio, and image analyses to discover top influencers' winning strategies based on 24 pretest videos of 6 influencers from Taobao with the follow-up experiment and 600 videos of 50 influencers from Douyin. Our study 1 (pretest on Taobao videos) identified three strategies – clarified jargon, voice arousal, and happy facial expression - employed by top influencers. In the follow-up experiment (study 2), we looked at the literature on persuasion and sales and formed the hypothesis that top influencers' strategies could benefit purchase intention through credibility, product evaluation, and audience emotions. We designed the experiment videos mocking top, middle, and bottom influencers, and our results supported all hypotheses. Lastly, our study 3 (big data on Douyin) demonstrates how sales performance is affected by clarified jargon, voice arousal, and happy facial expressions in the influencer's live-streaming content. We employ a multi-modality analysis and discover the significant result of employing clarified jargon, voice arousal, and happy facial expressions. Two interaction effects on sales were discovered: the interaction of clarified jargon and voice arousal, and the interaction of voice arousal and happy face. This research is one of the first to employ a multi-method approach to examine the content and non-verbal strategies in livestreaming marketing. It has important implications for influencers and influencers-to-be, for brands, and for platform companies hosting both influencers and brands.

Keywords: livestreaming, livestream, e-commerce, influencer, person brand, content analysis, facial expression analysis, audio mining

Introduction

Livestreaming e-commerce is an emerging form of online selling where influencers can rake in billions of dollars in a few hours. Consider the legendary Jiaqi Li in China as an example. He started as a salesperson at a cosmetics counter in a department store, but became a top influencer on Taobao, a major livestream selling platform in China. Western media picked up on his success and called him a "lipstick king," as he recorded $1.7 billion in sales in one day (November 11, 2021; the equivalent of Black Friday in terms of discounts). He has 30.3 million followers on Weibo (a Chinese social media platform), 60.8 million followers on Taobao (an e-commerce platform), and 11 million followers on The Red (a niche social media platform in China).

While there are numerous broadcasters who use livestream to sell products, there are only about 550 influencers who reached 100 million RMB in sales (Taobao annual report, 2021). Mr. Li stood out by consistently translating the livestream traffic into revenue. His personal brand with celebrity status helps transform his fans into paying consumers of the brands he is selling. Many brands leverage his success by seeking collaboration and adding "Jiaqi recommended" to their product label. On a popular social media app, the Red, a search with "Jiaqi recommended" yields 430,000 results of brands that Jiaqi has recommended and sold over time. Once they are promoted by Jiaqi, these brands will carry the label "Jiaqi recommended" to tap into the benefits of being associated with the celebrity streamer (Wang & Fan, 2021).

Inspired by the success of livestream influencers such as Jiaqi Li, our research examines the strategies of these influencers so that their successes can be replicated, to inspire future influencers-to-be, and to deepen our theoretical understanding of this novel and exciting retail phenomenon. This livestreaming retail format departs from the static forms of e-commerce such

as online shopping and social media marketing because it uniquely blends both the "show" and the "selling" action. It allows consumers to interact with sellers in real-time, resulting in an immersive interpersonal connection (Haimson &Tang, 2017; Wohn et al., 2018). At the same time, selling products is becoming a context via which this interpersonal and social interaction takes place. China's platform Taobao and Chinese TikTok are the two leaders in employing livestream selling as a new retail format. We utilized the tools of these two platforms to investigate influencers' strategies. We posit that when we contrast the videos with different sales performances, the winning content strategies are likely to be uncovered.

Past research has mostly focused on a single dimension of live streaming (e.g., interaction, motivation, attachments). Investigating pure interactive livestreaming, Lin et al. (2021) reported that the influencer's intentional use of happy facial expressions positively affected the amount of tips audiences gave during a livestream. There is now some limited research that investigates livestream commerce strategy (e.g., Bharadwaj et al., 2022). Livestreaming selling is different from interactive livestreaming; the former combines "show" and "selling" functions in order to achieve the selling goals, while the latter only focus on "showing" only.

Bharadwaj et al. (2022) found that there was a negative effect of emotional displays on sales; they suggested that salespeople should sell with a neutral facial expression (a straight face). They used a single-dimension perspective and examined the rather isolated effect of one key variable at a time. But livestream selling is an ongoing and dynamic format, and requires a comprehensive and multi-dimensional approach to understand the content and projection strategies of the influencers in a holistic way. To achieve the goal of comprehensively understanding live stream commerce, our research tries to see how the configuration of content and projection strategies (as opposed to a single static strategy) affects viewers' engagement and

responses.

Our study 1 was started by observing real influencers on the Taobao livestream. We first selected six influencers from the top, mid, and bottom tiers (two influencers in each tier) in the cosmetics sector on Taobao based on an existing influencer database (Huitun ranks top 500 livestreamers in a certain period – day, week, month, year). We downloaded a total of 24 livestream videos (four videos from each influencer during the same time frame) covering lipstick, face serum, make-up remover, and face masks and analyzed those data in terms of text, voice, and facial expressions. We hired two human coders to discover the unique content and rhetoric strategies (e.g., jargon) and employed machine learning algorithms and professional software to analyze audio and facial expressions data in these videos. Our study 1 of 6 influencers in the livestreaming category gave the answers to the two core questions: 1) what are the differences in the content, voice, and facial expressions between top livestreaming influencers and mid/bottom influencers and 2) what enables an online salesperson on the livestreaming platform to evolve into a person brand with fans and clout?

Based on our understanding of Chinese consumers' worship of knowledge and authority, we identified that top influencers used clarified jargon as the primary technique to attract an audience. In the past, literature has studied jargon without specifying the kind of different jargon. We named the jargon with straightforward explanations as clarified jargon, enriching the dimensions of linguistic studies on jargon. Further, since Chinese people love catching fun (courenao in Chinese), the fun vibe is also the key to drawing attention. Study 1 found that fun livestream rooms were more popular, generating more purchases. Our explanation of such observation was led by the unique habits of Chinese people in daily life, offering a new perspective to explain the effect of high arousal voice. Lastly, in study 1, we noticed the

importance of using happy faces in livestream selling. In Chinese culture, people use smiles to present their emotions in certain contexts, such as bitter smiles, compensation smiles, and fake smiles. Smile is not just a friendly signal but also a way to express emotions (even negative ones). Therefore, influencers usually present their smiles to the audience due to the ubiquitous and comprehensive usage of smiles. Such an explanation enriches non-verbal communication studies, providing new insights, especially in the Chinese context. The main contributions of Study 1 lie in the findings of livestream influencers' strategies and the new perspective to explain those findings.

Following the assessment of influencers' strategies of various tiers resulting from the study 1, we designed a subsequent experiment (study 2) and hired two actresses to be a mock influencer and produce mock livestream videos with different content strategies uncovered in the study 1. This allows us to test the effects of these strategies on viewers in form of their motivation to engage with the influencer and purchase intention of the influencer's product. Our study 2 supports five hypotheses proposed (H1 to H5), proving that different strategy combinations indeed generate different reactions from the audience. Specifically, videos with clarified jargon, high-arousal voices, and happy face can increase purchase intention by generating better product evaluations, more positive audience emotions, and higher credibility and authenticity. Study 2 provided the foundation of study 3 as it confirms that it is promising to look into influencer strategies in a more detailed angle.

Lastly, we identified 50 influencers on Douyin (Chinese TikTok) and recorded 600 videos from them (12 videos from each influencer from January to March 2023) covering both domestic and foreign product brands. We then analyzed these data regarding rhetoric, voice, and facial expressions to see if our hypotheses received support.

Study 3 supports four of the six hypotheses (H6 to H11) and underscores that videos with good sales performance tend to use more clarified jargon, employ low arousal voice, and display less happy faces. The surprising findings of the negative effect of voice arousal and happy face lead to further attention in the context of research. It is essential not to neglect that Chinese consumers switched their tastes based on their mental desires. Specifically, they prefer a more elegant and quiet vibe in livestream. Our Chinese context gives inspiration for further research on livestream selling that strategies serve to context, and they may not hold the same effect across different contexts.

Moreover, we employed multi-modality analysis in study 3 to explore the interactions among content, voice, and facial expressions. As a format-rich media, live stream selling has visual effects that convey detailed information through the combination of images, audio, and video (Elkin, 2002). Research reveals that combining these features with the textual content heightens the vividness and efficacy of communication (Appiah, 2006; Lee & Shin, 2022). Such multi-modality media can increase credibility and engagement intentions (Lee & Shin, 2022). It is therefore important to examine the interactions between textual elements and voice or facial expressions. We use multi-modality analysis to clarify our findings, which show how emotions expressed via voice or facial expressions can enhance content strategies. Our result shows significant interactions between jargon and voice arousal and between voice arousal and happy face. For example, we found the boundary condition of employing clarified jargon that the positive effect exists only when combined with a low-arousal voice.

To the best of our knowledge, the research reported here is one of the first to explore the content and delivery strategies of livestream influencers using a multi-dimensional and mixed-methods approach. Empirically, it sheds light on other influencers and platforms on how to build

viewer engagement and increase livestream sales. Theoretically, the novel approach to this emerging yet burgeoning phenomenon allows us to test the combined and synergistic effects of rhetoric, voice, and facial expression features and expand our understanding of this relatively new livestream sales strategy.

The remainder of this dissertation is organized as follows: We first conducted a study 1 to explore the unique content delivery strategies of top, middle and bottom influences. We then use the videos from the study 1 as the mocking stimuli to run an experiment (study 2). Next, we survey the relevant literature on livestreaming. Based on that literature, we propose six hypotheses that focus on the factors that influence the effectiveness of livestreaming. Furthermore, we describe our research methods and present our empirical results (study 3). The dissertation concludes with a general discussion of our findings.

Research on Live Streaming

Early research on livestreaming. Considerable research exists that focuses on the pure interaction-oriented forms of live streaming like video gaming (e.g., on Twitch) and influencers chatting with followers (e.g., on Instagram). This research category includes those studying the live streaming platform/features and influencers. For example, the experience of para-social interaction affects the intention to watch continuously on Chinese TikTok TV and YY Live (Hu et al., 2017). Influencer attractiveness is important in consumer engagement, but it differs slightly according to gender (Chen & Lin, 2018). Females have a more favorable impression of a livestream when they *admire* the influencer, but males develop a more favorable impression of a livestream when they *interact* with the influencer or the audience. Lin et al. (2021) reported that the influencer's intentional use of happy facial expressions positively affects the number of tips

audiences give during a live stream. By employing more happy faces, influencers received more tips from the audience. A study focused on Twitch (Zhao. et al., 2019) found that professional players are more likely to attract a larger viewer base, and that social affordance (e.g., profile building affordance, social connectivity) was positively associated with popularity.

However, the livestreaming retail format departs from the pure-interactive live stream because it uniquely blends both the "show" and the "selling" action. The ultimate goal is to use the interactive function with the audience to achieve the sales goals. Some research has examined the detailed strategies that influencers should utilize. For example, Bharadwaj et al. (2022) explored the effect of emotional displays on television sales and found a negative effect of emotional displays on sales. Their work suggested that salespeople on TV shows should sell with neutral facial expressions (straight faces).

A unidimensional perspective characterizes the literature on live stream research, and the isolated impact of individual key variables is common. This analysis strategy becomes questionable given the dynamic nature of livestream selling. A comprehensive and multi-dimensional approach is needed to understand the content and delivery strategies of the influencers in a holistic way. In addressing this gap, our study diverges from the traditional approach as we do an in-depth examination of livestream influencers' content strategies through multidimensional analysis, including rhetoric, audio, and facial expressions.

Multi-modality Communication. Research suggests that the incorporation of images and videos into textual content augments the vividness and efficacy of communication (Appiah, 2006; Lee & Shin, 2022). In our research framework, textual content is foundational, but we recognize that livestream selling is a rich medium that furnishes visual effects that convey complex information by amalgamating images, audio, and video (Elkin, 2002). Research on

multi-modality has shed light on how the combination of diverse communication modes (e.g., text combined with images, text combined with audio, and visual elements) augments the effectiveness of persuasion and communication (Lee & Shin, 2022). Empirical studies of multi-modality communication have demonstrated that individuals exposed to video-based deepfake news exhibited a heightened sense of source vividness compared to those who were exposed to fake news presented through alternative modalities (i.e., text-only and text-photo). This resulted in increased credibility and intentions for engagement with the deceptive content (Lee & Shin, 2022).

Human cognitive processing of information is enriched when information is presented through multiple sensory channels (van Wassenhove et al., 2005). Park et al. (2014) examined how the amalgamation of multiple communication modalities (i.e., audio, text, and visual) for predicting persuasiveness compared with the usage of a singular modality in isolation. They found that using multiple communication modalities (audio, text, and visual) significantly heightened predictive accuracy in persuasion in contrast to relying solely on using a single modality.

The multi-modality facet of live stream selling functions as a distinctive interface through which consumers engage during the shopping process. Research on multi-modality interfaces has demonstrated that internet browsers show a more favorable attitude towards a website featuring products when the site incorporates audio *and* video features, as opposed to those containing solely text (Appiah, 2006). This outcome is attributed to the vividness effects of multi-modality media, as information in a multi-modality format empowers advertisers to harness and capitalize on the multi-sensory engagement of sight, sound, and motion. Likewise, multi-modality media positively affects the attitude about online videos' attitude and sharing intention (Hsieh et al.,

2012). Multi-modality interfaces offer great flexibility and reliability to accommodate the diverse needs and preferences of users (Xiao et al., 2002; Xiao et al., 2003; Oviatt et al., 2005; Bohus & Horvitz, 2010). These interfaces also have greater adaptability in dynamically changing environmental conditions (Oviatt et al., 2000).

In our research, we emphasize the significance of the interplay between textual elements and emotional cues of vocal and facial expressions. The literature substantiates the advantages of multi-modality communication. We employed a sophisticated methodology known as multi-modality analysis aimed at discerning whether emotions conveyed through vocal cues or facial expressions could augment the efficacy of rhetorical strategy, thereby elevating sales performance. Because influencer marketing has been primarily characterized by text-based content, the content strategy is also a foundation of our investigation.

Study 1: Exploratory study

Sampling of Influencers. We adopted the stratification method (based on the top 500 list) to select the focal influencers for our analysis. The rank is provided by Huitun.com, a leading website in China for rankings and statistics of livestream influencers. Huitun generates rankings of livestream sales influencers based on their Huitun score, which is a weighted score of the fan size, comments, and likes in the live videos on influences' Taobao accounts within a period of time (could be a day, week, month, or year from the search date; here we use the year for maximum comprehensiveness and inclusiveness). The scores and rankings are dynamic.

We collected the ranking information in April 2020. The 1st to 100th ranked influencers constitute the top 20% of the list (top influencer group), the 150th to 300th ranked are the mid 30% (middle influencer group), and the 350th to 500th ranked are the bottom 30% of the list (bottom influencer group). We then randomly selected two influencers from each influencer

group within the same industry. For example, we randomly selected two influencers from the top 20% of the list (top influencer group).

In the prior work on social media influencers, the frequently used metric is the simple number of followers. Micro-influencers have 10,000 to 100,000 followers, and mega-influencers have more than 1 million followers (Britt et al., 2020; Park et al., 2021). However, we think this single measure is insufficient to represent influencer success because it only reflects engagement performance. The livestream influencers's performance is sales-driven. Therefore, Huitun provides a more impactful way to rank these influencers as both their abilities to drive social media engagement and sales are captured. Table 1 provides a descriptive profile of the six influencers we selected.

-- Insert Table 1 here --

Sampling of influencers' livestream videos. We monitored these six influencers and their livestreaming sessions between May and November of 2020. As the six randomly selected influencers were all selling in the cosmetics and personal care industry, we recorded the sessions when they sold lipsticks, serums, facial masks, and make-up removers. Table 2 lists the video sample we generated for our final analysis.

-- Insert Table 2 here --

Analytic plan. We adopted three-V analysis (visual, vocal, and verbal) widely used in social psychology and communication literature (Mehrabian, 1972; Strahan & Zytowski, 1976; Krauss et al., 1981; Knapp et al., 2013; Hu & Ma, 2021).

Content analysis to detect rhetoric styles. We hired two research assistants from a Chinese university to manually analyze the scripts of the videos. Their task was to detect distinguished rhetoric styles employed by the influencers. The main emerging rhetoric style the coders

identified manually was jargon usage. Two coders first marked all jargon in the text, labeling those with and without explanation, respectively. The patterns were easy to identify: top influencers used jargon with explanation more than middle and bottom influencers. For example, the bottom influencers kept repeating the jargon but did not provide any effective explanations.

Jargon use. Jargon is the special words or expressions used by a particular profession or group that are difficult for others to understand (Oxford English Dictionary). Livestream sales require the audience to trust the influencer and the product within minutes or even seconds; how to generate trust effectively is the key to success. In China, online paid knowledge is a significant trend in self-improvement. The desire to gain knowledge and seek professionalism drives this phenomenon (Zhou et al., 2022). One prominent example is the Fandeng Reading community (now called Fanshu), where paid members can read book reviews and share their critiques.

Associating with self-improvement through learning, speaking ability is a key benefit gained from acquiring knowledge. Research has demonstrated that people experiencing low status tend to use jargon to influence others' evaluations (Brown et al., 2020). This status compensation function is particularly effective in persuasion and sales. A heightened social status can amplify the perceived power difference between influencers and the audience, a phenomenon particularly salient in our examination given that Douyin's audience predominantly comes from smaller municipalities with comparatively lower educational attainment (2021 Douyin E-commerce Ecological Development Report). The heightened perceived power wielded by influencers bolsters favorable attitudes towards the brand through perceptions of influencer expertise and trustworthiness (Nafees et al., 2021), ultimately leading to a positive increase in purchase intention (Weismueller et al., 2020).

The heightened perceived power of influencers also brings authority, which enhances the persuasion effect. Authority is one of the principles that can enhance persuasion (Cialdini, 2001). Specifically, people are more likely to obey a person in a high authority position. Messages from credible sources (those with perceived authority) are more persuasive and result in greater attitude change compared to messages from less credible sources (Hovland & Weiss, 1951). Eagly and Chaiken (1993) suggest that authority cues can increase persuasion, especially when individuals are not motivated or able to process information systematically.

Moreover, the concern about audience evaluation drives the usage of jargon (Brown et al., 2020) because influencer and product evaluation directly affect audience purchasing behavior. The counselor's knowledge of psychology was rated higher when he employed jargon than when he used layman's language (Atkinson & Carskaddon, 1975). Similarly, participants have higher credibility of information when the expertise uses medical jargon (Zimmermann & Jucks, 2018). These literature studied in-person communication on the usage of jargon.

However, livestream selling presents a unique context where influencers must sell products quickly and with minimal audience involvement. Additionally, because livestreaming is often viewed as entertainment, audiences tend to treat it less seriously compared to face-to-face purchases. Despite these differences, theories from earlier literature provide a foundation for our hypothesis. Given the Chinese people's strong desire for and belief in knowledge, using jargon—an indicator of expertise and authority—can be crucial in winning the audience's favor.

Our coders identified the different jargon style among top, middle, and bottom influencers. Jiaqi Li is famous for his expertise in cosmetics, which is manifested by his great usage of jargon. Our coders also took a look at another top influencer, discovering that she also pays great attention to employing jargon when introducing products. Deeply investigating their jargon style,

we found that top-tier influencers use jargon with *clear and straightforward explanations*. For example, top influencer Jiaqi Li introduces one product in this way:

> The essence is 100%, and the gold foil in it has the purity of 98.1%. Gold foil can help you release negative ions. It can speed up water microcirculation in our skin and enhance the absorption capacity of our skin. The peptide is a small molecule of bisulphite; it can help us protect collagen, weakening fine lines and repair (Jiaqi Li, 2020).

He lets the audience know what gold foil and peptides are and what they can do to the skin. However, when mid-tier influencers employ jargon, they rarely provide additional explanations. Middle influencer Bingni Han, for example, introduces a product like this:

> The product I am holding is a product from the medical brand. Medical brand, whose production requirements are very high, conducts aseptic production, which is in a very safe workspace. I am holding a mask that is cold-compressed medical level. The normal mask cannot be called cold-compressed. The only functional mask will be called cold-compressed (Bingni Han, 2020).

After listening to her introduction, it is still unclear what the medical brand is, what aseptic production and cold-compressed features are, and how they can benefit the skin. Taken together, we concluded that the feature of jargon that top-tier influencers usually adopt is clarified jargon.

Voice feature. We noticed the different levels of voice arousal among the three tiers of influencers. For example, in our sample of influencers, top influencer Jiaqi Li is famous for his loud voice, which many people joke will turn his children deaf by three years old. Such observation leads us to pay attention to influencers' voice quality, especially the arousal dimension.

Chinese people love creating a vibe, such as celebrating traditional holidays, weddings, birthdays, etc. Often, Chinese people are drawn to the excitement (or "kanrenao" in Chinese), which means they enjoy engaging with and gathering around popular, lively events (Li, 2022). A vibrant and noisy livestream can also be "fun" and attract an audience online. We describe the “renao” (fun in English) as a high-arousal event where both the voice and acting can trigger the audience's emotions. The importance of vibe stems from the nature of livestream selling. Influencers must sell products within minutes. The time is minimal for the audience to evaluate their purchase rationally. In this case, impulsive shopping could be the goal of influencers. Studies have demonstrated that arousal drives buying impulsiveness (Herabadi et al., 2009). Specifically, when a consumer experiences high arousal during their shopping, they are likely to engage in impulsive shopping.

We collaborated with Nemesysco Ltd, an Israeli high-tech voice analysis firm whose Layered Voice Analysis (LVA™) system (Nemesysco, 2015) has been commercially applied in call centers and sold as “the most sophisticated, flexible, cutting-edge voice analysis technology available today” (Nemesysco, 2016). Signal-processing algorithms extract and combine attributes from voices to identify, among other things, evidence of cognitive processes and specific emotional reactions. The algorithm identifies the interval between two continuous wave peaks as each voice unit, and each unit may last for a different length. The LVA system is language-independent.

In addition to commercial uses in fraud detection and call center monitoring, audio mining technology has been featured in recent marketing research. For instance, to examine how vocal tones impact persuasion, a recent study measured persuaders’ vocal tones in Kickstarter video pitches using the novel audio mining technology QA5 from Nemesysco Ltd (Wang et al., 2021).

Mayew and Venkatachalanom (2012) investigated the underlying mechanism and intuitions of emotional analysis and concluded that the measures produced by LVA are consistent with some of the traditional measures of acoustic analysis. The scale of the voice arousal is from 0 to 30, with a higher number indicating a higher degree of arousal in voice. We combine two top-tier influencers' audio results to form the top group, two mid-tier influencers' audio files for the mid group, and two bottom influencers' audio to form the bottom group. Table 3 shows summary statistics.

-- Insert Table 3 here --

We found significant differences in voice arousal among three groups ($F(2,6037)=36.28$, $p<.001$). Specifically, top influencer group adopts significantly higher voice arousal ($M=19.27$, $SD=10.97$) than the middle ($M=17.92$, $SD=9.80$) and bottom ($M=16.38$, $SD=9.93$) groups. Figure 1 shows the post hoc comparisons of the three groups. This result echoes our initial observations that top influencers especially Jiaqi Li employed high arousal voice to attract audience.

--Insert Figure 1 here --

Facial expressions. Recent research has demonstrated the value of analyzing image data for marketing insights (Li & Xie, 2020; Ordenes et al., 2019). Based on our observation, top influencer Jiaqi Li shows more happy expressions with his high-arousal voice.

Traditionally, Chinese culture places significant emphasis on facial expressions to convey information. The term "lianse" (referring to the color of the face or facial expression) is frequently referenced in Chinese literature and publications (Lu & Ahlsen, 2011). Additionally, Chinese culture recognizes various types of smiles, such as "peixiao" (a compensating smile) and "gegexiao" (a giggle). Although current AI technology cannot yet differentiate between these

nuanced smiles, the importance of a smile in the livestream context is clear. Given that livestreaming is a quick and low-involvement shopping format, the positive energy conveyed through happy facial expressions can enhance the likelihood of impulsive purchases (Herabadi et al., 2009).

To investigate whether the observations of top influencers and the importance of smiles remain true, we submitted the image data to the face application programming interface (API) to identify emotions. We adopted Microsoft's Azure Cognitive Service for replicability and transparency. Microsoft provides APIs to which we feed our raw images and obtain a host of face-related measures constructed by Microsoft Azure's machine learning algorithms. Microsoft Azure API has the highest prediction accuracy among the top five facial detection APIs, including Rekognition, Kairos, and FaceX (Malone & Burns, 2021).

The adoption of facial expression analysis has increasingly appeared in recent academic research. For example, Hu & Ma (2022) studied the persuasiveness of delivery in start-up pitches where the Face++ API and Microsoft Azure Cognitive Service were used to process full-pitch videos and quantify persuasion across visual, vocal, and verbal dimensions. Lin, Yao, and Chen (2021) examined the role of happy emotion in livestreaming tips. They employed a pre-trained deep neural network, which can detect happiness, surprise, disgust, anger, sadness, fear, and neutrality.

For every second of video footage, we sampled one frame (Yang et al., 2021; Yue-Hei Ng et al., 2015; Rajaram & Manchanda, 2020). The face API processed our image data and generated the automated coding of 11,791 images after excluding those that either did not have faces or had multiple faces. The scale of the happy face is from 0 to 1, with a higher number indicating a higher degree of happiness on the face. We combine two top-tier influencers' facial results to

form the top group, two mid-tier influencers' audio files for the mid group, and two bottom influencers' audios to form the bottom group. Table 4 shows summary statistics.

-- Insert Table 4 here –

We found that there is a significant difference in happy faces among the three groups (F(3,11788)=458.64, p<.001, R^2=.072). However, out of expectation, the top group does not have the highest happy face (M=.03). Instead, the middle group (M=.20) has a significantly higher happy face than both the bottom group (M=.10, p<.001) and top group (M=.03, p<.001). Figure 2 shows the post-hoc comparison for three groups. This result does not echo our observations. Nevertheless, it did bring some insights as the middle group indeed has a higher happy face than the bottom group. We posit that the happy face conveys the message, "This product is great, and I like it." Moreover, our sample size is very small, which can cause bias in the result. We will carry the happy face variable for later studies, where the larger sample size can generate more reliable conclusions.

--Insert Figure 2 here --

Discussion of study 1 (exploratory study) and preparation for study 2 (experiment). This study 1 gave us a profile of the text, audio, and facial expression patterns for influencers in each category. It provided important insights with regard to how various elements of rhetoric styles, voice, and facial expression features configure to jointly constitute different influencer profiles based on real-life videos of livestream influencers. Because the data from 24 videos is too small, there is no way to see which variable has what kind of effect. The purpose of Study 1 is to briefly *identify* the variables that could benefit livestream sales. In study 2, we aim to see if there is a positive effect when variables are combined. If so, we will collect more data to see the effect of each variable separately. The task of study 2 is to test whether the three profiles of videos, as

exemplified by each tier of influencers, can experimentally lead to the expected positive outcome in a more tightly controlled setting.

Study 2: Experiment

Hypothesis development

Top, middle, and bottom influencer videos on purchase intention.

We identified the top influencers' unique content delivery strategies in Study 1. We believe using these strategies could promote their livestream performance. In the Chinese livestream selling context, the focal standard of an influencer is the selling ability. In other words, the more they sell, the higher their ranking. Since we cannot make participants purchase the products in the lab setting, purchase intention is the direct measure of selling ability. We try to investigate whether there will be differences in purchase intention by adopting different strategies. We believe the unique strategies including using clarified jargon, high arousal voice and happy face can lead to participants' different reactions to purchase intention.

Old studies indeed built the foundations for our study. Literature provided evidence that clarified jargon, voice arousal and happy face could promote sales performance. Jargon increases investment willingness for investors with minimal industry knowledge because it increases perceived product premiums (Tan et al., 2019). Clarified jargon can also cancel the negative effect of jargon (e.g., understanding issues, Dolan, 2019). Moreover, voice arousal enhances the audience's emotions and persuasion. Literature has defined high arousal voice as the one that has high energy (Mehrabian & Russell, 1974). An entrepreneur's high-arousal voice increases perceptions of their passion (Allison et al., 2022). Consumers reciprocate the enthusiasm of the retail salespeople with more positive emotions, which benefits sales performance and customer

satisfaction (Lee & Dubinsky, 2003). In the context of a phone survey, interviewers with high-pitched voices induce more favorable perceptions of enthusiasm (Oksenberg et al., 1986) and competence as well as lower refusal rates (Chebat et al., 2007; Oksenberg et al., 1986). As for happy face, literature demonstrated that the Duchenne (or genuine) smile can enhance the authentic persuasion of a celebrity endorser (Ilicic & Baxter, 2018). Besides, emotional experience has a significant positive impact on the purchase intention (Tellis et al., 2019; Nasermoadeli & Maghnati, 2013), we believe the positive emotional experience created by a happy face in livestream will have an impact on sales performance (which is highly correlated with consumers' purchase intentions).

One of the key differences between our study and previous research is that our context is very new and flexible. In face-to-face or online shopping scenarios, the audience has more time to evaluate their purchases. They can take the time to decide if the salesperson is trustworthy and if the product is as good as described. However, livestream selling only offers minutes for the audience to make decisions. While the price benefits of livestreaming are very tempting for the audience, the pressure on influencers is also high—they have to sell as many products as possible, quickly. In such a short time, the level of engagement is also much lower compared to traditional shopping. It is important for us to determine if the old theories still hold in this emerging context. Based on the logic of previous studies, we first proposed that:

H1: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention.

Credibility, authenticity

It has been supported that the counselor's knowledge of psychology was rated higher when he employed psychological jargon than when he used concrete layman's language (Atkinson &

Carskaddon, 1975). Moreover, interviewers with high-pitched voices induce higher competence and lower refusal rates (Chebat et al., 2007; Oksenberg et al., 1986). Lastly, people were more likely to believe the confederate's statement when it was paired with a deliberate Duchenne smile (Reed et al., 2018), showing that happy face improved perceived credibility.

Credibility and authenticity are heavily studied in the literature related to purchase intention. Many studies have pointed out the positive effect of endorser/influencer credibility and authenticity on product purchase intention. Kim and Song (2020) suggested that credibility can promote purchase intention via competence and authenticity. Similarly, Wang et al. (2017) added that endorser credibility positively impacts brand attitude, brand credibility, and purchase intention. Salesperson credibility positively affects product evaluation and buying intention when buyers have no prior knowledge of a brand, which is very similar to our livestream samples in which many unpopular brands are being promoted (Sharma, 1990). The message from a noncredible salesperson may not be fully processed, and the message content may be discounted. However, the message from a highly credible salesperson is considered true compared to a message from a less credible salesperson (Aaker & Myers, 1975). Such a high credibility of information from a highly credible salesperson can also enhance purchase. Literature indicates that trustworthiness and information quality have significant positive direct effects on the credibility of influencers and significant indirect effects on consumers' purchase intention (Saima & Khan, 2020). Trust and commitment to the salesperson are directly linked with purchase intention as well as indirectly through store attitude (Macintosh & Lockshin, 1997).

Credibility is closely linked with power. Individuals with high credibility are often perceived as having significant power (Neser et al., 1993). In our livestream context, credibility is crucial because Chinese culture places a strong emphasis on power and authority from an early age.

Chinese parents, particularly those who uphold traditional values, are often associated with authoritative and authoritarian parenting styles (Xu et al., 2005). They are frequently described as "controlling" or "authoritarian" (Chao, 1994). Consequently, Chinese people tend to trust figures of high power or authority. Influencers can enhance their perceived power and credibility by using clarified jargon, high arousal voice and a happy face, ultimately driving purchases.

H2: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention because of the perceived higher authenticity and credibility.

Product evaluation

Another factor affecting purchase intention is product evaluation (how consumers evaluate a product). As we mentioned earlier, using jargon can improve patients' evaluation of the psychology counselor because jargon represents knowledge (Atkinson & Carskaddon, 1975). Also, voice arousal increases perceptions of a salesperson's passion (Allison et al., 2022) and consumers reciprocate that passion with more positive emotions, which benefits sales performance and customer satisfaction (Lee & Dubinsky, 2003). Similarly, happy facial expressions can also boost product evaluation. A research demonstrated that the t-shirt was rated more positively when the model had a genuine smile (vs. neutral face) (Peace et al., 2006).

Literature often studies product evaluation and purchase intention simultaneously, treating them as downstream variables in consumer behavior studies. One of the works shows that a more positive product evaluation can yield a greater intention to purchase (Cavazza & Gabrielli, 2015). Similarly, when participants had positive product evaluations, they were more likely to have higher purchase intention (Kim et al., 2012). The literature supports the importance of product evaluation in traditional shopping, where purchases typically take longer, and consumer involvement is higher. However, no one has studied whether, in the context of livestream selling,

product evaluation remains key to purchase intention. We believe that positive product evaluations stem from the perceived high power and authority of influencers who use specific techniques (e.g., clear jargon). Our study points to the same conclusion as previous research, despite the different mechanisms described in the older literature.

H3: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention because they have more positive product evaluations.

Audience emotions

Audience emotion is also a key factor in livestream sales, which can be explained by the emotional contagion theory. Literature has supported that using jargon can increase perceived knowledgeability (Aktinson &vCarskaddon, 1975). Such expertise can generate consumer's positive emotions during purchasing (Lee & Dubinsky, 2013). Moreover, voice arousal benefits the perceived passion (Allison et al., 2022; Oksenberg et al., 1986). Consumers usually reciprocate the such passion and ethusiasm of the salespeople with more positive emotions (Lee & Dubinsky, 2003). Lastly, happy facial expression signal a positive emotions, such as desires to affiliate with and/or cooperate with another person (Rychlowska et al., 2017), and provide support and encouragement (Guyer et al., 2019).

Customer emotions have beneficial effects for both the store and the salesperson. (Lee & Dubinsky, 2003). Consumers' emotional state has impact on purchase intention in the online shopping environment (Kim & Lennon, 2013). Consumers will have positive attitudinal bias toward a product when they encountered happy delivery person (Howard & Gengler, 2001). Such positive emotions can lead to purchase intention (Cavazza & Gabrielli, 2015; Kim et al, 2012). The more that customers perceive retail salespeople to exhibit friendliness, the greater the degree of customers' positive emotions (Lee, & Dubinsky, 2003). This makes so much sense

because the emotional experience significantly positively impacts the purchase intention (Nasermoadeli et al., 2013).

Considering that livestream selling happens very quickly for each item and the audience is less serious, impulsive shopping is likely to be the typical shopping pattern. The audience may not have enough time or cognitive ability to evaluate their purchases rationally. Often, Chinese people are drawn to the excitement (or "kanrenao" in Chinese), which means they enjoy engaging with and gathering around popular, lively events (Li, 2022). A vibrant and noisy livestream can also be "fun" and attract an audience online. Therefore, in our study, audience emotion is crucial in driving purchases.

H4: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention because audience has more positive emotions.

H5: The main effect of video types on product purchase intention (as in H1) will be mediated by perceived authenticity and credibility, product evaluation, and audience emotions of the influencer. (see Figure 3 of theoretical model of the following hypotheses):

-- Insert Figure 3 here --

Preparation for Influencer Videos. In the experiment, we hired both a Canadian and an English-speaking Chinese actress who played an influencer role with the fictitious name, Hyler Prince, and recorded a short video selling lotion products (see Figure 4 for screenshots of three conditions by both Canadian and Chinese actresses). We created a background screen in the shooting studio with the influencer's name on it. We opted to choose the lotion product from the cosmetics sector. A fictitious brand for the lotion ("Amazing Journey") was selected for two reasons: (1) it allowed us to cleanly manipulate the independent variable while controlling for the brand name, brand familiarity, and product category across conditions; (2) it enabled us to test if

the effects discovered in study 1 are robust enough to relatively unknown or new brands. We created three video footage conditions - top-, mid-, and bottom-tier influencer videos.

--Insert Figure 4 and Table 5 here--

Scripts and training of actress. As shown in Table 5, in each video condition, we directly adopted the style elements from the study 1 to replicate the language styles of clarified jargon. We trained our actress in terms of how to play the role of a top-, mid-, and bottom-tier influencer, respectively, in terms of their voice features and their facial expressions. We used our findings from the study 1 and created a style sheet for her when she played each type of influencer. In the multiple training sessions, we also showed the original livestream videos of each influencer type so that the actress has ideas of the voice and facial expression features of each influencer profile. After a few training sessions, we recorded three versions of videos with the actress playing as the top, mid, and bottom tier influencer, respectively. Then, we used Microsoft Azure Cognitive Service to analyze the voice and facial expressions to make sure the desired differences among top-, mid-, and bottom-tier influencer styles were achieved. Table 6 summarizes the analyses of the actress's videos so that the desired differences in text, audio (arousal level) and facial expression (happy face) are achieved. Lastly yet importantly, we controlled the script length and video length so that each video is similar in its length. The word counts for top, mid, and bottom conditions were 795, 799, and 806, respectively. The video lengths of the Chinese actress turned out longer than of Canadian actress (05:03, 04:45, 04:51) due to the acting style (e.g., Chinese actress showing surprising face takes longer).

--Insert Table 6 here--

Power analysis. We used Monte Carlo Power Analysis (Schoemann et al., 2017) to estimate the sample size. Since our independent variable and the whole model are novel, we would

conservatively expect a small standardized coefficient (0.1) on every path (a1, a2, a3, b1, b2, b3, and c'). The estimation suggested sample size with a power of 0.9 is 1330. For direct queries, attention check (e.g., "please mark xx and move on"), which was what we plan to adopt, the average loss of data is 13.96 percent, with a median of 10.60 percent (Abbey & Meloy, 2017). Besides, since our video is long and participants are easily likely to get tired, we decided to collect 1,800 responses to achieve the expected statistical power.

Method. One thousand eight hundred participants were recruited globally from Prolific. Ninety one failed attention checks and were thus excluded. 1,709 final subjects were retained for the analyses (M_{age}=30.94, 49.7% female) and got randomly assigned to one of the six influencer video conditions (three to the Chinese actress and threeto the Canadian actress).

Participants were instructed to watch the entire video because the "next" button shows up only after the video ends. They were then asked to answer a survey with the dependent variables. Because the products sold in the video are body lotions with feminine scents, participants were asked about their gender first. For male participants, they were asked to imagine that they were going to buy a body lotion as a birthday gift for their female friend or acquaintance. For female participants, they were asked to imagine that they were going to buy a body lotion. Some scales were adapted to suit male participants. For example, each question about purchase intention for male participants (Kim & Kim, 2021) ended with "as a birthday gift."

Results. We created two dummy variables for the mid (X1) and the bottom (X2) conditions and treated the top condition as the baseline.

Audience gender. The result revealed that the moderation effect of audience gender was not significant between influencer videos and purchase intention ($F(2,1703)=.78$, $p=.46$), product attitudes ($F(2,1703)=.21$, $p=.81$), authenticity and credibility ($F(2,1703)=1.37$, $p=.26$), and

between videos and audience emotion (F(2,1703)=.11, *p*=.89). Hence, in the subsequent analyses, we treated audience gender as the control variable.

Influencer ethnicity. The moderation effect of influencer ethnicity was significant between influencer videos and product attitude (F(2,1703)=3.37, p=.03 ΔR2=.004; Chinese influencer condition X1: B=-.27, SE=.14, 95% CI -.54, .002; Chinese influencer condition X2: B=-.70, SE=.14, 95% CI -.96, -.43; Canadian influencer condition X1: B=-.52, SE=.14, 95% CI -.79, -.25; Canadian influencer condition X2: B=-.45, SE=.14, 95% CI -.96, -.19). The moderation effect of influencer ethnicity was marginally significant between influencer videos and emotional experience (F(2,1703)=2.81, p=.06, ΔR2=.003; Chinese X1: B=-.36, SE=.13, 95% CI -.60, -.11; Chinese X2: B=-.80, SE=.12, 95% CI -1.05, -.56; Canadian X1: B=-.50, SE=.12, 95% CI -.75, -.26; Canadian X2: B=-.54, SE=.12, 95% CI -.79, -.30). A non-significant interaction effect emerged between influencer videos and purchase intention (F(2,1703)=1.70, p=.18, ΔR2=.002), and between influencer videos and authenticity (F(2,1703)=.80, p=.45, ΔR2=.001).

We suspect that the strength of the effect of influencer videos on some variables (e.g., product attitudes) is significantly different because our two actresses have different styles. But as we expected, the effects of influencer videos are significant on dependent variables in both Chinese and Canadian influencer conditions. Hence, even with the slight variations in the strength of the focal effect, in the subsequent analyses, we treated influencer ethnicity as the control variable.

Purchase intention. A one-way ANCOVA with audience gender and influencer ethnicity as the control variables indicated that influencer videos significantly affected purchase intention (F(2,1703)=15.54, *p*<.001, R^2=.12). Respondents watching the top condition video (M=2.68, SD=1.73) reported higher purchase intention than the mid (M=2.27, SD=1.59, *p*<.001) or bottom condition respondents (M=2.22, SD=1.53, *p*<.001). H1 has been supported.

Credibility and authenticity. As expected, a one-way ANCOVA with audience gender and influencer ethnicity as the control variables demonstrated that influencer videos significantly affected authenticity and credibility ($F(2,1703)=25.36$, $p<.001$, $R^2=.09$). Participants in the top condition (M=3.37, SD=1.43) indicated higher authenticity and credibility than participants in the mid (M =2.97, SD=1.51, $p<.001$) or bottom condition (M=3.78, SD=1.46, $p<.001$).

Next, we run the mediation analysis with influencer videos (X1: top vs. mid) as the independent variable, purchase intention as the dependent variable, and authenticity and credibility as mediators while controlling for the audience gender and influencer ethnicity. A bootstrapping analysis (5,000 resamples) showed that authenticity and credibility (indirect effect=-.30, SE=.06, 95% CI -.43, -.18) mediated the relationship between influencer videos (X1) and purchase intention. Specifically, the influencer video (X1) has a significant negative effect on authenticity and credibility (b=-.40, t=-4.71, p<.001), which later has a positive effect on purchase intention (b=.76, t=41.04, p<.001). The direct effect of influencer video (X1) on purchase intention was insignificant (b=-.11, SE=.07, 95% CI -.24, .02).

We then ran a second parallel mediation analysis (Model 4) for our second dummy variable, X2 (top vs. bottom). A bootstrapping analysis (5,000 resamples) uncovered that authenticity and credibility (indirect effect=-.45, SE=.06, 95% CI -.57, -.32) explained the link between influencer videos (X2) and purchase intention. Specifically, the influencer video (X2) has a significant negative effect on authenticity and credibility (b=-.59, t=-7.00, p<.001), which later has a positive effect on purchase intention (b=.76, t=41.04, p<.001). Again, the relationship between influencer videos (X2) and purchase intention was insignificant when authenticity and credibility was included in the model *(*direct effect=-.01, SE=.07, 95% CI -.14, .11*).* These results supported H2.

Product attitudes. As anticipated, a one-way ANCOVA with audience gender and influencer ethnicity as the control variables showed that influencer videos had a significant impact on consumers' product attitude ($F(2,1704)=19.70$, $p<.001$, $R^2=.08$). That is, participants in the top condition (M=3.48, SD=1.67) indicated more favorable product attitude than participants in the middle condition (M=3.07, SD=1.67, $p<.001$) or the bottom condition (M=2.90, SD=1.59, $p<.001$). H3 did receive support.

Next, we run the mediation analysis with influencer videos (X1: top vs. mid) as the independent variable, purchase intention as the dependent variable, and product attitude as mediators while controlling for the audience gender and influencer ethnicity. A bootstrapping analysis (5,000 resamples) showed that product attitude (indirect effect=-.29, SE=.07, 95% CI -.43, -.16) mediated the relationship between influencer videos (X1) and purchase intention. Specifically, the influencer video (X1) has a significant negative effect on product attitudes (b=-.41, t=-4.29, p<.001), which later has a positive effect on purchase intention (b=.72, t=47.04, p<.001). However, the direct effect of influencer video (X1) on purchase intention was significant (b=-.12, SE=.06, 95% CI -.24, .-.00).

We finally ran a second parallel mediation analysis (Model 4) for our second dummy variable, X2 (top vs. bottom). A bootstrapping analysis (5,000 resamples) uncovered that product attitude (indirect effect=-.42, SE=.07, 95% CI -.55, -.29) explained the link between influencer videos (X2) and purchase intention. Specifically, the influencer video (X2) has a significant negative effect on product attitude (b=-.58, t=-6.13, p<.001), which later has a positive effect on purchase intention (b=.72, t=47.04, p<.001). The relationship between influencer videos (X2) and purchase intention was insignificant when product attitude was included in the model *(*direct effect=-.04, SE=.06, 95% CI -.16, .07*).* These results supported H3.

Audience emotion. A one-way ANCOVA with audience gender and influencer ethnicity as the control variables indicated that influencer videos did impact audience emotion ($F(2,1704)=30.47$, $p<.001$, $R^2=.06$). Respondents in the top condition (M=2.78, SD=1.63) reported higher audience emotions than participants in the mid condition (M=2.33, SD=1.49, $p<.001$) or the bottom one (M=2.10, SD=1.36, $p<.001$). H4 is supported.

Next, we run the mediation analysis with influencer videos (X1: top vs. mid) as the independent variable, purchase intention as the dependent variable, and audience emotion as mediators while controlling for the audience gender and influencer ethnicity. A bootstrapping analysis (5,000 resamples) showed that audience emotion (indirect effect=-.31, SE=.07, 95% CI -.45, -.19) mediated the relationship between influencer videos (X1) and purchase intention. Specifically, the influencer video (X1) has a significant negative effect on audience emotion (b=-.44, t=-4.93, p<.001), which later has a positive effect on purchase intention (b=.72, t=40.02, p<.001). The direct effect of influencer video (X1) on purchase intention was insignificant (b=-.10, SE=.07, 95% CI -.23, .03).

We finally ran a second parallel mediation analysis (Model 4) for our second dummy variable, X2 (top vs. bottom). A bootstrapping analysis (5,000 resamples) uncovered that audience emotion (indirect effect=-.49, SE=.07, 95% CI -.62, -.36) explained the link between influencer videos (X2) and purchase intention. Specifically, the influencer video (X2) has a significant negative effect on audience emotion (b=-.67, t=-7.71, p<.001), which later has a positive effect on purchase intention (b=.72, t=40.02, p<.001). The relationship between influencer videos (X2) and purchase intention was insignificant when audience emotion was included in the model *(*direct effect=.03, SE=.07, 95% CI -.11, .16*).* These results supported H4.

Lastly, we run the parallel mediation analysis (Model 4) with influencer videos (X1: top vs. mid) as the independent variable, purchase intention as the dependent variable, and authenticity and credibility, product attitudes as well as audience emotion as mediators, while controlling for the audience gender and influencer ethnicity. A bootstrapping analysis (5,000 resamples) showed that authenticity and credibility (indirect effect=-.09, SE=.02, 95% CI -.13, -.05), product attitude (indirect effect=-.17, SE=.04, 95% CI -.25, -.09) and audience emotions (indirect effect=-.20, SE=.03, 95% CI -.27, -.14) mediated the relationship between influencer videos (X1) and purchase intention. The relationship between influencer videos (X1) and purchase intention was insignificant when authenticity and credibility, product attitudes as well as audience emotion were included in the model (direct effect=-.03, SE=.05, 95% CI -.14, .07).

Next, we ran a second parallel mediation analysis (Model 4) for our second dummy variable, X2 (top vs. bottom). A bootstrapping analysis (5,000 resamples) uncovered that authenticity and credibility (indirect effect=-.13, SE=.02, 95% CI -.18, -.08), product attitudes (indirect effect=-.24, SE=.04, 95% CI -.32, -.16), and audience emotion (indirect effect=-.67, SE=.09, 95% CI -.85, -.50) explained the link between influencer videos (X2) and purchase intention. The relationship between influencer videos (X2) and purchase intention was insignificant when authenticity and credibility, product attitudes as well as audience emotion were included in the model (direct effect=.10, SE=.05, 95% CI -.00, .21). These results supported H5.

Discussion. Study 2 intended to experimentally replicate the core effects observed in the real livestream context in Study 1 while providing the preliminary mechanism evidence. Additionally, this experiment sought to extend the findings from Study 1 into a non-Chinese context. As expected, the product promoted by the top influencer received higher purchase intention than when promoted by the mid and bottom influencers, respectively. The top

influencer condition also showed higher authenticity and credibility and triggered more positive emotional arousal among the audience than the mid and bottom influencer conditions. Lastly, the product promoted by top influencers received a more favorable attitude from the audience than the product promoted by the bottom influencer but not the mid influencer.

In addition, the core effect of influencer videos on purchase intention is mediated by the top influencer's perceived authenticity and credibility, product evaluation, and the positive emotional arousal and experiences of the participants as compared to the mid and bottom influencer conditions (H2, H3, H4 and H5).

One limitation of the current study is that the dependent variable, on average, have low mean scores (below 3). We believe this is because we use a fictitious brand with acting (vs. real) influencer. This side effect of an experiment context is not unproceedented. For example, in Rajagopal and Montgomery's work (2011), they used a fictitious brand as the stimuli, receiving low score (M=2.5; M=2.3) on the product experience belief confidence (on a 7-point scale) where as this PEBC score is higher in the real brand condition (M=4.0; M=6.1). To this end, prior research also suggests that brand familiarity could influence a consumer's confidence towards the brand, which in turn affects their purchase intention of the brand (Laroche et al., 1996; Bergkvist, 2009). However, we take the risk of having a fairly lower average baseline evaluation of the product and influencer in the experimental setting so that we gain tight control and internal validity of the study context to observe the net differential effects of the content and delivery strategies of the influencer.

Since we received support that top influencers' content delivery strategies indeed promote their sales performance, our next step is to formally propose the hypothesis for each variable

based on literature and bring the study to a more realistic setting. We aim to collect more real livestream videos to deeply understand how these strategies work.

Study 3: Big Data Study of Douyin Livestream Videos

Study 2 (experiment) proved that those strategies (e.g., clarified jargon, voice arousal, and happy face) were very effective when mixed together, but we are still unclear which variables contribute more to sales performance. In study 3, we collected video data from Douyin to investigate those variables in detail. We began with literature review to find support for our exploratory findings for *each* variable. Based on the literature, we proposed six hypotheses. Next, we run statistic analysis on our 600 videos to verify if hypotheses were supported. Finally, we give our conclusion and discussion based on the findings.

Clarified jargon. Even though literature has pointed out the positive effect of using jargon, for example, Atkinson and Carskaddon (1975) concluded that a counselor was rated as more knowledgeable when employing abstract, psychological jargon than when using concrete, layperson's language. Similarly, jargon increased investment willingness for investors with minimal industry knowledge because it increased perceived product premium (Tan et al., 2019). A salesperson's credibility positively affects customers' product evaluation and purchase intention (Sharma, 1990) and expertise can generate consumer's positive emotions during purchasing (Lee & Dubinsky, 2013). A conflict of findings indicates that jargon can potentially cause confusion or decrease trust (Dolan, 2019), and only *clarified* jargon generates a positive effect. Castro et al. (2007) found that 81% of doctor-patient encounters contained at least one unclarified jargon term, leading to low ratings of patient comprehension. Therefore, to summarize the literature, we believe that only clarified jargon could contribute to sales. An

example of clarified jargon from live streamer Jiaqi Li is as follows: “The essence is 100%, and the gold foil in it has a purity of 98.1%. Gold foil can help you release negative ions. It can speed up water microcirculation in our skin and enhance the absorption capacity of our skin. The peptide is a small molecule of bisulphite; it can help us protect collagen, weakening fine lines and repair.” He lets the audience know what gold foil and peptides are and what they can do to the skin.

Besides, because of the worship of knowledge and authority of Chinese people, jargon as a way to showcase knowledge and enhance power should work greatly in driving persuasion (Zhou et al., 2022; Nafees et al., 2021; Weismueller et al., 2020; Cialdini, 2001).

Taken together, these points lead to the hypothesis:

H6: Clarified jargon has a positive effect on sales performance in livestream selling.

Audio analysis. We examined the voice arousal variable since it is as vital in sales. For example, Wiener and Chartrand (2014) examined the impact of an announcer's voice quality on purchase intentions in a retail setting. Female audience are most likely to purchase when they hear it advertised by a man who has a creaky voice instead of a woman with a creaky voice. Chattopadhyay et al. (2003) found that a voice with a faster-than-normal syllable speed can generate a more favorable response to an advertisement and improve brand attitudes. Consumers reciprocate the enthusiasm of the retail salespeople with more positive emotions, which benefits sales performance and customer satisfaction (Lee & Dubinsky, 2003).

Research in psychoacoustics and cognitive psychology indicates that high (vs. low), bright (vs. dull), and rough (vs. smooth) voices are more arousing (Laukka et al., 2005; Scherer, 1986), which means that such a voice can induce a subjective sense of energy (Mehrabian & Russell, 1974). Zoghaib (2019) notes that the subjective sense of energy induced by a speaker’s vocal

characteristics is called “voice arousal.” For example, an entrepreneur's high-arousal voice increases perceptions of their passion (Allison et al., 2022). In the context of a phone survey, interviewers with high-pitched voices induce more favorable perceptions of enthusiasm (Oksenberg et al., 1986) and competence as well as lower refusal rates (Chebat et al., 2007; Oksenberg et al., 1986).

As mentioned earlier, livestream selling is a unique format that includes both selling and entertainment. Thus, it is questionable if the previous findings in persuasion or phone sales can hold true. The abovementioned findings indeed support our study 1 that top influencers (who sold more) employed high arousal voices. However, when reviewing the literature on Chinese livestream sales, we found the opposite effect of voice arousal, that is, in the Douyin livestream research, consumers do not prefer the noisy live streaming style and find the quiet and gentle streaming more appealing (Lu et al., 2023).

One of the interesting shifting during our project is the livestream taste of Chinese consumers. Our samples in Study 1 were recorded in 2020 and 2021, and the follow-up experiment was done in 2022. Articles published in 2023 demonstrated the opposite conclusion that audiences prefer a low arousal vibe of livestream (e.g., Lu et al., 2023; Li, 2023). Indeed, gentle livestreamers such as Zhang Xiaohui, Dong Jie, and Dong Yuhui became popular in late 2022 or early 2023, leading to a new, quite luxurious style of livestream. Therefore, in this part, we believe that this shift should be considered in our Study 3 that voice arousal has a negative effect on livestream. The transition of livestream taste can be explained by the post-pandemic anxiety (Woodward et al., 2022), leading them to enjoy more elegant and peaceful livestream.

Research shows that the key success of Dongfangzhenxuan (Oriental Selection in English)’s live stream on Douyin is their style: “no yelling, no fighting, no vulgar words, and some are

poetic, civilized, romantic, comfortable, and elegant" (Zhang, 2022). Consumers regard livestream as a place of emotional dependence. The emergence and popularity of "slow live broadcast" demonstrates this point (e.g., "Zhang Xiaohui"'s slow live broadcast) (Li, 2023). Such livestream preference can be explained by consumers' mental states. There are considerable anxiety symptoms (e.g., poor sleeping quality, burnout) among residents of China (Bai et al., 2021). Research has shown that an increase in cognitive and physiological arousal increases anxiety (Noteboom et al., 2001). Since many Chinese consumers regard livestream as a place of emotional dependence (Li, 2023), it makes sense that Chinese consumers would not enjoy the high-arousal livestream, which could further increase their anxiety levels. Even though the study 1 shows a positive effect of high-arousal voice, our sample size is too small and only works as an exploratory study. In contrast, those papers studying Chinese livestreams are more empirically rigorous with larger sample sizes. Therefore, we propose that:

H7: Voice arousal has a negative effect on the sales performance in livestream selling.

Facial expression analysis. As the literature suggests, facial expression and voice are true representations of the mental state and self-state of the speaker (Schore, 2002), so they can increase emotional contagion, which facilitates persuasion, purchase intention, and attitudes (Lee & Dubinsky, 2003). Research has demonstrated the value of analyzing image data to gain marketing insights (Li & Xie, 2020; Ordenes et al., 2019). People often buy products not based on the product information or ratings, but rather on how the salesperson makes them feel (Fisher & Shapiro, 2005). The facial expressions of endorsers affect the perceived endorsers' social status and further affect consumers' reactions to the product being advertised (Chen & Wyer Jr., 2020). We considered the happy facial exprssion as it is vital in sales as the nonverbal communication. By definition, nonverbal communication means "communication without

words" (DeVito, 2005). The literature covered how effective leaders, salespersons, and customers use nonverbal communication to increase communication effectiveness. For example, facial expressions such as happiness, sadness, anger, fear, and surprise can be easily identified when conveying emotions (Salovey & Mayer, 1989). Postures can disclose comfort and engagement level (Kirmani & Campbell, 2004).

The significance of happy face should not be underestimated. A simple positive facial cue—the Duchenne (or genuine) smile—can enhance the authentic persuasion of a celebrity endorser (Ilicic & Baxter, 2018). Participants were more likely to believe the confederate's statement when it was paired with a deliberate Duchenne smile (Reed et al., 2018). A smiling (vs. frowning) communicator can generate more persuasion (Van Kleef et al., 2015). Smiling can signal a desire to affiliate with and/or cooperate with another person (Rychlowska et al., 2017), project confidence, communicate romantic interest, help others avoid embarrassment, and provide support and encouragement (Guyer et al., 2019). Because emotional experience has a significant positive impact on the purchase intention (Tellis et al., 2019; Nasermoadeli & Maghnati, 2013), we believe the positive emotional experience created by a happy face in livestream will have an impact on sales performance (which is highly correlated with consumers' purchase intentions). Therefore, we propose that:

H8: Happy facial expressions increase sales performance in livestream selling.

Multi-modality analysis. In research on Douyin platform, consumers dislike the noisy livestreaming style and find the quiet and gentle streaming more appealing (Lu et al., 2023). This is due to the fact that Chinese consumers have overall high anxiety levels, and high arousal environments could make their anxiety more severe. Therefore, high voice arousal livestream loses its charm and attention from the audience. This would likely counteract the positive effect

of clarified jargon since the audience does not pay attention anymore or simply just quits the live room. We therefore propose the following hypothesis:

H9: Voice arousal moderates the effect of clarified jargon on sales performance. When voice arousal is low, the positive effect of clarified jargon on sales is stronger. When voice arousal is high, the positive effect of clarified jargon is weaker.

Per our previous hypotheses, clarified jargon and happy faces can enhance persuasion and sales. When combining both strategies into a livestream, we believe sales performance can also be boosted even further. Opposite to a happy facial expression, a neutral face (which means a less happy face) has less attraction to the audience. Research demonstrates that faces with emotional expressions tend to attract more attention than neutral faces (Vuilleumiet, 2002). A neutral face can lead to high advertisement effectiveness only when it is a low-amusing theme (Lewinski et al., 2014), which is opposite to livestream selling as a highly entertaining format. Taken together, we believe that when there is a less happy face (more neutral facial expression), the audience's attention is low, which counteracts the positive effect of clarified jargon.

H10: A happy facial expression moderates the effect of clarified jargon on sales performance. When influencers display a happier face, the positive effect of clarified jargon on sales is stronger. When influencers display a less happier face, the positive effect of clarified jargon on sales is weaker.

Similarly, when an influencer has a less happy face (e.g., more neutral face), the negative effect of voice arousal is enhanced due to the loss of attention from the audience (see Figure 5 for the hypothese diagram).

H11: A happy facial expression moderates the effect of voice arousal on sales performance. When the happy face is high, the negative effect of voice arousal on sales is weaker. When the happy face is low, the negative effect of voice arousal on sales performance is stronger.

--Insert Figure 5--

To study the effect of each individual variable on sales performance, we collected 600 video segments on Douyin livestream. According to the *2021 Douyin E-Commerce Ecological Development Report*, the audience of Douyin is mostly Gen-Z, millennials, and the elder generation. The majority audience is mostly located in third or fourth-tier cities (e.g., Weifang, Zhenjiang, Baoding, etc. from Yicai news). We believe the audience group does not have high education. According to the seventh national census bulletin (No. 6), the average years of education for the region's population aged 15 and above are higher for first-tier city such as Beijing (12.64), Shanghai (11.81) compared with the province which has many third to fourth tier cities (e.g., 9.84 in Hebei province). The power distance between the influencer and the audience is large, since the audience is mostly less educated people from small cities.

Data on influencers and videos are extracted from Huitun, a professional third-party data provider of livestream platforms. Huitun ranks influencers based on sales, which helps us see how our independent variables directly affect sales.

Sampling. Livestream on Douyin has different genres. We focus on live streamers who are qualified based on the following criteria: they must have individual person accounts, they must livestream regularly, they must sell a great variety of products, and they must show faces. We identified fifty such influencers and collected videos from each of them. Our data collection period was from 7 January to 7 March 2023. Because Douyin does not offer recorded videos, we

recorded live-stream videos manually. Table 7 shows the descriptive information on the fifty influencers.

--Insert Table 7 here--

Due to schedule conflicts, internet speed, and pop-up live stream schedules, we could not randomly record live stream sessions of each influencer. Instead, we recorded four hours of livestream sessions for each influencer (see Figure 6 for the process of recording). To facilitate the recording process, our team enlisted the assistance of a professional technician consultant who provided a recording interface capable of handling up to twenty concurrent recordings. We refreshed the livestream pages once every hour to capture the available sessions. This process was repeated until each influencer's cumulative recorded time reached the four-hour mark. Within these four-hour recordings, multiple sessions were included, with variations observed among different influencers. For instance, some influencers had three sessions that collectively spanned four hours, while others had one session that extended the entire duration. The minimum number of sessions in the four hours was one, while the maximum was four. During each session, product videos of varying lengths were presented. Some sessions featured over twenty product videos, while others included only five. The minimum number of products sold in one session was seventeen, while the maximum was fifty-seven. Within the 4-hour videos of each influencer, we randomly select twelve items. These videos may be from different sessions of that influencer (if they have multiple sessions within the four hours) or from the same session (e.g., an influencer only has one session, but it lasts four hours). Ultimately, our dataset consisted of six hundred video clips (video length ranged from thirteen seconds to one thousand and one hundred and fifty-two seconds) sourced from ninety-four sessions. This sample size surpasses the upper thresholds observed in the existing literature on Chinese livestream selling, such as the

587 video clips examined by Luo et al. (2021), the 40 shopping sessions studied by Wu et al. (2022), and the 27 influencers analyzed by Lee & Theokary (2021).

--Insert Figure 6 here--

Methods.

Clarified jargon. We hired two research assistants from a Chinese university to manually and individually code the clarified jargon employed by the influencers. They first identified all the jargon usage and then labeled those with explanations. They compared their respective judgments, identified inconsistencies, and reached full consensus.

Audio and facial expression. We used the same methods as in Study 1 for voice and facial expression analysis. For each video, we take the average of its voice arousal and happy face score, which will be used in later regression analysis. For every second of video footage, we sampled one frame (Yang et al., 2021; Yue-Hei Ng et al., 2015; Rajaram & Manchanda, 2020) as Study 1.

Variables and model specification. The *dependent variable* is the sales that was generated by each video.We use the ln format of sales in the analysis since it will be more normally distributed, and the coefficient will not be too large or too small. The *independent variables* and their operationalizations are shown in Table 8. We controlled video duration (in second and in ln format), product final price (in ln format), assistant (have vs. not have), and influencer gender (male vs. female). We also use brand origin (domestic vs. foreign brand) as a control variable. Past studies have noted the impact of consumer ethnocentrism, illuminating Chinese consumers' favorable attitudes and heightened purchase intention towards domestic products (Wu et al., 2010; Xin & Seo, 2020). See Table 9 for the coding processing of control variables. X in Xb

denoted the matrix of duration, price, assistant, brand origin, and influencer ranking, influencer gender.

$$Ln(Sales) = \beta_0 + \beta_1 clarified\ jargon + \beta_2 voice\ arousal + \beta_3 happy\ face + \beta_4 clarified\ jargon * happy\ face * voice\ arousal + \beta_5 clarified\ jargon * happy\ face + \beta_6 clarified\ jargon * voice\ arousal + \beta_7 happy\ face * voice\ arousal + Xb + \varepsilon$$

--Insert Table 8, 9, 10, and 11 here—

Power. We also did a power calculation with a medium effect size to ensure that our sample size met the minimum requirements. Our model has three independent variables, four interactions, and six control variables. According to G Power, the estimated sample size is one hundred and eighty-nine. Our six-hundred sample size far exceeds the requirements.

Results. Table 10 shows the descriptive statistics. Before running the regression analysis, we checked the multicollinearity issue (Franke, 2010) by using VIF statistics. Results show that no multicollinearity issue existed, since the VIF of the three IVs are all below 1.4 (VIF of jargon = 1.24; VIF of voice arousal = 1.04, VIF of happy face = 1.02). Other control variables also showed no problems (VIF of ranking = 1.08, VIF of assistant = 1.10, VIF of product origin = 1.10; VIF of gender = 1.24; VIF of price = 1.21; VIF of duration = 1.32).

Full model. We put clarified jargon, voice arousal, and happy face with interaction terms together to run the full model with control variables. In Table 11, the clarified jargon (b = 1.06, t(13,586) = 2.44, p=.015) and voice arousal (b = -.22, t(13,586) = 2.06, p = .01) are significant on sales performance in the expected direction; this supports H6 and H7. Happy facial expression (b = -.82, t(13,586) = -3.14, p=.002) negatively affects sales performance, which is opposite to the

hypothesized direction. Therefore H8 is not supported. The interaction of clarified jargon and voice arousal (b = -.51, t(13,586) = -2.95, p=.003) is negatively significant, which supports H9. Lastly, the interaction of voice arousal and happy facial expression (b = .88, t(13,586) = 3.05, p=.002) is positively significant, and therefore support H11. The model R square is 0.32,

Robustness check. First, we analyzed the curvilinear relationships between independent and dependent variables with all control variables. This ensures that linear is the most logical relationship between independent and dependent variables. We ran the quadratic regression by looking at the independent variables individually; only the quadratic term of the happy face showed a significant result (b=-.065, t=-3.32, p=.001). The quadratic terms of clarified jargon (b=.77, t=.29, p>.05) and voice arousal (b=.04, t=.77, p>.05) remain non-significant. When we added three variables to the model simultaneously with the same dependent variable and control variables, the quadratic term of happy face still showed significant results (b=-.068, t=-3.40, p=.001), while the clarified jargon (b=.023, t=.23, p>.05) and voice arousal (b=-.21, t=-.80, p>.05) remain non-significant. The model R square is 0.28. In the range of 0 to 1 of the happy face, the overall tendency of both linear and quadratic are negative (as indicated in the result), and the vertex did not exist in this range. We excluded the non-linear relationships between independent variables and dependent variables. The later analysis focuses on linear relationships only.

Our second robustness check aims to assess the generalizability of our results and suggests whether the influencer's ranking moderates the effect of the IVs on the DV. Here is the question: Will bottom-tier influencers employing the same strategy generate the same effect as top influencers? This post-hoc analysis can give us an insight into what influencers can expect as they keep growing their follower numbers; this increases the robustness of our results. To

facilitate the analysis, we treat influencer ranking as the new moderator. The independent variables, dependent variables, and control variables remain the same. Thus, the new model is:

$$Ln(Sales) = \beta_0 + \beta_1 clarified\ jargon + \beta_2 voice\ arousal + \beta_3 happy\ face + \beta_4 ranking + \beta_5 clarified\ jargon * happy\ face * voice\ arousal * ranking + \beta_6 clarified\ jargon * happy\ face * voice\ arousal + \beta_7 clarified\ jargon * happy\ face * ranking + \beta_8 clarified\ jargon * voice\ arousal * ranking + \beta_9 happy\ face * voice\ arousal * ranking + \beta_{10} clarified\ jargon * ranking + \beta_{11} ranking * voice\ arousal + \beta_{12} happy\ face * ranking + \beta_{13} clarified\ jargon * happy\ face + \beta_{14} clarified\ jargon * voice\ arousal + \beta_{15} happy\ face * voice\ arousal + Xb' + \varepsilon$$

We did not find a significant moderation effect of ranking on any independent variables and interactions (ranking x jargon x voice x face: b = .01, t = 1.05, p = .29; ranking x jargon x voice: b = -.001, t = -.89, p = .37; ranking x jargon x face: b = -.02, t = -1.09, p = .27; ranking x voice x face: b = .003, t = .79, p = .43; ranking x jargon: b = .003, t = 1.26, p = .21; ranking x voice: b = .00, t = -.09, p = .92; ranking x face: b = -.007, t = -.72, p = .47). This illustrates that our findings are robust. The effect of our independent variables on sales performance does not depend on influencer ranking.

We did a second series of robustness checks using a multi-level approach, which treats the influencer as the higher-level variable and allows the randomness of both video-level and influencer-level variables. In this case, we can investigate whether influencers employing our strategy could achieve better sales performance.

We first run the empty model with influencers as the class. The results shows significant different on sales across influencer (b=10.84, t=70.81, p<.001). Next, we treat three independent

variables (clarified jargon, voice arousal, happy face) as the level 1 predictor and influencer ranking as the level 2 predictor. The intercepts and slopes in both level 1 and level 2 equations are allowed to vary randomly across groups. The result only shows the significant positive effect of clarified jargon (H6) (b=.28, t=2.67, p=.01). The voice arousal (H7) (b=-.08, t=-.76, p=.45) and happy face (H8) (b=-0.03, t=-.65, p=.51) are non-significant. Meanwhile, the influencer ranking was significant (b=-.00, b=-2.82, p=.007). See Figure 7 for the result. These results demonstrated that influencers who adopt clarified jargon more frequently have better sales performance.

$Level$ 1:Sales=β_0i+β_1i×$jargon$+β_2j $arousal$+β_3j $happy$+ε_ij

$Level$ 2: β_0i=γ_00+γ_01×ranking+μ_oj

β_1i=γ_10+μ_1j

β_2i=γ_20+μ_2j

β_3i=γ_30+μ_3j

--Insert Figure 7 here--

When adding interaction terms of clarified jargon and voice arousal, voice arousal and happy face, and happy face and clarified jargon and the three-way of clarified jargon, voice arousal and happy face, the clarified jargon has marginal significant effect on sales performance (b=.85, t=1.78, p=.075). However, the effect of voice arousal (H2) (b=-.06, b=-.47, p=.64) and happy face (H3) (b=-.40, t=-1.23, p=.22) were still non-significant, so as all the interaction terms (jargon x voice arousal: b=-.29, t=-1.53, p=.13; jargon x happy face: b=2.46, t=.79, p=.43, voice arousal x happy voice: b=.41, t=1.41, p=.25). See Figure 8 for the result.

--Insert Figure 8 here--

Lastly, we tried to see if influencer ranking moderate the effect IVs to DVs. As expected, the interaction term between influencer ranking and clarified jargon is non-significant (b=.00, t=.46, p=.65), so as the interaction between influencer ranking and voice arousal (b=.00, t=.09, p=.93) and between influencer ranking and happy face (b=.00, t=.60, p=.55). See Figure 9 for the result.

--Insert Figure 9 here--

Discussion. Our main study looked into the impact of each variable and their interaction. We found a positive effect of clarified jargon (H6), which boosts influencers' credibility and expertise (Atkinson & Carskaddon, 1975), thereby increasing positive emotions and buying intentions (Sharma, 1990; Lee & Dubinski, 2003).

We found a negative effect of voice arousal (H7). This finding contrasts with the literature, which suggests that voice arousal can increase perceptions of passion and a higher acceptance rate of competence (Allison et al., 2022; Chebat et al., 2007; Oksenberg et al., 1986). However, in the context of Chinese livestream selling, consumers prefer a more peaceful “no yelling,” “no vulgar words,” and a slow environment (Zhang, 2022; Li, 2023). Existing research shows that an increase in cognitive and physiological arousal increases anxiety (Noteboom et al., 2001). However, anxiety is often experienced by residents in China, so it makes sense that Chinese consumers would not enjoy a high-arousal livestream, which could further increase their anxiety levels.

As for H8, we found that a happy face has a negative effect on sales performance. This counter-intuitive finding could be caused by the high arousal that a happy face brings to the audience. According to the literature, happiness belongs to the high-arousal positive emotions (Sato & Yoshikawa, 2007; Bhimavarapu et al., 2020). Therefore, the happy emotions brought by the happy faces of the influencer could increase consumers’ arousal and further increase anxiety

(Noteboom et al., 2001). Another explanation brought by Kim (2022) is that exposing to happiness can stop consumers from consuming happiness-claimed products because of a decrease in perceived happiness potential that consumers attribute to the products. Even though not all the products were hedonic, the happy influencer could increase the positive expectation of consumers on the products. As a consequence, consumers may experience a decrease in perceived positive emotions when receiving and using the product.

H9 predicts that the interaction of clarified jargon and voice arousal will be negative, and this was supported. When voice arousal is low, the positive effect of clarified jargon on sales is stronger. When voice arousal is high, the positive effect of clarified jargon is weaker. In general, Douyin's audience does not prefer the noisy live-streaming style and finds the quiet and gentle streaming more appealing (Lu et al., 2023). In this case, the audience will not continue their involvement to see if influencers are good. Instead, they will switch to another live room to find an ideal livestream atmosphere, especially since most audiences regard livestream as a place of emotional dependence (Li, 2023).

Unfortunately, H10 did not receive support that there was no significant interaction between clarified jargon and happy facial expressions. We assume that the negative effect of happy facial expression was not as strong as voice arousal, which could counteract the benefit of clarified jargon.

For H11, we found a significant positive interaction between voice arousal and a happy face. When the happy face is at a high level, there is a positive effect of voice arousal on sales performance. When happy face is at a low level, such positive effect of voice arousal is weaker (the effect of voice arousal on sale performance is more negative). This result supports the H11. Even though we found the negative overall effect of happy face, the interaction term shows that happy face, to some extend, can reverse the negative effect of voice arousal on sales

performance. This could be explained that their combination may provide the emotion consistency to consumers. People who are perceived as emotionally sincere have consistency between feelings and actions (Caza et al. ,2015). Such consistency will make them honest and have the qualities of integrity. Table 12 shows the summary of findings.

The robustness check by two different approach confirmed our believe that influencer ranking does not moderte the effect of IVs on DV, even though it remains significant in its main effect. Besides, we observed that the effect of clarified jargon is more stable and stronger than other effects. Influencers who adopt clarified jargon more frequently have better sales performance. This analysis provides insight into future research that the subject could be video (600 videos in our study), influencers (50 influencers in our study), or even MCN company.

--Insert Table 12 here--

General Discussion

Theoretical Contributions and Implications. In this research, we examined a growing yet surprisingly under-examined livestreaming e-commerce phenomenon in which some influencers are far more successful than others at both selling products in the livestreams and generating online traffic. These influencers transformed from salespeople selling product brands into personal brands reflecting their own with celebrity status and millions of loyal followers. Our major contribution to the prior research on livestreaming is that we shed light on the unique content delivery strategies in influencers' live videos to communicate with an audience and promote products that reflect their unique personal brand styles. We provide a first comprehensive empirical examination that tests the effectiveness of top influencers' strategies in rhetoric, voice, and facial expressions; we also provide guidance for influencers-to-be.

We started out by downloading the actual videos of livestream influencers in the cosmetic sector and extracting the unique content strategies of top influencers (vis-à-vis mid- and bottom-tier influencers) and then conducted an experiment to rigorously test the differential effects of these content strategies on consumer loyalty, purchase intention, and product attitudes. Our lab experiments controlling for the noisy characteristics of the real world replicated the positive effects of the unique configuration of content strategy elements. This gives us confidence in claiming our core contribution to the prior literature that the success of an influencer in the livestreaming category is the outcome of a combination of design elements in their speech and communication, which theoretically and conceptually is far more complex than a single-dimension approach predominantly adopted in the extant work (e.g., Li et al., 2021; Bharadwaj et al., 2022). Our multi-dimension approach leveraging both the real-world video analyses and the experiment video production is notable because it offers stronger ecological validity which gives us faith to claim that the unique content styles of top influencers both in the lab and in the real world translate into positive results for the influencers and product sales alike.

In this emerging research area of livestream selling, we go beyond prior research. Existing research has focused on a specific dimension, such as the audience's attachment to the platforms and influencers (Li et al., 2021) or to specific facial expressions of the online salesperson in the livestream videos (Bharadwaj et al., 2022). The multi-dimensional approach and design allowed us to extract the content and delivery strategies of the winning influencers by anatomizing their unique linguistic styles, voice features, and facial expressions. Novel tools such as voice software and facial expression analysis have enabled us to test the combined effects of these features, which was impossible in the past.

Given that livestream selling is a new and rapidly evolving context, our study examined established theories within this novel setting. Unlike traditional shopping, where consumers have more time to evaluate their purchase and make decisions, livestream audiences must act quickly. Influencers also need to drive purchases efficiently within this limited timeframe. in this fast-paced environment, are credibility, product evaluation, and audience emotion as crucial as in traditional shopping scenarios? Our study 2 tested the construct and confirmed that credibility, product evaluation, and audience emotion are significant in driving purchase. Additionally, our study introduced new mechanisms, such as the reverence for knowledge and authority and the enjoyment of entertaining content, which influence these factors in the context of livestream selling.

Our work also reveals several new insights that challenged past literature. Table 13 summarizes the result of each hypothesis, and Table 14 contrasts our findings with the conclusions in past literature. Old literature studied Western participants, who have different mental health conditions than Chinese participants. Our study demonstrated the negative effect of voice arousal due to the preference of Chinese consumers and the negative consequences (e.g., anxiety) of experiencing a high arousal environment, contradicting the past literature that supported the positive effect of high-arousal voice (Allison et al., 2022; Chebat et al., 2007; Oksenberg et al., 1986). Similarly, we also found a negative effect of a happy face on sales performance, challenging the old findings of the positive effect of a happy face (Ilicic & Baxter, 2018; Reed et al., 2018; Van Kleef et al., 2015). That finding also reflects the Chinese consumers' mental state that high anxiety people do not prefer high-arousal environments.

--Insert Table 13 and Table 14 here--

Furthermore, our research recognizes the pivotal role of multi-modality in the realms of persuasion and communication. Our study examines the interactive effect between content strategy (i.e., clarified jargon) and emotion strategies that encompass voice and facial expression. Through our rigorous analysis, we discovered a notable augmentation in sales performance due to the employment of both voice and facial expression. For instance, our investigation revealed the positive impact of employing clarified jargon in isolation but, conversely, suggested a weaker effect when coupled with a high-arousal voice. In a similar vein, we observed that when influencers exhibited expressions of happiness infrequently, the negative effect of voice arousal on sales was weaker. In contrast, frequent displays of such expressions amplified the negative influence of voice arousal on sales. These empirical revelations provide valuable insights for the domain of livestream commerce, and highlight the need for future research to dedicate greater attention to how strategies within a singular dimension can significantly enhance overall performance.

We collected six hundred live videos from fifty influencers on Douyin, enabling us to enhance the generalization of our findings and help compare the strength of how each variable affected sales performance. We defined unique jargon usage, and the clarified jargon that was rarely studied in the past. This study also contributes to the boundary condition investigation. For example, we found significant interactions between clarified jargon and voice arousal, and voice arousal and happy face. These findings demonstrated the dynamic characteristics of livestream selling and the importance of considering the configuration in our work. Our study also included a robustness check to investigate if influencer ranking moderates the effect of IVs on sales performance. We did not find a significant moderation effect of influencer ranking on any

independent variables and interactions, illustrating that all influencers should receive equal benefits by using these strategies.

Managerial Implications. Influencers, products, and brands have increasingly adopted new strategies as the world rapidly moves to the digitalized domain. More than ever, product and service brands need powerful and strong personal brands (Smith & Fischer, 2021) to advocate for and promote their existence and sales. Our research is very timely as this emerging format (livestream sales) has successfully kicked off in Asia, which can spill over to, and leveraged by, other e-commerce platforms such as Amazon. One cannot ignore the explosive growth of the multi-channel network (MCN) companies in China like Dayu Media, Mei One, and Hive Media (Li, 2022). The knowledge of how to build top influencers in a category is increasingly important for the platforms and for individual influencers, and our specific configuration of optimal content and delivery strategies has the potential to equip brands, platforms, and influencers with practical tools to be successful.

Our findings are instrumental for these companies and influencers-to-be as they hone their skills in livestream communication. Our findings provide the do's and don'ts in influencer training (see Table 15 for the contrasting of top and bottom influencers), with the hope that the knowledge can empower these influencers' personal brand-building. After all, being a social media influencer has become a viable career option for many ambitious young people (Appel et al., 2020; Suciu, 2020).

--Insert Table 15 here--

Limitation and future research. We did not find significant results in the three-way interaction (study 3). We did not propose this as a formal hypothesis because it is a complex configuration. In reality, livestreaming is more complicated and has more dimensions than we could analyze

using our methodology. It is almost impossible to cover all aspects and generate a perfect guidebook. The non-significant finding uncovered the dynamic and unpredictable part of the live stream.

Additionally, another unique feature of livestream selling platforms is that audience can interact with others in real time, which can strongly contribute to person brand building of the influencers and the success of product sales. However, due to our limited technical resources, we are unable to scrutinize the dynamics of audience and influencer interactions and their subsequent impact on influencer brand building and product sales. Given that prior work has shown the positive effect of social presence on persuasion (Gefen & Straub, 2004), future work should test whether the way influencers respond and interact with the audience can affect the felt connections and engagement of the audience in the livestream sessions, which can subsequently impact the influencer's person brand building and product sales.

Lastly, our findings are specific to Chinese live-stream selling, involving Chinese influencers and audiences (study). Extending our results to other contexts (e.g., non-Chinese languages) is beyond the scope of this dissertation, as live-streaming is still nascent in foreign countries, and their consumers may have different preferences from Chinese consumers. It is impossible to collect hundreds of livestream videos in other countries to run the analysis. An unexplored but potentially significant avenue for future research is whether strategies employed by Chinese influencers can serve as foundational learning tools for foreign influencers, who may eventually develop their own unique approaches.

Reference

Aaker, D. A., & Myers, J. G. (1975). Advertising Management: By David A. Aaker, John G. Myers. Prentice-Hall.

Abbey, J. D., & Meloy, M. G. (2017). Attention by design: Using attention checks to detect inattentive respondents and improve data quality. *Journal of Operations Management, 53*, 63-70.

Allison, T. H., Warnick, B. J., Davis, B. C., & Cardon, M. S. (2022). Can you hear me now? Engendering passion and preparedness perceptions with vocal expressions in crowdfunding pitches. *Journal of Business Venturing*, 37(3), 106193.

Atkinson, D. & Carskaddon, G. (1975). A prestigious introduction, psychological jargon, and perceived counselor credibility. *Journal of counseling psychology*, 22(3), 180

Appel, G., Grewal, L., Hadi, R., & Stephen, A. T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing Science*, 48(1), 79-95.

Appiah, O. (2006). Rich Media, poor Media: The Impact of audio/video vs. text/picture testimonial ads on browsers' evaluations of commercial web sites and online products. *Journal of Current Issues & Research in Advertising*, 28(1), 73-86.

Bai, S., Chang, Q., Yao, D., Zhang, Y., Wu, B., & Zhao, Y. (2021). Anxiety in residents in China: prevalence and risk factors in a multicenter study. *Academic Medicine*, 96(5), 718-727.

Bergkvist, L. (2009). The role of confidence in attitude–intention and beliefs–attitude relationships. *International Journal of Advertising*, *28*(5), 863-880.

Bharadwaj, N., Ballings, M., Naik, P. A., Moore, M., & Arat, M. M. (2022). A new livestream retail analytics framework to assess the sales impact of emotional displays. *Journal of Marketing*, 86(1), 27-47.

Bhimavarapu, J. P., Kalyan, S. S. S., & Mittal, V. K. (2020, February). Discriminating high arousal and low arousal emotional speech using mahalanobis distance among acoustic features. *In 2020 National Conference on Communications (NCC) (pp. 1-6). IEEE*.

Britt, R., Hayes, J., Britt, B., & Park, H. (2020). Too big to sell? A computational analysis of network and content characteristics among mega and micro beauty and fashion social media influencers. *Journal of Interactive Advertising*, *20*(2), 111-118.

Brown, Z. C., Anicich, E. M., & Galinsky, A. D. (2020). Compensatory conspicuous communication: Low status increases jargon use. *Organizational Behavior and Human Decision Processes*, *161*, 274-290.

Castro, C. M., Wilson, C., Wang, F., & Schillinger, D. (2007). Babel babble: physicians' use of unclarified medical jargon with patients. *American journal of health behavior*, 31(1), S85-S95.

Cavazza, N., & Gabrielli, V. (2015). Affordant shapes of product holder influence product evaluation and purchase intention. *Current Psychology*, *34*, 447-465.

Caza, A., Zhang, G., Wang, L., & Bai, Y. (2015). How do you really feel? Effect of leaders' perceived emotional sincerity on followers' trust. *The Leadership Quarterly*, *26*(4), 518-531.

Chen, C. C., & Lin, Y. C. (2018). What drives live-stream usage intention? The perspectives of flow, entertainment, social interaction, and endorsement. *Telematics and Informatics*, 35(1), 293-303.

Chao, R. K. (1994). Beyond parental control and authoritarian parenting style: Understanding Chinese parenting through the cultural notion of training. *Child development*, 65(4), 1111-1119.

Chen, Y., & Wyer Jr, R. S. (2020). The effects of endorsers' facial expressions on status perceptions and purchase intentions. *International Journal of Research in Marketing*, 37(2), 371-385.

Cialdini, R. (2001). Principles of persuasion. *Arizona State University, eBrand Media Publication*.

Dolan, G. (2019). The hidden consequences of using jargon. *Governance Directions*, 71(4), 212-214.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt brace Jovanovich college publishers.

Elkin, Tobi (2002 August 6), Automakers biggest users of online rich media ads. [Online] Available: <http: / /www.adage.com/ news.cms?newsId=35670>.

Franke, G. R. (2010). Multicollinearity. *Wiley international encyclopedia of marketing*.

Freeman, K. S., & Spyridakis, J. H. (2004). An examination of factors that affect the credibility of online health information. *Technical communication*, *51*(2), 239-263.

Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega*, 32(6), 407-424.

Guyer, J. J., Briñol, P., Petty, R. E., & Horcajo, J. (2019). Nonverbal behavior of persuasive sources: A multiple process analysis. *Journal of Nonverbal Behavior*, 43, 203-231.

Haimson, O. L., & Tang, J. C. (2017, May). What makes live events engaging on Facebook Live, Periscope, and Snapchat. *In Proceedings of the 2017 CHI conference on human factors in computing systems (pp. 48-60).*

Herabadi, A. G., Verplanken, B., & Van Knippenberg, A. (2009). Consumption experience of impulse buying in Indonesia: Emotional arousal and hedonistic considerations. *Asian Journal of Social Psychology, 12*(1), 20-31.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public opinion quarterly, 15*(4), 635-650.

Hsieh, J. K., Hsieh, Y. C., & Tang, Y. C. (2012). Exploring the disseminating behaviors of eWOM marketing: persuasion in online video. *Electronic Commerce Research*, 12, 201-224.

Hu, A., & Ma, S. (2021). *Persuading investors: A video-based study* (No. w29048). *National Bureau of Economic Research.*

Hu, M., Zhang, M., & Wang, Y. (2017). Why do audiences choose to keep watching on live video streaming platforms? An explanation of dual identification framework. *Computers in Human Behavior*, 75, 594-606.

Ilicic, J., Kulczynski, A., & Baxter, S. M. (2018). How a smile can make a difference: Enhancing the persuasive appeal of celebrity endorsers: Boosting consumer perceptions of celebrity genuineness through the use of a "Duchenne Smile" in advertising. *Journal of Advertising Research*, 58(1), 51-64.

Kim, A. (2022). The paradox in happiness sales: How can happiness primes backfire?. *Journal of Business Research, 146,* 540-552.

Kim, D. Y., & Kim, H. Y. (2021). Trust me, trust me not: A nuanced view of influencer marketing on social media. *Journal of Business Research, 134,* 223-232.

Kim, J., & Lennon, S. J. (2013). Effects of reputation and website quality on online consumers' emotion, perceived risk and purchase intention: Based on the stimulus-organism-response model. *Journal of Research in Interactive Marketing*, *7*(1), 33-56.

Kim, K., Ko, E., & Lee, Y. I. (2012). Art infusion in fashion product: The influence of visual art on product evaluation and purchase intention of consumers. *Journal of Global Fashion Marketing*, *3*(4), 180-186.

Knapp, M. L., Hall, J. A., & Horgan, T. G. (2013). Nonverbal communication in human interaction. *Cengage Learning.*

Laroche, M., Kim, C., & Zhou, L. (1996). Brand familiarity and confidence as determinants of purchase intention: An empirical test in a multiple brand context. *Journal of business Research*, *37*(2), 115-120.

Laukka, P., Juslin, P., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, *19*(5), 633-653.

Lee, J., & Shin, S. Y. (2022). Something that they never said: Multimodal disinformation and source vividness in understanding the power of AI-enabled deepfake news. *Media Psychology*, 25(4), 531-546.

Lee, S., & Dubinsky, A. (2003). Influence of salesperson characteristics and customer emotion on retail dyadic relationships. *The International Review of Retail, Distribution and Consumer Research*, 13(1), 21-36.

Lee, M. T., & Theokary, C. (2021). The superstar social media influencer: Exploiting linguistic style and emotional contagion over content?. *Journal of Business Research*, 132, 860-871.

Lewinski, P., Fransen, M. L., & Tan, E. S. (2014). Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli. *Journal of Neuroscience, Psychology, and Economics*, *7*(1), 1.

Li, E. (2022, May 8). 4 MCNs You Need to Know in China. Jing Daily. Retrieved from https://jingdaily.com/4-mcn-china-influencer/

Li, J. (2022). The hot noise of open-air cinema. In *Sensing China* (pp. 199-228). Routledge.

which means they love Sometimes they also said “approaching/gathering hot noise

Li, Y., Li, X., & Cai, J. (2021). How attachment affects user stickiness on live streaming platforms: A socio-technical approach perspective. *Journal of Retailing and Consumer Services*, 60, 102478.

Li, Y., & Xie, Y. (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.

Lin, Y., Yao, D., & Chen, X. (2021). Happiness begets money: emotion and engagement in live streaming. *Journal of Marketing Research*, 58(3), 417-438.

Lu, J., Allwood, J., & Ahlsén, E. (2011). A study on cultural variations of smile based on empirical recordings of Chinese and Swedish first encounters. *jia*, *46*(31), 7862791.

Lu, Q., Liu, C., & Chen, Y. (2023). A qualitative study of how consumers’ purchase intention is influenced by the livestreaming shopping streamer under negative incidents caused by streamers.

Luo, H., Cheng, S., Zhou, W., Yu, S., & Lin, X. (2021). A study on the impact of linguistic persuasive styles on the sales volume of live streaming products in social e-commerce environment. *Mathematics*, 9(13), 1576.

Macintosh, G., & Lockshin, L. S. (1997). Retail relationships and store loyalty: a multi-level perspective. *International Journal of Research in marketing*, *14*(5), 487-497.

Mayew, William J., and Mohan Venkatachalam. (2012) The power of voice: Managerial affective states and future firm performance. *Journal of Finance*, 67(1), 1-43.

Nafees, L., Cook, C. M., Nikolov, A. N., & Stoddard, J. E. (2021). Can social media influencer (SMI) power influence consumer brand attitudes? The mediating role of perceived SMI credibility. *Digital Business*, 1(2), 100008.

Nasermoadeli, A., Ling, K. C., & Maghnati, F. (2013). Evaluating the impacts of customer experience on purchase intention. *International Journal of business and management*, *8*(6), 128.

Nesler, M. S., Aguinis, H., Quigley, B. M., & Tedeschi, J. T. (1993). The Effect of Credibility on Perceived Power 1. *Journal of Applied Social Psychology*, *23*(17), 1407-1425.

Noteboom, J. T., Barnholt, K. R., & Enoka, R. M. (2001). Activation of the arousal response and impairment of performance increase with anxiety and stressor intensity. *Journal of applied physiology.*

Paing, K. S. (2019). The effect of brand experience on brand love and brand loyalty of Erke sportsweak users in Yangon (Doctoral dissertation, MERAL Portal).

Park, J., Lee, J. M., Xiong, V. Y., Septianto, F., & Seo, Y. (2021). David and Goliath: when and why micro-influencers are more persuasive than mega-influencers. *Journal of Advertising*, *50*(5), 584-602.

Park, S., Shim, H. S., Chatterjee, M., Sagae, K., & Morency, L. P. (2014, November). Computational analysis of persuasiveness in social multimedia: A novel dataset and

multimodal prediction approach. *In Proceedings of the 16th International Conference on Multimodal Interaction (pp. 50-57).*

Rajagopal, P., & Montgomery, N. V. (2011). I imagine, I experience, I like: The false experience effect. *Journal of Consumer Research, 38*(3), 578-594.

Rajaram, P., & Manchanda, P. (2020). Video influencers: Unboxing the mystique. arXiv preprint arXiv:2012.12311.

Reed, L. I., Stratton, R., & Rambeas, J. D. (2018). Face value and cheap talk: How smiles can increase or decrease the credibility of our words. *Evolutionary Psychology*, 16(4), 1474704918814400.

Rychlowska, M., Jack, R. E., Garrod, O. G., Schyns, P. G., Martin, J. D., & Niedenthal, P. M. (2017). Functional smiles: Tools for love, sympathy, and war. *Psychological science,* 28(9), 1259-1270.

Saima, & Khan, M. A. (2020). Effect of social media influencer marketing on consumers' purchase intention and the mediating role of credibility. *Journal of Promotion Management*, *27*(4), 503-523.

Sato, W., & Yoshikawa, S. (2007). Enhanced experience of emotional arousal in response to dynamic facial expressions. *Journal of Nonverbal Behavior*, *31*, 119-135.

Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychological bulletin*, 99(2), 143.

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, *8*(4), 379-386.

Schore, A. N. (2002). Advances in neuropsychoanalysis, attachment theory, and trauma research: Implications for self psychology. *Psychoanalytic inquiry*, 22(3), 433-484.

Sharma, A. (1990). The persuasive effect of salesperson credibility: conceptual and empirical examination. *Journal of Personal Selling & Sales Management,* 10(4), 71-80.

Smith, A. N., & Fischer, E. (2021). Pay attention, please! Person brand building in organized online attention economies. *Journal of the Academy of Marketing Science*, 49(2), 258-279.

Suciu, P. (2020). Is Being A Social Media Influencer A Real Career? Forbes.

Tan, H. T., Wang, E. Y., & Yoo, G. S. (2019). Who likes jargon? The joint effect of jargon type and industry knowledge on investors' judgments. *Journal of Accounting and Economics,* 67(2-3), 416-437.

Tellis, G. J., MacInnis, D. J., Tirunillai, S., & Zhang, Y. (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.

Ordenes, F., Grewal, D., Ludwig, S., Ruyter, K., Mahr, D., & Wetzels, M. (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.

Peace, V., Miles, L., & Johnston, L. (2006). It doesn't matter what you wear: The impact of posed and genuine expressions of happiness on product evaluation. *Social Cognition*, *24*(2), 137-168.

Van Kleef, G. A., Van den Berg, H., & Heerdink, M. W. (2015). The persuasive power of emotions: Effects of emotional expressions on attitude formation and change. *Journal of Applied Psychology*, 100(4), 1124.

Wang, L., & Fan, Y. (2021, October 18). China's Biggest Influencer Pushes a New Message: Buy Chinese. Sixthtone. Retrieved from https://www.sixthtone.com/news/1008694/chinas-biggest-influencer-pushes-a-new-message-buy-chinese

Wang, S. W., Kao, G. H. Y., & Ngamsiriudom, W. (2017). Consumers' attitude of endorser credibility, brand and intention with respect to celebrity endorsement of the airline sector. *Journal of Air Transport Management*, *60*, 10-17.

Wang, X., Lu, S., Li, X. I., Khamitov, M., & Bendle, N. (2021). Audio mining: the role of vocal tone in persuasion. *Journal of Consumer Research*, 48(2), 189-211.

Weismueller, J., Harrigan, P., Wang, S., & Soutar, G. N. (2020). Influencer endorsements: How advertising disclosure and source credibility affect consumer purchase intention on social media. *Australasian marketing journal*, 28(4), 160-170.

Wohn, D., Freeman, G., & McLaughlin, C. (2018). Explaining viewers' emotional, instrumental, and financial support provision for live streamers. *In Proceedings of the 2018 CHI conference on human factors in computing systems (pp. 1-13).*

Woodward, S. F., Bari, S., Vike, N., Lalvani, S., Stetsiv, K., Kim, B. W., ... & Katsaggelos, A. K. (2022). Anxiety, post–COVID-19 syndrome-related depression, and suicidal thoughts and behaviors in COVID-19 survivors: cross-sectional study. *JMIR Formative Research*, *6*(10), e36656.

Wu, J., Zhu, N., & Dai, Q. (2010). Consumer ethnocentrism, product attitudes and purchase intentions of domestic products in China. EBM, Chengdu, China, 2262, 2265.

Xin, L., & Seo, S. (2020). The role of consumer ethnocentrism, country image, and subjective knowledge in predicting intention to purchase imported functional foods. *British Food Journal*, 122(2), 448-464.

Xu, Y., Farver, J. A., Zhang, Z., Zeng, Q., Yu, L., & Cai, B. (2005). Mainland Chinese parenting styles and parent-child interaction. *International Journal of Behavioral Development*, 29(6), 524-531.

Yang, J., Zhang, J., & Zhang, Y. (2021). First law of motion: Influencer video advertising on TikTok. Available at SSRN 3815124.

Yue-Hei Ng, J., Hausknecht, M., Vijayanarasimhan, S., Vinyals, O., Monga, R., & Toderici, G. (2015). Beyond short snippets: Deep networks for video classification. *Proceedings of the IEEE conference on computer vision and pattern recognition.*

Zhou, S., Li, T., Yang, S., & Chen, Y. (2022). What drives consumers' purchase intention of online paid knowledge? A stimulus-organism-response perspective. *Electronic Commerce Research and Applications, 52*, 101126.

Zimmermann, M., & Jucks, R. (2018). How experts' use of medical technical jargon in different types of online health forums affects perceived information credibility: Randomized experiment with laypersons. *Journal of medical Internet research*, 20(1), e30.

Zoghaib, A. (2019). Persuasion of voices: The effects of a speaker's voice characteristics and gender on consumers' responses. *Recherche et Applications en Marketing* (English Edition), 34(3), 83-110.

2021 Douyin E-commerce ecological development report (2021).

李雅莉. (2023). 抖音电商主播的直播话术分析. Modern Linguistics, 11, 5018.

张娜娜. (2022). 基于直播电商产业链的东方甄选带货现象研究. E-Commerce Letters, 11, 63.

2017“新一线”城市排行榜发布成都、杭州、武汉蝉联三甲郑州、东莞新晋入榜_第一财经.

www.yicai.com. Retrieved 2018-04-01.

于珊. (n.d.). 第七次全国人口普查公报（第六号）_部门政务_中国政府网.
https://www.gov.cn/xinwen/2021-05/11/content_5605789.htmTables and Figures

Appendix A

Three questions captured purchase intention (Kim & Kim, 2021; α=.93): “I am likely to purchase this lotion,” “I am inclined to purchase this lotion,” and “I am willing to purchase this lotion”. The authenticity and credibility scale (Kim & Kim, 2021; Freeman & Spyridakis, 2004) included the evaluation of both the influencer and the product information offered by the influencer (e.g., “This influencer is reliable,” and “The product information given by this influencer is reliable”; α=.96). The product attitude questions were about how desirable, interesting, and favorable the lotion is (Kim & Kim, 2021; α=.95). Lastly, the audience emotion was measured by five questions (Paing, 2019; α=.97): “I have strong feelings about this live video,” “When watching this live video, I experience many emotions,” “This live video triggers positive emotions,” “This live video appeals positively to me on an emotional level,” and “This live video appeals positively to my feelings and sentiment.” To suit male participants, some scales were adapted. For example, each question of purchase intention for male participants (Kim & Kim, 2021) ended with “as a birthday gift”.

Tables and figures.

Table 1. Six influencers' descriptive information in terms of engagement and sales

Ranking	3	42	172	222	359	496
Name	Jiaqi Li	Xiaoyue	Han Bingni	Jiuer	shiyang	Wan Xin
Huitun score	70,289	3,838	1,871	1,600	1,182	924
Sector	cosmetic	cosmetic	cosmetic	cosmetic	cosmetic	cosmetic
Level	expert	expert	expert	expert	expert	expert
Fans	27,458,000	561,000	1,005,000	302,000	518,000	226,000
Likes	729,000	12,802	1,506	2,962	4,136	832
Comments	2,196,251	713,817	24,976	91,167	676,109	69,607
Sessions/Month	24	49	8	18	29	30
Audience online	272,918	354	201	231	1,575	702
Sales in RMB/Month	8,355,441	110,374	21,859	732,375	70,532	110,382
Views/Session	14,533,095	60,142	278,971	81,222	97,478	50,030
Likes/Session	24,180,243	739,569	1,144,058	460,548	79,958	45,701
Comments/Session	39,219	6,864	1,189	1,572	10,732	1,200

Table 2. Word count and video length of randomly selected influencers

Influencer	Product	Length in minute	date	Word count	Word count total	Length in total (in minutes)
Li Jiaqi (Top)	3CE matte lipstick	14.25	2020.08.09	4224	12,625	39.11
	MT anti-aging serum	6.73	2020.08.09	2289		
	Hanshu Gold Mask	14.05	2020.08.09	4773		
	Huaxizi makeup remover	4.08	2020.08.10	1339		
Xiao Yue (Top)	#333 matte lipstick	8.05	2020.11.02	2491	8,471	25.10
	Lantong serum	7.55	2020.08.10	2682		
	Anti-aging mask	6.62	2020.08.10	2496		
	MJC makeup remover	2.88	2020.10.03	802		
Han Bingni (Mid)	Kating lipstick	6.62	2020.11.29	1896	10,233	35.79
	Vivian serum	4.57	2020.05.03	1266		
	Xiumeile Mask	17.25	2020.08.09	5083		
	Unny makeup remover	7.35	2020.04.18	1988		
Jiuer (Mid)	Sivanna matte lipstick	29.68	2020.08.08	7406	20,359	68.70
	Dr.Brag serum	8.55	2020.08.08	2671		
	KACA mask	25.97	2020.08.08	8644		
	Red Earth makeup remover	4.5	2020.08.08	1638		

Table 2. Word count and video length of randomly selected influencers (continued).

Influencer	Product	Length in minute	Date	Word count	Word count total	Length in total (minutes)
Wan Xin (Bottom)	3CE matte lipstick	7.65	2020.11.15	2142	9,077	31.78
	Shuyu serum	4.28	2020.08.11	1298		
	Syrene mask	13.8	2020.11.23	3999		
	Rose makeup remover oil	6.05	2020.08.11	1638		
Shi Yang (Bottom)	Mikeer matte lipstick	4.083	2020.11.04	717	3,374	19.74
	Salmon serum	1.22	2020.08.22	327		
	Gaozi mask	8.42	2020.11.10	653		
	Jojoba oil makeup remover	6.02	2020.08.22	1677		

Table 3. Audio segments and arousal scores

	Segment #s	Voice arousal mean
Top	2,232	19,27
Mid	2,269	17.92
Bot	1,540	16.38
Total	6040	

Table 4. Image segments and the scores of happy and smile indicators

	Top	Mid	Bottom	Total
Valid image #	3,395	5,474	2,922	11,791
Happy mean	0.02746552	0.18372732	0.09456843	

Table 5. Stimuli of experiment – clarified jargon

Content	Top influencer script	Mid tier influencer script	Bottom tier influencer script
Jargon	The **gold foil (jargon)** in it has the purity of 98.1%. **Gold foil can help you release negative ions. It can speed up water circulation in our skin and enhance the absorption of our skin.** *(Clarified Explanation)*	The **gold foil (jargon)** in it has the purity of 98.1%.	The ingredients can release negative ions, which speed up water circulation in our skin and enhance the absorption of our skin.

Table 6. Manipulation checks of text, audio, and facial expression in S2A videos

	Top Influencer Video	Mid Influencer Video	Bottom Influencer Video
Jargon type	Clarified	Unclarified	None
Jargon frequency	8	5	0
CA_Voice arousal (mean)	11.9469 (high)	7.8911 (mid)	7.2837 (low)
CA_Happy (mean)	0.3279 (mid)	0.3541 (high)	0.0911 (low)
CHN_Voice arousal (mean)	13.3373 (high)	39.2188 (high)	0.1708 (high)
CHN_Happy (mean)	11.5705 (mid)	36.3206 (mid)	0.0609 (low)

Table 7. Descriptive information on 50 influencers

Ranking	Influencer	Gender
7	董先生	Male
9	小小 101	Female
10	隋心	Female
17	刘媛媛	Female
22	贾乃亮	Male
24	丫头 baby	Female
26	邱莹莹	Female
31	澳门小源	Female
34	孙铱	Female
46	小风子老板娘（放假倒计时）	Female
68	舒畅爱美丽	Female
71	王红权星	Male
75	小熊出没	Female
102	小贝饿了	Female
109	是可乐	Female
118	主持人王芳	Female
140	黄圣依	Female
142	罗凤魁	Male

Table 7. Descriptive information on 50 influencers (continued)

Ranking	Influencer	Gender
164	子安	Female
180	大奔姐	Female
184	小小莎老师	Female
199	深夜徐老师	Female
205	晴子-	Female
209	黄宥明	Male
230	马帅归来	Male
234	4 火姐姐	Female
272	曹征 Alvin	Male
290	小燕 517 美食派	Female
297	Christy 钟丽缇	Female
299	笑笑姐	Female
307	张喜喜	Female
318	一喜推荐官—每早七点直播	Female
327	梁昕	Female
329	希姐传媒（今晚 9 点直播）	Female
338	田恬 TTyilia	Female
345	王小川	Male

Table 7. Descriptive information on 50 influencers (continued)

Ranking	Influencer	Gender
371	菩小提	Female
417	刘天佐	Male
432	曹颖	Female
433	方琼	Female
480	密子君	Female
486	千颂妮	Female
489	吴小白的游乐园	Male

Table 8. The operationalization of independent variables

IV, DV	Operationalization
Sales in each video (DV)	in ln format in analysis
Clarified jargon (IV)	Frequency, human coding by two coders
Voice arousal (IV)	Scale from 1 to 30. A higher number means a higher degree of …
Happy face (IV)	Scale from 0 to 1. A higher number means a higher degree of …

Table 9. The operationalization of control variables

Control variable	Operationalization
Ln_Duration	In seconds
Ln_Price	The final price in RMB
Brand origin	Brand's original country, 0-foreign brand, 1-domestic brand
Gender	Influencer gender, 0-female, 1-male
Assistant	0-without assistant, 1-with assistant
Influencer ranking	From 7 to 489

Table 10. Descriptive statistics

	Variable name	**N**	**Min**	**Max**	**Mean**	**SD**
DV	Sale per video (original)	600	10.0	60000.0	1204.5	3071.4
	Sale per video (ln format)	600	6.0	14.9	10.8	1.6
IV	Clarified jargon	600	0.0	3.0	0.4	0.7
IV	Happy face	600	0.0	24.5	.2	1.0
IV	Voice arousal	600	0	5.45	2.6	1.0
Control	Price per product (ln)	600	0.0	10.0	5.2	1.5
Control	Video duration (ln)	600	2.6	7.0	5.5	0.8
Control	Assistant	600	0.0	1.0	.49	.50
Control	Brand origin	600	0.0	1.0	0.7	0.5
Control	Influencer ranking	600	7	489	199.34	148
Control	Influencer gender	600	0	1	.26	.44

Table 11. result table

Variable	b	t	Sig
(Constant)	7.25	14.35	0.000
(Control)influencer ranking	-.003	-7.94	0.000
(Control)assistant	.09	.75	0.439
(Control)brand origin	.64	4.70	0.000
(Control)gender	.74	5.12	0.000
(Control)ln_price	.29	7.15	0.000
(Control)ln_duration	.43	5.170	0.000
(IV)Clarified jargon (H6)	**1.06**	**2.44**	**0.015**
(IV)Voice arousal (H7)	**-.22**	**-2.60**	**0.10**
(IV)Happy face (H8)	**-.82**	**-3.14**	**0.002**
(Interaction)Jargon*face (H10)	-.86	-.29	0.772
(Interaction)Jargon*voice (H9)	**-.51**	**-2.95**	**0.003**
(interaction)face*voice (H11)	**.88**	**3.05**	**0.002**
face*voice*jargon	1.13	1.08	0.280

Table 12. Summary of findings (significant results)

Modality	Variable	Findings	Hypothesis support
Content (single)	Clarified jargon	Positive	Support for H6
Voice (single)	Voice arousal	Negative	Support for H7
Facial (single)	Happy face	Negative	Opposite to H8
Dual modality	Clarified jargon, voice arousal	When voice arousal is low, the positive effect of clarified jargon on sales is stronger. When voice arousal is high, the positive effect of clarified jargon is weaker.	Support for H9
Dual modality	Voice arousal, happy face	When there is a less happy face, the negative effect of clarified jargon on sales is stronger. When there is a happier face, the negative effect of clarified jargon on sales is weaker.	Support to H11

Table 13. Hypothesis table.

Hypothesis	**Study**	**Result**
H1: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention.	2	Supported
H2: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention because of the perceived higher authenticity and credibility.	2	Supported
H3: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention because they have more positive product evaluations.	2	Supported
H4: Videos with greater usage of clarified jargon, voice arousal and happy face can generate higher purchase intention because audience has more positive emotions.	2	Supported
H5: The main effect of video types on product purchase intention (as in H1) will be mediated by perceived authenticity and credibility, product evaluation, and audience emotions of the influencer.	2	Supported
H6: Clarified jargon has a positive effect on sales performance in livestream selling.	3	Supported
H7: Voice arousal has a negative effect on the sales performance in livestream selling.	3	Supported
H8: Happy facial expressions increase sales performance in livestream selling.	3	Not supported

Table 13. Hypothesis table (continued)

Hypothesis	Study	Result
H9: Voice arousal moderates the effect of clarified jargon on sales performance. When voice arousal is low, the positive effect of clarified jargon on sales is stronger. When voice arousal is high, the positive effect of clarified jargon is weaker.	3	Supported
H10: A happy facial expression moderates the effect of clarified jargon on sales performance. When influencers display a happier face, the positive effect of clarified jargon on sales is stronger. When influencers display a less happier face, the positive effect of clarified jargon on sales is weaker.	3	Not supported
H11: A happy facial expression moderates the effect of voice arousal on sales performance. When the happy face is high, the negative effect of voice arousal on sales is weaker. When the happy face is low, the negative effect of voice arousal on sales performance is stronger.	3	Supported

Table 14. Integration with existing literature.

Modality	Variables	Our novel findings	Past literature
Content	Clarified jargon	Positive effect	(In-person persuasion) N/A; Unclarified jargon terms impeded understanding
Voice	Voice arousal	Negative	(non-livestream occasion) Positive (Laukka et al., 2005; Scherer, 1986; Mehrabian & Russell, 1974; Allison et al., 2022),
Facial	Happy face	Negative	(traditional endorsement, in-person persuasion) Positive (Ilicic & Baxter, 2018; Reed et al., 2018; Van Kleef et al. 2015)

Table 15. Contrast of top and bottom influencers in employing text, voice, and facial expression

	Top influencers	Bottom influencers
Ranking	High	Low
Sales	High	Low
Clarified jargon	Frequent	Rarely
Voice arousal	Low	High
Happy face	Low	High

Figure 1. Post-hoc analysis on voice arousal

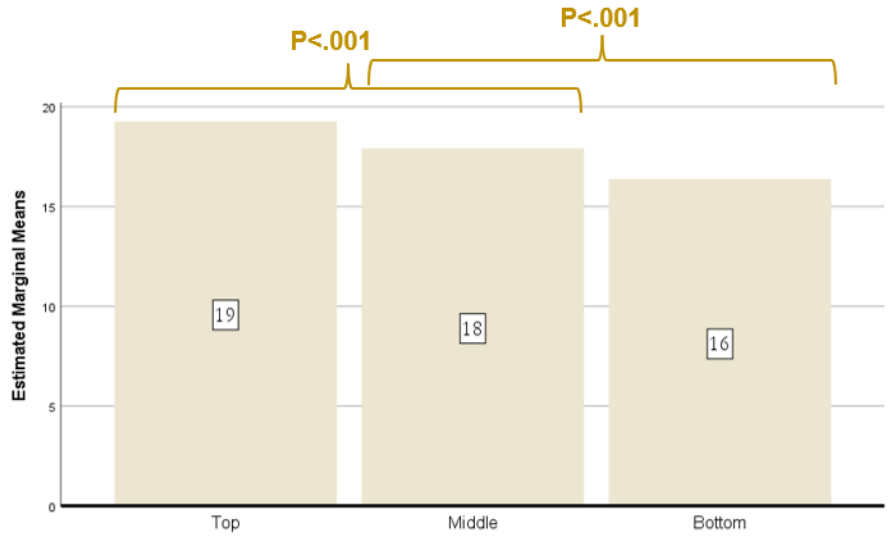

Figure 2. Post-hoc analysis on happy face

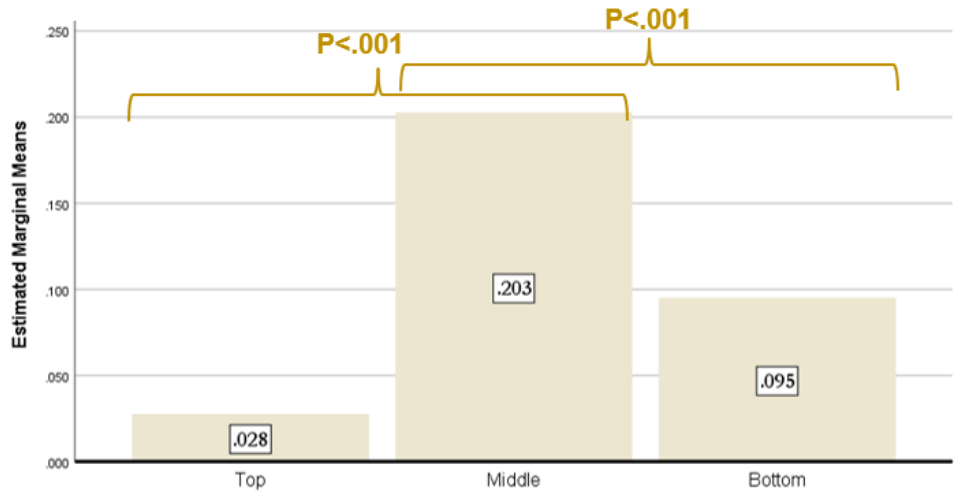

Figure 3. Model of experiment

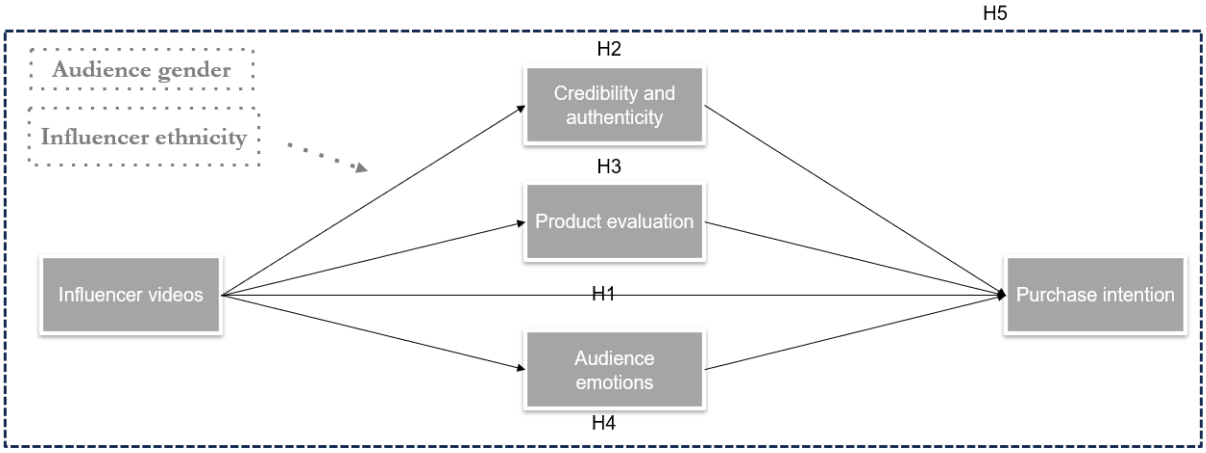

Figure 4. Screenshots of the acting influencer in top condition in Study 2

Canadian actress performing top influencer VS. Chinese actress performing top influencer



Figure 5. hypothese diagram

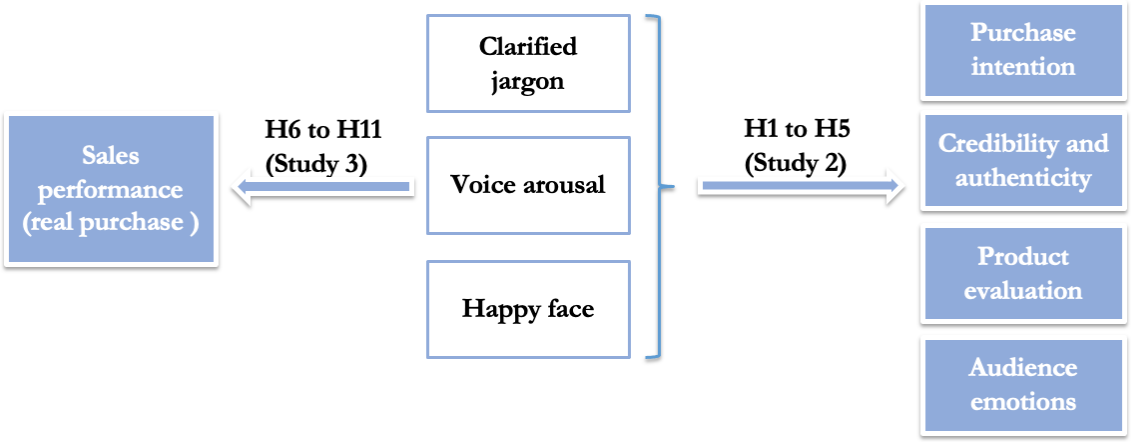

Figure 6. Sampling process

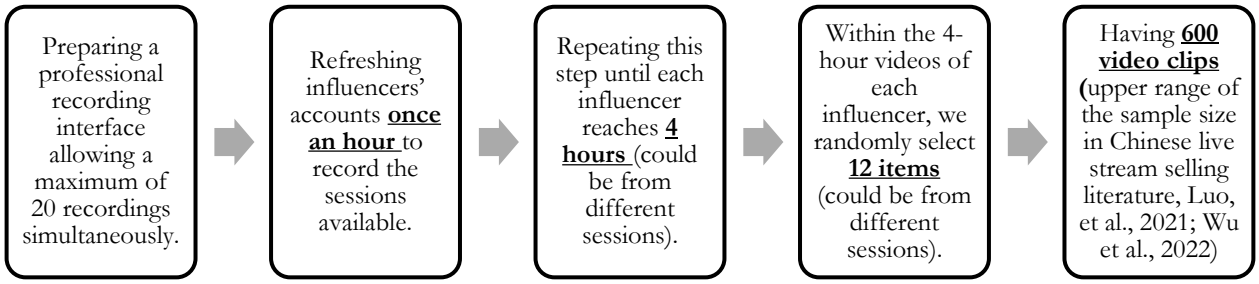

Figure 7. Result of multi-level model.

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	11.4949	0.3212	48	35.79	<.0001
Jargon_Frequency	0.2858	0.1071	547	2.67	0.0079
Ave_happiness	-0.03448	0.05275	547	-0.65	0.5136
Voice_arousal	-0.07690	0.1011	547	-0.76	0.4471
ranking_influencer	-0.00274	0.000972	48	-2.82	0.0069

Figure 8. Result of multi-level model with interactions.

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	11.4067	0.3298	48	34.59	<.0001
Jargon_Frequency	0.8519	0.4783	543	1.78	0.0755
Ave_happiness	-0.4513	0.3244	543	-1.39	0.1648
Voice_arousal	-0.08865	0.1197	543	-0.74	0.4594
Jargon*Ave_ha*Voice_	-0.1822	1.0811	543	-0.17	0.8662
ranking_influencer	-0.00274	0.000936	48	-2.93	0.0052
Jargon_Fr*Ave_happin	2.1852	3.1470	543	0.69	0.4877
Ave_happi*Voice_arou	0.4705	0.3613	543	1.30	0.1934
Jargon_Fr*Voice_arou	-0.3036	0.1875	543	-1.62	0.1059

Figure 9. Result of multi-level model with inflluencer ranking as the moderator.

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	11.5794	0.4803	48	24.11	<.0001
Jargon_Frequency	0.2282	0.1701	544	1.34	0.1802
Ave_happiness	-0.3788	0.5741	544	-0.66	0.5097
Voice_arousal	-0.08921	0.1840	544	-0.48	0.6280
ranking_influencer	-0.00331	0.001925	48	-1.72	0.0921
Jargon_Fr*ranking_in	0.000317	0.000697	544	0.46	0.6489
Ave_happi*ranking_in	0.002560	0.004253	544	0.60	0.5474
Voice_aro*ranking_in	0.000064	0.000689	544	0.09	0.9255