Copenhagen Business School
MSc in Business Administration and E-Business

Master's Thesis
Supervisor: Rob Gleasure

Between Peer and Promoter: Why Do We Care So Much About Influencers?

Students:
Suzanne Alley - 121297
Olivia Engelbach - 133909

Page Count: 100
Character Count: 213.256

Copenhagen, May 14th, 2021

Abstract

With millions of influencers around the world finding success on social media, this sought-after occupation is not just a fleeting trend. Influencers amass followings by exposing their personalities and private lives on social media and building friendship-like bonds with their audiences. Simultaneously, influencers also monetize their following through integrating sponsorships into their content, leading them to take part in two conflicting worlds - friendship and business. Motivated by this paradoxical setup, this study's objective is to explore why users are motivated to interact with influencers, even though interactions are one-sided and often heavily seasoned with advertisement. Therefore, this study sets out to answer the question "What motivates social media users to interact with influencers, whose attention might be diffused and commercially motivated?".

Approaching this phenomenon from two existing theories, the main assumption is that social media users interact with influencers out of (perceived or real) similarity and/or likability as argued by the similarity-likability theory. Contrarily, an aspirational motivation as captured in the self-expansion theory is considered, assuming that social media users are inclined to interact because of dissimilarities and luring personal growth opportunities.

From these theories, three sets of hypotheses are derived and quantitatively analyzed by looking at a dataset of 25,000 recently collected posts from 250 fashion influencers on Instagram. Producing a variety of interesting, multi-layered findings, overall results contradict the study's main assumption and show that social media users predominantly interact with influencers out of self-expansion motives.

As one of the rare existing studies that examines the user-influencer relationship from the user perspective, this research fills a notable research gap, and also holds meaningful implications for influencers, businesses, social media platforms, and regulating authorities.

Table of Contents

List of Figures

List of Tables

Introduction

While the term *influencer* may only have been added to the dictionary in 2019 (Lugo, 2019), the number of influencers and the popularity of influencer marketing has rapidly expanded over the past 10 years (Martineau, 2019). Influencers, or individuals who have amassed a large following on social media platforms (Haenlein et al., 2020), are opinion leaders in the social media world and have created the opportunity for themselves to communicate with large quantities of people despite not being a celebrity (Breves, Liebers, Abt, & Kunze, 2019).

Due to their powerful influence over consumers, influencer marketing has become the fastest-growing online customer acquisition method (Digital Marketing Institute, 2018). Valued at US$10 billion in 2020 (Haenlein et al., 2020), the influencer marketing industry is expected to be worth US$15 billion by 2022 (Business Insider, 2021). With 3.6 billion people using social media worldwide in 2020 (Tankovska, 2021a), influencer marketing has become unavoidably relevant for all businesses, especially business-to-consumer firms, and almost all companies in industries ranging from beauty to fashion, travel, food, and many more, are utilizing influencers as one of their main marketing methods (Haenlein et al., 2020).

While there is no shortage of brands wanting to collaborate with influencers, there is also no scarcity of individuals aspiring to become influencers and pursuing it as a full-time career (Haenlein et al., 2020). In total, across Instagram, YouTube, and TikTok, there are estimated to be anywhere between 3,200,000 and 37,800,000 influencers (Mediakix, 2021). In a 2019 study conducted in the UK, the job titles *influencer* and *YouTuber* made up two of the top three career choices sought after by teenagers (Haenlein et al., 2020). According to another study, over half of Gen Z and millennials would become influencers if the opportunity arose, and over 85% would share sponsored content (Locke, 2019).

Not only are Gen Z and millennials most likely to become influencers themselves, but they are also most likely to interact with them on social media platforms (Morning Consult, 2019). Around three-fourths of Gen Z and millennials follow social media influencers, with Gen Z being even more likely to follow and interact with a higher number of influencers (Morning Consult, 2019).

While these groups may be the most susceptible to influencer marketing, studies show that 60% of all consumers have previously been impacted by social media in regards to their in-store shopping decisions and that almost 75% of consumers use their social media networks for product recommendations (Digital Marketing Institute, 2018).

What may be especially intriguing about the job description of an influencer is the fact that it, in an entirely unique and novel way, obscures the lines between a perceived friendship and an advertising channel, facets of both of which are exhibited by the influencer constantly and, most of the time, simultaneously.

On the one hand, influencers share details of their personalities and private lives with their community as if they were close friends, talking about their daily activities, to-dos, and struggles. Doing so in an often charismatic or otherwise capturing way allows influencers to show that they are like everybody else and to build a loyal community of followers who often know an influencer at least as well as real-life friends would. As a result of this perceived friendship, followers establish high levels of trust with influencers and therefore tend to listen to and be influenced by them when it comes to advice and recommendations (Jin, Muqaddam, & Ryu, 2019).

On the other hand, this behavior naturally opens up the opportunity for influencers to monetize their occupation and enter into a unique, intermediary role between a peer and a promoter, where they represent both a consumer and company at the same time. Hereby, influencers turn their social media feeds into one-person marketing channels on which they get paid to broadcast messages on behalf of companies to their followers, who are much more inclined to trust an influencer’s recommendation over direct corporate marketing (Jin et al., 2019). Thus, influencers not only lend their looks and charisma to companies like models or celebrity endorsers but more so their entire package, which comes with a loyal community of like-minded followers.

Summing up, the phenomenon of social media influencers is essentially built on a paradoxical setup in which influencers are friends who share personal details about their private lives and personalities with their community, but simultaneously, they are also sales-people knocking on their followers’ doors with targeted product messages to deliver and plenty of things to sell.

While this setup is highly desirable for both influencers and businesses, it may however seem too conflicting to create value for followers at first glance. Neither the perceived friendship with an influencer is real, nor may their genuine product recommendations be. In spite of everything, influencers are as popular and in-demand as ever (Haenlein et al., 2020).

Fascinated by this contradiction and the overwhelming success of influencers in the social media landscape in recent years, this study's motivation is to investigate the appeal that users see in the one-sided interaction with influencers on Instagram, who are neither actual friends with whom personal relationships are maintained online, nor celebrities whose extravagant lifestyles and personal lives may be particularly fascinating to the outside world.

Therefore, this study sets out to examine the contradictory phenomenon of user-influencer interaction on social media by asking the research question:

➢ "What motivates social media users to interact with influencers, whose attention might be diffused and commercially motivated?"

In order to investigate this question, this study focuses on the particular user-influencer interactions on social media platform *Instagram*, which is selected as the focal context of this study as it not only attracts around one billion monthly active users (Tankovska, 2021b) but is also considered the most popular social media platform for influencer marketing (Evans, Hoy, & Childers, 2018). Due to Instagram's visual-centric, aesthetic appeal, and its ability to reach worldwide audiences, especially fashion content ranks among the most popular and thriving niches on the platform (Jin et al., 2019). As a result, Instagram is the social network with the highest amount of user engagement with fashion-related content, as well as the largest quantities of diverse fashion influencers (Pathak, 2015). Given these beneficial conditions and in order to delimit this study to an appropriate scope, this research specifically looks at the context of fashion influencers on Instagram.

Furthermore, since Instagram is a platform of global nature, the context of this study does not focus on one specific region but selects fashion influencers from a variety of English-speaking countries. Moreover, considering the platform and user dynamic on Instagram, a cross-sectional, highly recent context is selected in order to generate useful findings, which led to the collection of 25,000 Instagram posts shared in or shortly before February 2021.

More specifically, in order to answer the research question, this study takes a mostly positivist and deductive approach by testing two existing theories, that is, the similarity-likability and self-expansion theory, in the context of social media interaction. In order to confirm or reject these theories and expand existing knowledge, a quantitative, non-experimental research design based on a large, real-life sample of user behavior observations is conducted. The collected sample consists of a variety of recent, publicly observable Instagram data, including three kinds of engagement rates, that is, like rate, comment rate, and overall engagement rate, which are used to quantify the interaction between users and influencers. Then, an empirical analysis of the dataset is conducted to evaluate the study's hypotheses, examine causalities of interaction on social media, and derive transferrable generalizations with important repercussions for various stakeholders.

Looking at existing research, it becomes clear that influencers have not only been a recent hot topic in the marketing world but that there have also been conducted a number of studies examining the phenomenon. However, unlike the focus of this study, most previous works have focused on the influencer perspective, dissecting their role and its unique characteristics (e.g. Abidin, 2016; Duffy & Hund, 2015; Duffy & Pooley, 2019; Khamis, Ang, & Welling, 2017; Ge & Gretzel, 2018; Jin et al., 2019; Raun, 2018; Audrezet, de Kerviler, & Moulard, 2020) and how influencers are used as a marketing mechanism (e.g. Argyris, Wang, Kim, & Yin, 2020; Breves et al., 2019; Campbell & Farrell, 2020; Evans et al., 2018; Folkvord, Roes, & Bevelander, 2020; Kádeková & Holienčinová, 2018; Haenlein et al, 2020; Konstantopoulou, Rizomyliotis, Konstantoulaki, & Badahdah, 2019; Stubb & Colliander, 2019). While there exist a few preceding studies that focus on a user perspective, they are mainly concerned with the negative repercussions of interacting with influencers, such as being exposed to deceitful advertising (Van Reijmersdal & Van Dam, 2020; Evans et al., 2018) or unrealistic images, which diminish mental health (Chae, 2018). To our knowledge, this study, therefore, is the first to bring a new perspective to influencer literature and attempt to explore why social media users interact with influencers in the first place. This study will, thus, close a gap in preceding research and can be used as a springboard to delve deeper into the motivational perspective of the user-influencer relationship beyond just the negative repercussions.

The paper is structured as follows: first, a literature review of previous studies surrounding influencers will be presented, introducing relevant topics such as the contextual distinction of influencers, the history and types of influencers, social media platforms, and influencer marketing. Then, the study's underlying theory of what motivates social media users to interact with influencers, whose attention might be diffused and commercially motivated, will be presented within the two main presumptions, that is, the similarity-likability theory and the self-expansion theory, followed by the presentation of three sets of research hypotheses. Afterward, the study's methodology will be addressed, followed by an overview of the empirical results emerging from the data analyses conducted. Further robustness checks will be reviewed, and subsequently, the results will be analyzed and discussed by gradually moving from detailed to broader implications. Herein, the study's contributions to knowledge, future research ideas, implications for practice, and limitations are put forward. Finally, a brief conclusion will be drawn, summarizing the study and answering its research question.

Literature Review

In the past years, the label *influencer* has become a ubiquitous buzzword of everyday life, where social networking has come to be one of the most popular pastimes (Tankovska, 2021a). In 2020, over 3.6 billion people used social media platforms worldwide, a number that is expected to increase to 4.41 billion by 2025 (Tankovska, 2021a), more than half of the world population.

While there exist a variety of definitions for influencers, they are generally defined as individuals who have gained a large following on social media (Jun & Yi, 2020; Kádeková & Holienčinová, 2018; Jin et al., 2019; Haenlein et al., 2020; Bakshy, Hofman, Mason, & Watts, 2011; Abidin, 2016; Argyris et al., 2020; Breves et al., 2019), whereas *following* or *followers* are defined as the individual profiles that subscribe to an influencer's content (Argyris et al., 2020). Influencers are regarded as trusted tastemakers in any niche or area of interest they may focus on (Jin et al., 2019; Breves et al., 2019; Jun & Yi, 2020; Van Reijmersdal & Van Dam, 2020), which may include for example health and fitness, fashion and beauty, video games, or travel (Argyris et al., 2020; Campbell & Farrell, 2020).

Along with creating capturing content about a specific topic or lifestyle, influencers also gain a large following through connecting with their community on a highly personal level (Argyris et al., 2020), typically involving personal aspects of their lives that would usually be inaccessible to the public (Abidin & Ots, 2016). Moreover, influencers also usually embody a mixture of desirable physical or personal attributes that allow them to capture and directly impact their audiences (Bakshy et al., 2011; Brown & Hayes, 2008).

Besides creating and sharing curated content in a personal manner with a community of followers, another central aspect of being an influencer is the monetization of this very occupation, which most commonly takes the shape of advertising specific products or brands in exchange for compensation (Kádeková & Holienčinová, 2018; Abidin, 2016; Campbell & Farrell, 2020). These advertisements are often referred to as *advertorials* (a combination of the terms *advertisement* and *editorial*) and are professionally known as “highly personalized, opinion-laden promotions of products/services that influencers appear to personally experience and endorse for a fee” (Abidin, 2015, p.1). Advertorials constitute the main way individuals become full-time social media influencers (Campbell & Farrell, 2020), which, according to a 2019 study, ranks among the top three most sought-after career aspirations of British teenagers (Haenlein et al., 2020). It is worth mentioning that the faces of influencers do not necessarily have to be human, and the term’s definition has continuously expanded with countless new variations brought into being, such as pet accounts (e.g. Doug the Pug (@itsdougthepug), over 4,000,000 followers) or even AI-generated influencers, such as Lil Miquela (@lilmiquela), who has amassed over 3,000,000 followers and secured influencer marketing campaigns for various renowned fashion labels (Kumpumäki, n.d.; Campbell & Farrell, 2020).

Influencers Versus Celebrities

In an attempt to further frame the term influencer, there has been some controversy over whether celebrities should also be considered influencers, particularly in regards to the distinctions between the two titles. While some researchers have argued that influencers also include people who are famous and exert certain influence through institutional settings outside of social media, such as athletes, actors, and singers (Haenlein et al., 2020; Breves et al., 2019), most definitions describe influencers as individuals who started their online careers as ordinary, every-day Internet users (Abidin, 2016) and became increasingly well-known on social media without institutional mediation, that is, merely by the continuous production of capturing content. Therefore, influencers, in contrast to traditional celebrities and other personas of public interest with significant influence, are most often thought of as products of social media or someone who would not be known without and outside of social media (Haenlein et al., 2020; Breves et al., 2019; Argyris et al., 2020).

Adapting this view, this study therefore generally considers social media influencers and celebrities as being mutually exclusive. While celebrities rely on and are often subjected to traditional media outlets reporting on their work and lives, such as magazines and television, along with using digital media outlets to share their own content, influencers first and foremost rely on social media to communicate with their audience in a typically more open and proactive way (Breves et al., 2019). Social media influencers are not able to initially generate positive images through successful work in the public eye, instead, they are solely being judged based on their social media presence, where every action is critical in increasing and maintaining their following (Breves et al., 2019).

While social media influencers are thus distinguished from traditional celebrities and vice versa, influencers are nevertheless referred to as *micro-celebrities* in the particular context of social media (Abidin & Ots, 2016; Raun, 2018; Marwick, 2013). A micro-celebrity is essentially characterized by an individual's popularity linked almost exclusively to the Internet and an audience they can strategically maintain through creating and sharing their identities online, as well as through ongoing interaction and communication (Abidin & Ots, 2016; Khamis et al., 2017; Raun, 2018; Marwick, 2013; Senft, 2013). Unlike celebrities, micro-celebrities are famous only to a niche group of people and are perceived as more authentic (Abidin & Ots, 2016), and while celebrities are typically more distanced from their audience for the sake of maintaining their privacy, the popularity of micro-celebrities depends on remaining connected to their audience, or else they will lose their audience's attention (Abidin & Ots, 2016). Due to their similar definitions, the terms micro-celebrity and influencer are thus often used interchangeably (Jin et al., 2019).

History of Influencers

Generally, the idea of influencing others in buying decisions is not a novel concept, and customers have always been influenced by charismatic salespersons, celebrity endorsers, or simple word-of-mouth recommendations from people within their environments (Burns, 2020).

Today, this has primarily taken the shape of social media influencers, who have been found to have a bigger effect on purchasing decisions than traditional product endorsers (Van Driel & Dumitrica, 2021; Arnold, 2017) and therefore play a vital role in the modern marketplace. Over the past 20 years, various factors have contributed to the emergence of influencers as they are known today (Burns, 2020), the most significant of which will be presented shortly in the following.

Web 2.0

Coined by Tim O'Reilly in 2005, the term *Web 2.0* explains how, in the early 2000s, the web shifted from being "based on a network of hypertexts" (Barassi & Treré, 2012, p.1270) to a construct which harnesses participation, facilitating "the co-production of information, social networking and rich user experiences" (Barassi & Treré, 2012, p.1271; O'Reilly, 2005).

With more and more capabilities of personal computers and other devices available to users, new business models like blogs and social networks started emerging (Laudon & Traver, 2019), and especially the increase of shared communication services allowed for a more decentralized type of user-generated content on the Internet, which caused a revolution "toward increasing interactivity, social networking, and the culture of sharing" (Pedroni, 2015, p.180). Within a very short period of time, Web 2.0 applications became extremely popular and attracted huge audiences (Barassi & Treré, 2012). By giving users the ability to share their own point of view and the potential to easily reach millions of people (Pedroni, 2015), Web 2.0 laid the core foundation for the emergence of social media influencers (Raun, 2018; Konstantopoulou et al., 2019).

Blogs

The very first version of social media influencers arose when *blogs*, short for *weblogs*, became popular in the early 2000s (Hayes, 2018; Perlmutter, 2005). Blogs are personal web pages created by individuals that usually consist of a number of chronological entries posted by their author (Laudon & Traver, 2019), wherein entries are typically used by bloggers to log their thoughts or the ongoings of their everyday lives in the form of a digital journal (Siles, 2011). Blogs were especially novel because they shifted the power away from the big media and instead allowed regular people to post content they cared about (Duermyer, 2021). Furthermore, due to the interactivity of blogs, where users have the ability to comment and like posts, blogging quickly started gaining popularity and garnering large communities of followers (Barlow, 2007).

Social Media Platforms

Another product of the rapid growth of the Web 2.0 set of applications was the launch of social media platforms Facebook and Twitter in 2005 (Laudon & Traver, 2019). Just shortly after blogs became popular, social media platforms made it increasingly easy for bloggers and content creators to be discovered by users interested in similar topics, and thus reach a wider audience (Burns, 2020). Simultaneously, Youtube, which was also launched in 2005, first allowed amateur video makers (so-called *vloggers*) to upload their own video material for free, and thus provided an easily accessible and centralized video consumption platform which immediately recorded strong growth and big popularity (Christensen, 2007).

Reality-TV

Another factor which plays a major role in the history of influencers was the rise of reality TV, that is, television content revolving around real, everyday people instead of made up characters (Andrejevic, 2004). Just like traditional figures of the entertainment industry, reality TV personalities quickly attract large audiences. However, due to the number of personal moments shared on-screen, audiences perceive them to be more authentic and real (Jin et al., 2019). Thus, reality TV personalities tend to be much more relatable than movie stars who traditionally try to protect their privacy (Driessens, 2012). This is due to the fact that, in contrast with Hollywood movies, reality TV revolves around the very publication of personal details, the reputation, and everyday lives of its cast who often gains celebrity status and huge popularity themselves, like the members of the Kardashian family, whose lives were documented over many years in the juggernaut reality TV show *Keeping up with the Kardashians* (Bernstein, 2019; Driessens, 2013). Therefore, reality TV stars garner stronger connections with their audiences, which was found to result in higher purchase intentions for products endorsed by them, because consumers were found to personally identify with them and often try to imitate them (Jin et al., 2019).

Similar to the concept of reality TV, today, ordinary social media users can become influencers by sharing details about their charismatic personalities and everyday lives with their audiences (Van Driel & Dumitrica, 2021). However, unlike reality TV stars, they do not need the resources required for an actual reality TV production. Garnering a loyal following by being relatable, open, and sharing personal struggles that regular people might face on a daily basis, social media influencers have become self-branded and self-produced online personalities slowly accumulating “celebrity capital” (Driessens, 2013, p.543) similar to reality TV personalities (Evans, 2009).

Recent History

Over the years, increasingly visual-based social media applications like Instagram, Snapchat, or Tik Tok have launched and, besides loosely connecting individuals (Gleasure, 2019) as well as entertaining their users, they have embraced their role as a democratized exchange between companies and consumers (Burns, 2020). The wide variety of different platforms to use and share content continues to fuel both the supply and demand of social media influencers (Argyris et al., 2020). Furthermore, the recent shift to mobile has greatly facilitated both the creation of content as well as its consumption, which is possible from anywhere anytime and nowadays, enabling influencers to share whichever details of their everyday lives they would like to let their loyal audiences be part of (Droesch, 2019).

Self-Branding

A central aspect of becoming and being an influencer on social media is the activity of self-branding. In the recent past, self-branding has become common practice, with an increasing number of individuals making regular efforts to curate the image they present to the world (Duffy & Pooley, 2019). Scholars have been linking this trend not only to the growth of social media (Khamis et al., 2017) but more so to the rise of individualist, freelance work (Duffy & Hund, 2015; Beals, 2010) and “the unsettling experience of you’re-on-your-own workplace capitalism” (Duffy & Pooley, 2019, p.28) that encourages aspiring online entrepreneurs to manage their brands.

Given the increasing competition to secure lucrative advertising deals in the crowded marketplace that social media is today, most influencers, therefore, have made substantial efforts to market themselves to drive their success (Gandini, 2016), creating carefully curated professional images of themselves to highlight their unique selling points and to ultimately create a singular, charismatic public and social identity or personal brand that is responsive to the needs and interests of their target audience and companies (Khamis et al., 2017).

Commonly defined as a cultural, ideological, and sociological object that holds value (Gandini, 2016), a brand signifies a certain quality or characteristic associated with it and ultimately aims at influencing the consumer’s decision-making process (Khamis et al., 2017). In human form, a brand is defined as any widely known personality who is engaging in marketing communication efforts (Thomson, 2006), a concept that has become increasingly prominent ever since the widespread use of cable television, the internet, and social media (Jun & Yi, 2020).

In contrast to the traditional celebrities that are often contracted to endorse products, influencers however have yet to make a name for themselves and gain what Bourdieu (1986) defines as social capital. Social capital, which does not generate immediate pecuniary return but may be converted into economic capital only at a later point in time, is crucial for the commercial success of these micro-celebrities (Gandini, 2016; Khamis et al., 2017). As Gleasure & Morgan (2017) put it, physical capital is enabled by human capital, which itself is enabled by social capital. Therefore, one of the core aspects of social capital is the acquisition of a reputation, which is the return that is expected of self-branding (Gandini, 2016). Put differently, the practice of self-branding can essentially be defined as the construction of social capital (Gandini, 2016). Thus, from a more practical point of view, self-branding can be considered a strategic effort aimed at the generation of tangible, economic return. Important to emphasize here is that while the construction of social capital is mostly unremunerated, it is nevertheless indispensable to self-construct and secure employment as an influencer (Hearn, 2008).

The importance of garnering social and cultural capital as an influencer has become especially important given the saturation of the online influencer marketplace, which is also deemed an *attention economy* (Fairchild, 2007), and where the value of distinctiveness and visibility has constantly increased (Brody, 2001). Thus, every individual needs to manage their own brand identity in order to stand out from the competition and attract and deliver value to customers and employers alike, and in doing so become commercially successful (Peters, 1997; Chen, 2013). Some even go as far as to equate this strategic self-expression as the “monetization of ‘being’” (Hearn, 2011, p.315). In the context of the attention economy, scholars also equal the process of self-branding with exposure (Marshall, 2016), attention-seeking, and status-enhancing behavior (Marwick, 2013; Shepherd, 2005), directed at achieving celebrity status (Trammell & Keshelashvili, 2005). Simultaneously, self-branding is also aimed at building a deep, intimate, and loyal relationship between influencers and their followers (Abidin & Ots, 2016).

Importantly, self-branding through social media does not require initial affiliation with the already powerful and frees individuals from the dependence on traditional gatekeepers of fame (Marwick, 2015), which may be considered another appealing reason responsible for the rapid increase of aspiring micro-celebrities. Thus, influencers generate celebrity capital through authentic self-branding, trying to garner as much attention as possible on their social media channels, which they can then utilize as their own marketing platform on behalf of advertisers (Hearn & Schoenhoff, 2016).

Besides authenticity (Duffy & Pooley, 2019) and the overall desire to make a favorable impression on others (Petroni, 2019), further important aspects of successful self-branding and earning strangers’ trust are the continuity and consistency of narrative and image (Duffy & Hund, 2015). Various authors have underlined the significant, unremunerated efforts this entails, pointing out the high emotional labor, vigilance, and thick skin required (Marwick, 2013; Chae, 2018; Audrezet et al., 2020; Petroni, 2019), while others have highlighted that, despite the perks enjoyed when successful, relying on one’s own brand also quickly shifts the burden of labor security to the individual (Khamis et al., 2017) and thus makes becoming an influencer not a risk-free endeavour.

Types of Influencers

When thinking and writing about influencers, it is crucial to consider that not all influencers are equal and that different influencers have different self-branding techniques, topics of focus, engagement rates, sizes and types of following, ways of influencing their communities, as well as monetary demands for collaboration (Campbell & Farrell, 2020; Ge & Gretzel, 2018). Considering these different aspects, a variety of terms have been tossed around in the social media world to describe and categorize influencers (Haenlein et al., 2020), which typically not only define follower size, but also "perceived authenticity, accessibility, expertise, and cultural capital" (Campbell & Farrell, 2020, p.471). The five categories selected to classify influencer types in this study are: mega-influencer, macro-influencer, mid-tier-influencer, micro-influencer, and nano-influencer, as visualized in Figure 1 below.

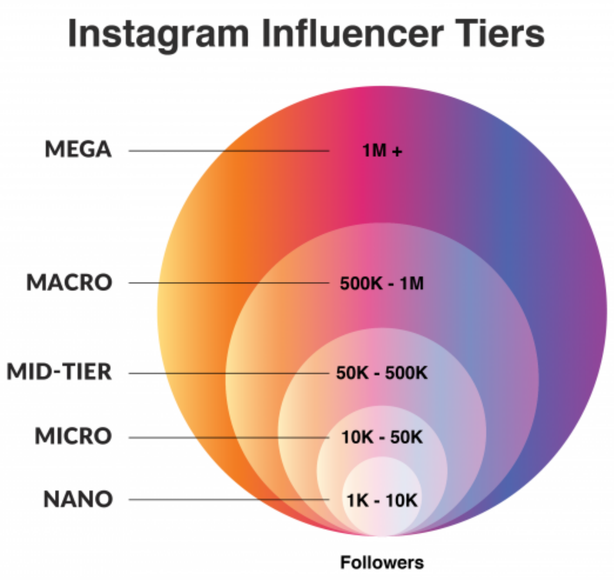

Figure 1: *Instagram Influencer Tiers (Mediakix, n.d.)*

Mega-Influencers

With 1,000,000 or more followers, *mega-influencers* have the largest audience on social media and are therefore the most well-known of their kind (Campbell & Farrell, 2020; Mediakix, n.d.; Ismail, 2018; Britt, Hayes, Britt, & Park, 2020). Mega-influencers have the most widespread impact, giving them the ability to shape contemporary pop and media culture, steer trends and drive hashtags, generate significant demand for products, and direct attention to a variety of topics (Mediakix, n.d.). Naturally, their audience is more diverse and global, and they typically also cover a broader area of interest compared to smaller influencer types (Ismail, 2018). Due to their large following and resulting number of impressions they create, mega-influencers are compensated the most to advertise brands and products on their channels (Campbell & Farrell, 2020; Mediakix, n.d.).

However, while mega-influencers may have the most power and impact on social media platforms, they are also more prone to skepticism in regards to their authenticity, trustworthiness, and the sincerity and genuineness of their endorsements (Britt et al., 2020). As consumers are less likely to trust product or brand recommendations from a celebrity than a non-celebrity (Ismail, 2018), this skepticism is increased when an influencer's popularity expands beyond the social media world (Ismail, 2018). As a result, mega-influencers typically record the lowest engagement rate among the different influencer types (Mediakix, n.d.).

Macro-Influencers

Amassing between 500,000 and 1,000,000 followers, *macro-influencers* have the second-largest audience (Mediakix, n.d.) and typically have stronger engagement rates than mega-influencers, while still reaching and impacting a wide audience that often aspires to be like them (Campbell & Farrell, 2020). Given their perception as content leaders and experts in their area of interest, macro-influencers charge the second-highest for brand partnerships and product exposure on their channels (Campbell & Farrell, 2020). Macro-influencers usually do not have celebrity status outside of social media as some mega-influencers do, but instead gain and maintain their popularity through the Internet and their curated online brand (Campbell & Farrell, 2020; Ismail, 2018). Therefore, while they may be famous to large masses on the Internet and usually pursue influencing as full-time careers, they are not usually household names beyond their following (Campbell & Farrell, 2020; Mediakix, n.d.).

Mid-Tier-Influencers

Next, *mid-tier-influencers* share content with between 50,000 to 500,000 followers (Izea, 2020; Mediakix, n.d.). Not as well-known as mega-influencers and macro-influencers, mid-tier influencers have risen in popularity, have higher engagement rates, and are generally perceived as more trustworthy and less distant from their audience than larger influencers (Izea, 2020; Mediakix, n.d.). Mid-tier influencers often constitute the stage where individuals transition into becoming full-time influencers (Izea, 2020; Mediakix, n.d.), and are characterized by attracting both niche audiences as well as more general audiences.

Micro-Influencers

Micro-influencers are smaller than the previous influencer types mentioned regarding both their scale and their scope (Campbell & Farrell, 2020). Their audience ranges from 10,000 to 50,000 followers and tends to consist of users from the same geographic region as the influencer (Izea, 2020; Mediakix, n.d.; Campbell & Farrell, 2020). Primarily focusing on specific areas of interest or aesthetics, they are able to grow a tight community around their niche (Izea, 2020; Mediakix, n.d.). Micro-influencers harness more trust and are perceived as more accessible, passion-driven, and in touch with their audience, which is considered to cause higher engagement rates (Campbell & Farrell, 2020; Britt et al., 2020; Izea, 2020; Mediakix, n.d.) that have made them attractive marketing channels for brands looking to target specific market segments (Mediakix, n.d.). However, given their limited reach, there has been a certain stigma around micro-influencers, as they have been known to pay for fake followers in order to appear more popular (Mediakix, n.d.).

Nano-Influencers

Lastly, *nano-influencers* are characterized as having between 1,000 and 10,000 followers, which consist of mostly friends, acquaintances, and other local users (Izea, 2020; Mediakix, n.d.; Campbell & Farrell, 2020). Nano-influencers often have a highly personal relationship with their audience and therefore usually have the highest perceived authenticity and engagement rates out of all the influencer types (Izea, 2020; Mediakix, n.d.; Campbell & Farrell, 2020), which may also be caused by higher influencer-sided interaction rates, such as personal replies to comments or private messages (Izea, 2020). Like micro-influencers, nano-influencers are also more niche-focused, often with an additional local focus (Izea, 2020).

Given their comparatively small follower numbers, nano-influencers naturally have the most limited reach and impact on social media (Campbell & Farrell, 2020; Mediakix, n.d.) Frequently, they are just starting out their social media careers or pursue it as a hobby, and are therefore perceived to have the least amount of industry knowledge and established brand relationships (Campbell & Farrell, 2020). Instead of monetary compensation, these influencers are most open to pecuniary unremunerated partnerships and instead accept free product samples in return for social capital through professional networking and increased exposure on social media (Campbell & Farrell, 2020; Mediakix, n.d.). In order to launch their careers, nano-influencers are also more likely to approach brands themselves about partnership opportunities rather than waiting to be approached (Campbell & Farrell, 2020).

For a full overview of the previously presented influencer types and their general characteristics, see Figure 2 below.

Figure 2: *Influencer Types & Characteristics*

Social Media Platforms

Constituting the core foundation for the existence of influencers today, social media platforms, such as Facebook, Twitter, YouTube, Instagram, and TikTok, are all essentially open forums where users connect with one another and, to varying extent, compete to gather audiences (Campbell & Farrell, 2020; Haenlein et al., 2020), wherein every platform differs in user demographics and usage purposes (Haenlein et al., 2020). For example, Facebook and Twitter users generally tend to be middle-aged, whereas Instagram attracts millennials and TikTok the youngest users, who are mostly in their teens or early 20s (Haenlein et al., 2020). As for usage purposes, Facebook often represents a popular tool to stay connected with friends and family, while Twitter is mostly used to stay up-to-date with news and to share opinions (Haenlein et al., 2020). YouTube, Instagram, and TikTok primarily serve as entertainment and self-presentation platforms, which also makes it easier for users to gain popularity and influencer status (Haenlein et al., 2020). In order to delimit this study to an appropriate scope, Instagram is selected as the platform of particular interest to meet the underlying research objectives, and is therefore focused on and explored in more detail in the following.

Founded by Kevin Systrom in 2010 and acquired by Facebook in 2012 (Instagram, n.d.a.; Facebook, n.d.), Instagram enables users to share visuals with their followers (Evans et al., 2018) and to engage with other users by following, liking, commenting on, or sharing their posts (Haenlein et al., 2020). Given the platform's design revolving around social interaction and aesthetical presentation, which allows users to "build personal narratives and showcase identities that attract audiences" (Jin et al., 2019, p.568), it is one of the most popular platforms for (aspiring) influencers (Campbell & Farrell, 2020; Argyris et al., 2020). More than a third of American adults and two-thirds of Americans between 18-29 years use Instagram, numbers that have been continuously trending upwards (Perrin & Anderson, 2019).
Instagram's popularity has also spurred the regular use of the term *Instafamous*, which refers to users that gain a large following on and through their content on Instagram, and therefore become well-known and influential on the platform (Jin et al., 2019; Evans et al., 2018). Out of the platform's more than one billion monthly active users (Facebook, n.d.), 21.5% have a following of over 10,000, and 5.6% have a following of over 50,000 (Haenlein et al., 2020).

A user's personal Instagram feed typically contains a mix of posts from followed accounts and followed hashtags, as well as sponsored posts inserted in between by the company itself (Haenlein et al., 2020, p.23). This content is today no longer presented chronologically, but in an order determined by the platform's AI algorithm which prioritizes content the user is most likely to be interested in (Haenlein et al., 2020). Therefore, when users share a post, not all their followers may actually see it, as post visibility depends on many factors hidden in the Instagram algorithm, which has been subject to frequent changes and adaptations over the years (Haenlein et al., 2020; Warren, 2021).

While in the early beginnings of Instagram only squared picture posts with captions were possible, the company has since added multiple features aimed at providing its users with rich opportunities to create diverse, customized, and engaging content. These include for example differently formatted pictures, videos, stories (i.e., pictures or videos that are only visible for 24 hours), highlights (i.e., permanently visible stories), live videos, IGTV (i.e., longer videos comparable to those found on Youtube), and, most recently, Reels (i.e., short videos in portrait mode), capturing the growing popularity of brief and easily browsable video clips (Haenlein et al., 2020; Instagram, n.d.b.). Adding to this, myriads of special effects and creative tweaks enable users to generate and share extremely rich and entertaining content.

Furthermore, Instagram provides all business profiles, including most influencers, with *Insights,* a detailed content performance analysis with which professional users can monitor their reach, interactions, follower numbers, and many more KPIs in order to steer their content and drive their business, for example through promoting posts and thus becoming more visible (Facebook for Business, n.d.a.). Insights is not only a welcome feature used to identify one's target audience and track community size over time, it also plays a crucial role in influencer marketing.

Influencer Marketing

On social media, consumers have the power to control what content they see and who they follow (Mathew, 2018). Because of this, it has become more relevant, but at the same time more difficult, for brands to reach their desired audiences, especially since the number of social media platforms continuously grows and users become widespread (Mathew, 2018). Eventually, marketers found a solution to this problem - influencer marketing (Mathew, 2018).

Key Facts & Figures

Generally defined as when brands collaborate with influencers to target a specific audience and promote specific products or increase awareness (Bailis, n.d.), *influencer marketing* has become extremely popular on Instagram (Breves et al., 2019). Also deemed as highly credible word of mouth (Evans et al., 2018; Konstantopoulou et al., 2019), influencer marketing and the industry that has established around it has seen rapid growth in recent years, reaching US$8 billion in 2019, and is expected to be worth up to US$15 billion by 2022 (Business Insider, 2021). Today, nearly 80% of companies have dedicated a budget to using influencers to nudge the purchasing decisions of potential buyers (Influencer Marketing Hub, 2021a; Brown & Hayes, 2008). Based on the presumption that influencer content is posted by other ordinary consumers and that most posts are thought to be noncommercial, influencer marketing has been considered a more trustworthy form of advertising compared to traditional, anonymous corporate marketing (Mudambi & Schuff, 2010), and has benefited from the particularly strong bonds that influencers establish with their following over time (Park, Ciampaglia, & Ferrara, 2016; Konstantopoulou et al., 2019). As a result, reports have shown that large companies nearly doubled the number of creators involved in every campaign in the past two years and that the use of micro-influencers, whose posts are perceived as the most authentic and real (Kowalczyk & Pounders, 2016), has tripled since 2016 (Influencer Marketing Hub, 2021a; Burns, 2020).

Due to its visual nature, aesthetic appeal, and wide reach (Jin et al., 2019), Instagram has become one of the most viable social media platforms for influencer marketing, including for example selfie advertorials, product placements, or brand collaborations (Abidin, 2016), and has clearly outranked any other social media network with 80% to 87% of companies predominantly choosing Instagram for influencer campaigns (Business Insider, 2021; Influencer Marketing Hub, 2021a). Furthermore, the platform has continued to embrace and foster influencer marketing by providing users with a rich range of content formats going way beyond the creative freedom of other platforms (Haenlein et al., 2020), along with insightful performance analysis tools. Furthermore, Instagram's user base has been another reason for its popularity as a marketing channel, as the majority are millennials who are not only projected to surpass baby boomers as the most substantial purchasing age group and are simultaneously particularly attracted to influencers, but who are also more likely to make purchases directly from their phones (Fry, 2018).

Instagram Functions & Features

Recognizing these marketing opportunities, in recent years, Instagram has established a variety of new functionalities fostering e-commerce on the platform (Haenlein et al., 2020; Argyris et al., 2020). These include the *swipe up* feature, where verified users with more than 10,000 followers can add a link to their story for viewers to visit by just one simple upward swiping gesture (Kircher, 2020). This allows influencers to directly link to product websites of the items they are promoting and significantly facilitates shopping from within the app. Furthermore, the continuously improved *Search & Explore* tab aims to make it very easy for users to discover new content and is highly customized to every individual profile (Instagram, n.d.c.). By showing posts, videos, and products that are similar to what a user engages the most with, Instagram thus supports content creators and businesses alike in growing their audience and expanding their reach.

Introduced in 2017 and launched by 2019 (Facebook, n.d.), *Instagram Shopping* was created to increase the continuous growth of e-commerce on the app, where reportedly 70% of users have used the platform to discover products, and 87% have previously been inspired by influencers to make a purchase (Facebook for Business, n.d.b.). Specifically, Instagram Shopping allows businesses or influencers to tag products in their posts, enabling users to discover and shop items directly from a post without having to leave the post or app (Facebook for Business, n.d.b.; Instagram, n.d.d.). Clicking on a tagged product leads users to a product page showing more details, such as price and pictures, along with a call-to-action button to view the product on the store website, which, when clicked upon, opens the business's website within the Instagram app, facilitating on-the-spot purchases. This feature alone is expected to increase Instagram's revenue by US$10 billion in 2021 (Argyris et al., 2020). In March 2019, Instagram also started testing *Checkout* with a number of influencers and fashion labels, which allows users to place orders directly from the product page without even having to load an external store website within the app to make a purchase (Facebook for Business, n.d.b.). As of late 2020, tagged products, along with locations and posts, can also be included in *Guides*, which can be curated by any user to share easily digestible content and recommendations on any topic (Sehl, 2021). This function has been especially valuable to influencers, as it has opened up new opportunities for them to put the spotlight on recommended or sponsored products (Warren, 2020a).

With influencers being able to shop-tag products in their posts, stories, or guides, Instagram Shopping has not only improved the user experience for both creators, shops, and their followers, but it has significantly contributed to the continued rise of influencer marketing (Haenlein et al., 2020; Argyris et al., 2020) and particularly the increase of spontaneous, visually inspired e-commerce on Instagram (Triwidisari, Nurkhin, & Muhsin, 2017).

Shapes of Influencer Marketing

In practice, influencer marketing has taken on many shapes, with lines that are often blurry, as commercial posts and activities are usually woven seamlessly into the content posted by influencers (Campbell & Grimm, 2019; De Veirman, Cauberghe, & Hudders, 2017).

All forms of influencer marketing however generally share the underlying principle of social media micro-celebrities using their influence for remuneration through entering into partnerships with brands (Audrezet et al., 2020). In doing so, influencers lend their audience and their content creation expertise to brands who often struggle to create engaging content or to attain a large organic reach on their own (Cooper, 2021). In exchange for their abilities and the resulting value that is created (Audrezet et al., 2020), the most successful influencers earn several million dollars per year (McCoole, 2018). Important to highlight in this context is the unique position influencers have taken on, which is situated in between companies and consumers, where the influencer acts as both parties (Ge & Gretzel, 2018; Kozinets, de Valck, Wojnicki, & Wilner, 2010).

Overall, the most popular implementation of influencer marketing is product placement, which has been found to improve brand memorization, brand attitude, brand choice, and increase purchase intentions (Audrezet et al., 2020). Presenting a product in a usage context and relatable situation (Russell & Stern, 2006), influencers incorporate brand messages with their posts in return for varying amounts of compensation (Hearn & Schoenhoff, 2016; Lu, Chang, & Chang, 2014). Within product placement, there exist multiple degrees of brand impingement: at minimal impingement, marketers typically send free products to influencers, hoping that they will share information about the sampled products with their following (De Veirman et al., 2017). Betting on goodwill and many influencers' dependence on such gifting to foster continuous brand relationships, influencers are often not compensated in ways other than the free product and social capital garnered through such basic sponsorships (McQuarrie & Phillips, 2014; Argyris et al., 2020). At maximum impingement, influencers are paid as contractually stipulated in return for one or more posts whose content, frequency, time, text, and picture are often fully determined by the advertising company (Audrezet et al., 2020). Many different shapes of product placement exist on every point of the continuum between minimum and maximum impingement, for example where influencers are provided with products to give away to their followers, or paid with credit to purchase company products (Burns, 2020).

Other formats of influencer marketing range from affiliate links as a form of commission payment (Warren, 2020b) to social media platform takeovers that primarily serve the goal of garnering followers and increasing brand awareness, which influencers can do in a more effective and personal manner than brands themselves (Campbell & Farrell, 2020). Another popular way of conducting influencer marketing are influencer events hosted by brands who pay for travels and garner influencers at photo-worthy locations where they organize experiences and fun activities, throughout which the influencers create content with company products (Agrawal, 2016; Contestabile, 2018; De Veirman et al., 2017). Many companies also work with so-called *ambassadors*, an influencer partnership typically characterized by its longevity, shared values, and history, and the influencer's active advocacy of and engagement with the company products over time (Entwistle, 2017). Lastly, the closest form of partnership between brands and influencers is found within design collaborations, where product lines or collections are created and marketed together as common endeavour (Burns, 2020).

Regardless of the shape chosen for an influencer marketing campaign, one key aspect in the interest of both brands and influencers is the good match between their respective brand identities, which has proven to be crucial for campaign success (Koernig & Boyd, 2009). Therefore, given the wide variety and high demand for partnerships, many social media personalities work with agencies or content networks who select the best-fitting brands for their protégés and manage their various financial and contractual relationships with advertisers (McCoole, 2018; Abidin & Ots, 2016).

Ad Disclosure

When influencers incorporate sponsored content on their social media channels, they are legally required to be transparent and indicate that this content is commercially motivated advertisement (Breves et al., 2019; Evans et al., 2018; Haenlein et al., 2020; Audrezet et al., 2020). This indication usually comes in the form of a disclosure label within the text element of the content, such as using *#sponsored* or *#ad* at the end of a caption (Evans et al., 2018; Breves et al., 2019). Ad disclosures are necessary in order to protect users and make them aware of the commercial background of the content they are consuming, and they often are the only indicators distinguishing an advertising post from a regular post (Van Reijmersdal & Van Dam, 2020). In such cases, the sponsored content is called *native advertising*, as it seamlessly blends into the surrounding, unpaid content and is entirely different from traditional advertisements, such as TV commercials or billboards, and might therefore be easily mistaken by unsuspecting consumers (Evans et al., 2018; Audrezet et al., 2020; Stubb, Nyström, & Colliander, 2019). Thus, native advertising easily obscures the paid relationship between influencers and businesses, leading to the influencers' opinions to be mistaken as objective instead of subjective to compensation (Evans et al., 2018).

In order to protect consumers, authorities across the globe have implemented advertisement disclosure guidelines. For example, due to blurry lines and grey zones existing around ad disclosure, the US Federal Trade Commission (FTC) developed stricter guidelines for native advertising in 2017, which have set out to eliminate "the possibility of paid brand endorsements masquerading as organic, unpaid posts" (Evans et al., 2018, p.139) by requiring that any type of promotional social media content must include a disclosure (Audrezet et al., 2020). This disclosure should at least come in the form of *#ad* and be displayed sufficiently visible, that is, within the first three lines of the caption (Audrezet et al., 2020).

Although these guidelines have been put into place, more than 85% of influencers have reportedly not complied with the regulations (Haenlein et al., 2020), which may be related to the finding that ad disclosures have a serious negative impact on the success of the sponsorship as well as an influencer's perceived credibility (Evans et al., 2018; Van Driel & Dumitrica, 2021). Therefore, in order to preserve their own brand, engagement, and income, influencers - sometimes coerced by brands (Lee, 2020) - were found to hide the commercial nature of their posts (Van Driel & Dumitrica, 2021).
However, studies have also found that obviously sponsored content that does not contain ad disclosure or sponsorship information negatively affects source and message credibility (Stubb & Colliander, 2019), which encourages sponsorship transparency.

Penalties for deceptive advertisements as implemented by the FTC include orders to remove the bespoke content as well as monetary fines, corrective advertising, which would require the influencer to correct deceptive claims in the original ad and include disclosure in future ads, and civil penalties, which, depending on the severity of the violation, require the influencer to compensate consumers who bought the advertised product (Federal Trade Commission, n.d.).

While the FTC guidelines only apply to the US, and legislation regarding advertisement disclosure differs from country to country, the same underlying principles can be found in legislations worldwide (Podlegal, n.d.; Purtill, 2017; Government of Canada, 2019; Advertising Standards Authority, 2020; Competition & Markets Authority, 2019). These principles require general transparency about brand relationships and promotions, and manifest that influencers need to disclose when they are being paid, gifted, or loaned items from a brand to avoid misleading their followers (Competition & Markets Authority, 2019). Some countries, such as Canada, implemented stricter guidelines that are, for example, more specific as to where the ad disclosure label should be located, ruling that disclosure labels located after the *Read More* button are not visible enough (Hale, 2020). Penalties for breaching ad disclosure rules also vary from country to country, so that breaches of the ad disclosure guidelines as specified in the Australian Consumer Law (ACL) can result in fines of up to AU$ 220,000 per post (Podlegal, n.d.; Purtill, 2017).

In conclusion, influencers take on a complex role. They have evolved and continue doing so as technology changes, and adapt to the size of their audience and what their following expects from them. They move constantly along the continuum between ordinary social media users and micro-celebrities, navigating their own brands between genuine community interaction and paid sponsorship deals, which need to be embedded into their normal social media content, while yet standing out enough to meet legal requirements.

Recognizing the dynamic nature of the term *influencer,* the following section will delve in further into what makes influencers so appealing and why users interact with them.

Theory

Due to the previously discussed nature of influencers and the archetypical broadcasting of aspects of their private lives online, interactions between ordinary social media users and influencers can generally be classified as *parasocial interactions* and, developed over time, *parasocial relationships* (Breves et al., 2019; Folkvord et al., 2020; Jun & Yi, 2020; Yuan, Moon, Kim, & Wang, 2019). Characterized by their unreciprocal cognitive and affective nature, parasocial relationships are considered long-term and cross-situational, however, to some extent, imaginative bonds from the users' perspective (Breves et al., 2019). Despite the fact that parasocial relationships are generally characterized as unilateral communications lacking face-to-face interaction, followers tend to think of an influencer as a friend (Yuan et al., 2019). However, while users interact with influencers through subscribing to, consuming, and commenting, sharing, or liking their content (Breves et al., 2019), influencers hold the control over reciprocation in this type of relationship (Yuan et al., 2019), making it significantly less equal than it may be perceived by users.

In order to examine what motivates social media users to interact with influencers, whose attention might be diffused and commercially motivated, given the underlying, parasocial nature of their relationship, the following section will present two theories of interaction research, that is, the *Similarity-Likability Theory* and the *Self-Expansion Theory*, which will constitute the theoretical foundation of the subsequent research.

The Similarity-Likability Theory

When looking at follower-influencer relationships, a plethora of previous studies have pointed towards the relevance of perceived similarity among the two (Jin et al., 2019; Abidin & Ots, 2016; Argyris et al., 2020; Breves et al., 2019; Folkvord et al., 2020). For example, research has found that when influencers share certain demographics, life stages, events, or similar interests with their community, and embrace this relatability, users will perceive them as more credible and more like peers (Campbell & Farrell, 2020). As a result, users will be more likely to interact with these influencers, and in a more positive manner, as opposed to with less relatable influencers or traditional celebrities (Jin et al., 2019; Folkvord et al., 2020; Argyris et al., 2020; Campbell & Farrell, 2020). Furthermore, influencers perceived as similar by their following have also been found to be perceived as more authentic, familiar, and likable (Argyris et al., 2020; Campbell & Farrell, 2020).

Originally, this relationship between similarity and likability goes beyond parasocial relationships on social media and has its roots in *Social Identity Theory*, which was first coined in the 1970s and explores intergroup relations and behaviors on the basis of perceived group status similarities and differences (Tajfel & Turner, 1979), as well as the identity of the self in and between groups (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Generally, a *social identity* is defined as “an individual’s knowledge that he belongs to certain social groups together with some emotional and value significance to him of this group membership” (Tajfel, 1972, p.292). Consequently, social identity theory is based on the idea of in-groups and out-groups in which members of society cluster, and it furthermore defines and evaluates one’s self-concept, how the self is embedded within an in-group, and how one will be treated and thought of in a social context. This concept emerged from various experiments where study subjects were consistently found to strongly favor the (often random) group they had been assigned to, even when they had no knowledge about the other groups (Hogg, 2016). From these findings, it was concluded that “the mere fact of being categorized as a member of a group” produced ethnocentrism and competitive intergroup behavior (Hogg, 2016, p.5; Diehl, 1990), whereas it is important to note that individuals were found to favor their in-group rather than discriminate against the out-group (Hogg, 2016).

This behavior and social identity theory in general can further be traced back to the very fundamental evolutionary human motives, based on the belief that groups go to great lengths to protect and promote the well-being of their own members (Hogg, 2016). This seems like a natural behavior considering the value and significance that individuals derive from being part of a group as explained by Tajfel (1972), and the fact that social groups provide members with a shared identity that prescribes and evaluates who they are, what they should believe, and how they should behave. As one’s self is thus defined and evaluated through the group, and the shape of the group directly reflects an individual’s status, prestige, and social valence, the promotion of in-group wellbeing is a logical phenomenon within social behavior and effectively a struggle over the relative power of one’s own in-group (Hogg, 2016).

In general, groups are usually characterized by the existence of substantial agreement on in-group and out-group prototypes, which are reinforced by self-categorization of members into these norms, assimilating the self to the in-group and promoting homogeneity through deeper processes of internalizing and enacting a group’s prototype (Abrams & Hogg, 1990). Common bases for group formation typically include perceived, self-categorized, cultural, or any other inherent or perceived similarities.

Moreover, in-group members and norms not only capture intragroup similarity, they simultaneously highlight intergroup distinctiveness from relevant out-groups, which further establishes the similarity-likability theory as a logic phenomenon as seen in the early experiments of social identity research, which causes people of all ages and cultures to generally like others who are similar to themselves, including those “who share similar attitudes (...), personality traits (...), economic status (...), and even names or birthdates” (Collisson & Howell, 2014, p.385), and rather dislike those who are perceived as dissimilar (Collisson & Howell, 2014). This similarity-likability theory, also called similarity-attraction theory (Berscheid & Hatfield-Walster, 1969), is validated consistently throughout studies in social psychology (Byrne, 1971; Collisson & Howell, 2014; Hampton, Fisher Boyd, & Sprecher, 2019; Montoya & Horton, 2012).

Importantly, it has been found that not only does similarity breed liking, but that also liking breeds similarity (Alves, Koch, & Unkelbach, 2016; Collisson & Howell, 2014). From this perspective, people tend to perceive others that they like as more similar to themselves than others whom they dislike (Collisson & Howell, 2014). This can be explained when looking at the differences between positive and negative perceptions (Alves et al., 2016), where studies have shown that positive perceptions are less diverse than negative perceptions and that positive qualities are often defined by the lack of certain negative qualities (Alves et al., 2016). Thus, while there are limited ways a person can be positively perceived, there are many and more diverse ways individuals can be perceived negatively (Alves et al., 2016). Therefore, while people can dislike others for a variety of reasons, the reasons as to why individuals like others should be similar (Alves et al., 2016).

In conclusion, people like and are attracted to others who are similar, rather than dissimilar, to themselves, and at the same time people are more likely to perceive others they like as similar to themselves. A common and related theory to describe this bilateral link between similarity and likability is the *Balance Theory*, which will be introduced in the following section (Collisson & Howell, 2014).

Balance Theory

Introduced by Fritz Heider in 1958, the balance theory is based on the *naive theory of action*, which is a conceptual framework that is used to understand, interpret, and predict the behavior of others, and is centrally focused on intentional concepts, such as beliefs and desires among other things (Khanafiah & Situngkir, 2004). The balance theory conveys that people regulate their attitudes in order to achieve or maintain *cognitive consistency*, a balance that is reached when individuals do not perceive stress or pressure to realign their attitudes (Collisson & Howell, 2014). Therefore, according to the balance theory, when an individual discovers that another individual holds the same opinion regarding a topic, they will naturally like the other person in order to create and maintain cognitive balance (Collisson & Howell, 2014). Contrarily, if an individual discovers that another's opinion regarding a specific topic conflicts with their own, they will dislike that person to create and maintain their cognitive balance (Collisson & Howell, 2014). To achieve cognitive balance, people thus like those who are similar and dislike those who are dissimilar to themselves (Collisson & Howell, 2014).

In addition to that, the balance theory also conveys that a state of cognitive balance can equally be achieved by certain inferences made toward likable or dislikable individuals (Collisson & Howell, 2014). In other words, when people do not know a lot about a likable person, they still perceive them as similar in order to maintain cognitive balance (Collisson & Howell, 2014). Vice versa, when people do not know a lot about a dislikable person, they automatically perceive them as dissimilar in order to facilitate the state of balance (Collisson & Howell, 2014). Through this theory, the similarity-likability theory can be explained further and its effect rationalized.

Mediators of the Similarity-Likability Theory

Previous studies have identified and presented several mediators to further elucidate the similarity-likability theory (Hampton et al., 2019; Collisson & Howell, 2014), a number of which will be discussed in the following.

Cognitive Evaluation

One such mediator is *cognitive evaluation*, which, given that people assume information about others beyond what they already know and given that they also regard themselves positively, further explores the concept that individuals tend to infer positive attributes to and information about others they perceive as similar to themselves (Hampton et al., 2019). Both likability and cognitive evaluation are impacted by negative and positive information, and while likability is impacted more by negative information, cognitive evaluation is equally or even more so impacted by positive information as opposed to negative information (Singh, Ho, Tan, & Bell, 2007). In agreement with the theories about cognitive balance, cognitive evaluation studies have shown that when one perceives similarity in another, one associates other positive attributes with that person, which further prompts liking (Montoya & Horton, 2004; Hampton et al., 2019). Furthermore, research has also found that while attraction based on similarity is automatic, making a cognitive evaluation of another person's qualities before a heuristic evaluation can decrease or control the automatic nature of attraction (Singh et al., 2007), an approach tracing back to the dual or heuristic-versus-cognitive processing theory as studied by for example Evans (2008) and Chaiken & Manis (1980). Overall, increased cognitive evaluation has thus shown to mediate the rather impulsive effects of interpersonal attraction based on similar attitudes (Montoya & Horton, 2004; Singh et al., 2007).

Self-Concept Clarity

Generally, *self-concept clarity* represents how clearly defined, stable, and consistent one's perceived personal attributes are, and is considered an assessment of an individual's beliefs about themselves, that is, the concept of their self (Campbell et al., 1996). Importantly, self-concept clarity is not the same as self-knowledge, as clarity does not deal with the accuracy of the beliefs individuals hold about themselves (Campbell et al., 1996). It should also be noted that people with high self-concept clarity, meaning more stable self-beliefs, are ”less likely to change their self-descriptions over time or endorse mutually exclusive self-descriptive traits such as careless and careful” (Lewandowski, Nardone, & Raines, 2010, p.418). While this moderator has not been researched as a moderator specifically of the similarity-likability theory, it can naturally be assumed that people need to have a clear understanding of themselves and the beliefs they hold in order to project their own characteristics onto strangers (Collisson & Howell, 2014).

Self-Esteem

Connected to self-concept clarity, self-esteem constitutes another mediator of the similarity-likability theory (Collisson & Howell, 2014) and represents an individual's “beliefs about what they think they need to be or do to be a person of worth” (Park & Crocker, 2008, p.185).

Levels of self-esteem differ from person to person and can greatly affect the way they form relationships (Park & Crocker, 2008). In regards to balance theory, the level of one's self-esteem, that is, attitude towards oneself, has also been suggested to influence balance principles (Collisson & Howell, 2014). People who dislike themselves may reject positive others on the grounds of being too good for them and, according to balance theory, a person with negative self-perception does not achieve balance when matched with a positive other (Collisson & Howell, 2014). Low self-esteem may also affect people's beliefs about how they are perceived by others, and individuals with low self-esteem may assume that others dislike them since they dislike themselves (Collisson & Howell, 2014). This phenomenon is connected to the certainty of being liked, which is another moderator of the similarity-likability theory.

The Certainty of Being Liked

In particular, the certainty of being liked mediates the similarity-likability theory because when people know they are liked by their opposite, they are more likely to reciprocate this liking (Hampton et al., 2019; Sprecher, Treger, Hilaire, Fisher Boyd, & Hatfield, 2013). Tracing back to the very fundamental, evolutionary motives of humankind (e.g. Durante & Griskevicius, 2016), people inherently want to belong, so they are inclined to seek the company of those around whom they feel loved, valued, and cared for (Hampton et al., 2019). Moreover, research has found that people seem to be inherently aware of the effects of the similarity-likability theory and therefore believe that they will be more popular among those perceived as similar to them (Hampton et al., 2019; Sprecher et al., 2013).

Consensual Validation

Consensual validation is generally defined as the agreement of two or more perspectives in regards to a specific topic or event (Vogt, 2005). According to consensual validation theory, it is important for individuals to be able to foresee situations within one's environment and to be confident that one is right (Sigall, 1970). Thus, when people learn that others share their opinion, they feel rewarded, because their need for consensual validation is achieved (Sigall, 1970). Throughout their lives, people look to others in order to validate their beliefs and generally feel an inherent desire to connect with others through a mutually shared worldview (Sprecher et al., 2013; Hampton et al., 2019). Since it feels much more rewarding for one's beliefs to be validated as opposed to invalidated or contradicted, people often strategically look to others who they perceive as similar and thus expect to share similar views that ultimately help them achieve cognitive balance (Hampton et al., 2019). Thus, when two people are similar, especially in regards to attitudes, consensual validation leading to positive affect and liking are elicited (Hampton et al., 2019).

Preference for Consistency

While the way some people think and subsequently act may not always be aligned (Collisson & Howell, 2014), most people inherently attempt to think and behave consistently in an effort to reduce cognitive dissonances causing stress, tension, and discomfort (Collisson & Howell, 2014). This type of behavior is also called *preference for consistency* (Cialdini, Trost, Newsom, & Geen, 1995). Not only are people who have a stronger preference for consistency more likely to strive for cognitive balance, but they are also more likely to make inferences about other individuals out of a motivation to actively achieve balance (Collisson & Howell, 2014). As a result, differences in individuals' preference for consistency frequently predict behaviors and phenomena linked to cognitive balance, such as the similarity-likability theory (Cialdini et al., 1995). Therefore, on the one hand, people who value consistency more are also more likely to exhibit the similarity-likability effect, and on the other hand, people with a weaker preference for consistency will most likely not play into this theory (Collisson & Howell, 2014; Cialdini et al., 1995).

The Expectation of Enjoyable Interactions

Lastly, another mediator of the similarity-likability theory is the expectation of enjoyable interactions. When people perceive similarity in others, this often leads them to expect that their interactions with these individuals are more pleasant than with those perceived as dissimilar (Hampton et al., 2019). Often referred to as *rewards of interaction*, studies have shown that people like others who are enjoyable or fun to interact with or who they expect to have enjoyable interactions with (Hampton et al., 2019; Sprecher et al., 2013). Additionally, previous studies have also found that pleasant interactions with others frequently contribute to the formation of new relationships and liking (Hampton et al., 2019).

All in all, the similarity-likability theory has garnered longstanding and significant support both in theory and practice (Sprecher et al., 2013), and not only accounts for real-life interactions but also the increasing interactions conducted online (Jin et al., 2019; Abidin & Ots, 2016; Argyris et al., 2020; Breves et al., 2019; Campbell & Farrell, 2020; Folkvord et al., 2020). Therefore, it enables this study to explain the parasocial interactions between users and influencers, whose underlying philosophy and business model is based on being relatable and similar to ordinary users. Conflictingly, however, this perspective does not provide an explanation for user interaction with influencers who clearly distinguish themselves by portraying seemingly unattainable luxury lifestyles through their social media profiles, which the following section will therefore attempt based on the concept of self-expansion.

The Self-Expansion Theory

While the similarity-likability theory is "one of the most well-established findings in the study of interpersonal attraction" (Aron, Steele, Kashdan, & Perez, 2006, p.387), there is a body of research that has discovered certain conditions where similarity has less of an effect, no effect, or even a negative effect on likability (Aron et al., 2006; Aronson & Worchel, 1966; Jones, Bell, & Aronson, 1972; Izard & Smith, 1960; Goldstein & Rosenfeld, 1969; Nahemow & Lawton, 1975). For example, preceding studies have found that individuals who formed relationships based on who they frequently interacted with were more likely to have friends with dissimilar characteristics, such as age and race (Nahemow & Lawton, 1975; Aron et al., 2006). Researchers have also discovered that similarity is less important for people who were found to have lower fear of rejection or need for approval than others (Aron et al., 2006; Goldstein & Rosenfeld, 1969). Yet again, other studies have found that the effect of similarity is dependent on an individual's maturity, which is considered to play a significant role in a person's ability to build and maintain relationships (Aron et al., 2006; Izard & Smith, 1960). Therefore, younger and typically less mature individuals are more likely to seek relationships with those who are similar to themselves as opposed to their older and more mature counterparts (Izard & Smith, 1960).

The main perspective that suggests that the similarity-likability theory is not important under certain circumstances is the *self-expansion theory*, according to which individuals seek to ameliorate themselves and their potential efficacy (Aron et al., 2006). In order to do so, relationships with others who enable the acquisition of beneficial social or physical assets, perspectives, or knowledge are formed (Hampton et al., 2019). In other words, self-expansion involves developing close, personal relationships aimed at growing, learning, and expanding the self, which has been found to lead to the other person's resources, mindset, and identity being partially internalized (Aron et al., 2006). Therefore, according to the self-expansion theory, individuals essentially tend to be attracted to others who are considered beneficial for expanding the self (Aron et al., 2006).

Consequently, the self-expansion theory suggests that people have a tendency to like those who are *dis*similar to themselves and by that provide room for uni- or bilateral personal growth, which directly opposes the similarity-likability theory (Hampton et al., 2019; Aron et al., 2006). From a social identity theory perspective, this would mean striving for inter-group rather than intra-group relationships, expanding the horizon of both the self and one's in-group alike (Ketay, Beck, & Welker, 2020). Forming relationships with those who are similar would lead to fewer opportunities to expand the self, and differences are therefore considered a desirable base for building relationships, providing new perspectives and knowledge that may not be exhibited by similar others (Hampton et al., 2019; Aron et al., 2006).
In this context, existing studies have particularly looked at dissimilarity in regards to interests, since differing interests entail rich self-expansion possibilities and simultaneously increase one's own potential of fostering expansion among others later on (Aron et al., 2006). Especially in close relationships, individuals with different interests are thus likely to experience novelty and overcome challenges together while self-expanding, reinforcing their bonding and relationship depth, and thus making dissimilarity attractive (Aron et al., 2006).

Furthermore, previous studies have found that the effect of similarity on liking was not significant when one believes that another likes them (Aronson & Worchel, 1966) and that there exists a tendency to reciprocate affection to attitudinally dissimilar, as opposed to attitudinally similar, individuals (Jones et al., 1972). Other studies have moreover discovered that when given the option between a person who was likely to form a relationship with a participant and a person who shared similar interests with the participant, participants chose the person who was likely to form a relationship (Aron et al., 2006). This may be because the formation of a relationship was probable, which rendered additional reasons to predict if a relationship was possible, such as similarity, irrelevant, and led to the consideration of other factors, such as how another would contribute to the participant's self-expansion (Aron et al., 2006).

Despite these findings, there also exist several studies that support self-expansion as a mediator of the similarity-likability theory (Hampton et al., 2019; Sprecher, Treger, Fisher Boyd, Hilaire, & Grzybowski, 2015; Aron et al., 2006), as it has been found that similarity may also generate possibilities for the expansion of the self due to a higher overlap among two individuals which may, for example, result in mutual inspiration to grow (Hampton et al., 2019). Moreover, studies have also found positive correlations between similarity, self-expansion opportunities, and likability (Sprecher et al., 2015) which positively affected the establishment of friendships, leading to similarity often being used as a predictor of how much potential a particular relationship has to form and prosper in the long-term (Aron et al., 2006). Thus, in some cases, the self-expansion theory can also support the similarity-likability theory, since, in order to self-expand, a relationship must first be formed, which is empirically more probable with others who are similar as opposed to dissimilar (Aron et al., 2006). In general, no two individuals are completely alike, meaning that everyone differs at least slightly if not more, which means that already the act of forming a relationship with another individual entails substantial potential to expand the self (Aron et al., 2006).

In the context of this research, the self-expansion theory will be considered in support of dissimilarity leading to attraction, while the similarity-likability theory will serve as the primary theory used to deduct hypotheses and assumptions from.

Theory Application to the Context of Instagram Influencers

Summing up, user interactions with influencers are generally classified as parasocial relationships, and there are various possible explanations as to why users are motivated to form such unilateral, nonreciprocal relationships and interact with social media personalities.

First and foremost, the similarity-likability theory is expected to provide a good explanation to this question because perceived similarity is considered the foundation of an influencer’s business model, and studies have shown that relatability translates into more and more positive follower engagement (Campbell & Farrell, 2020; Van Driel & Dumitrica, 2021). The similarity-likability theory can be traced back to the study of social identity theory about intra- and inter-group relations based on similarities and differences, from which strong in-group preferences were found to emerge (Tajfel, 1972). Moreover, it has been found that not only does similarity breed liking, but that liking also breeds similarity, so those thought of as likable are also naturally perceived as similar to the self in comparison to less liked others (Alves et al., 2016). This bilateral relationship between similarity and likability may be explained by the balance theory, which argues that individuals strive for cognitive consistency and regulate their attitudes based on this inherent desire. For example, sharing an opinion about a topic creates cognitive balance and therefore liking, so those with similar views become more likable.

Nevertheless, the question of whether users interact with influencers on the verge of being a celebrity or those who portray seemingly unattainable luxury lifestyles on social media remains uncovered by the similarity-likability theory. This however can be explained by the self-expansion theory, which revolves around the idea that in order to grow themselves, individuals seek relationships with others who can foster the acquisition of desirable characteristics, and who are therefore dissimilar from themselves.

In the case of Instagram, the app's main user base is rather young, with more than 70% under the age of 35 (Tankovska, 2021b), and the platform's super users are mostly teenagers and young adults (Suciu, 2019). With this demographic in mind, it can be assumed that the need for validation and fear of rejection, which have been found to be disproportional to maturity and peak during teenage years (Izard & Smith, 1960; Burrow & Rainone, 2017), is overall high on Instagram regarding both followers as well as influencers. Thus, especially young Instagram users can be assumed to fulfill the characteristics supporting the similarity-likability theory, meaning that they are expected to turn to others perceived as similar in order to seek meaningful relationships, where they may feel like part of an in-group. This could be an influencer of the same age, origin, look, culture, or with the same interests, but also this influencer's community. In that case, the parasocial relationship between follower and influencer may even be expanded onto other members of the influencer's community with which users may connect over comments and content, making the relationship less unilateral and expanding the pool of similar others to enter into online relationships with. Thus, further similarity, likability, enjoyable interactions, and validation may be found within the community tied to an influencer, which is likely to increase the cognitive balance of Instagram users even further. Moreover, a follower's preference to be exposed to content that one is likely to find interesting based on similarity of interests, both in an unsponsored and influencer marketing context, may be another simple yet strong supporting argument of the similarity-likability theory in the specific context of Instagram.

Simultaneously, these arguments may also represent reasons for influencers to foster the depth and strength of their relationships with their followers. Influencers too may seek validation, possibly even more so than their followers, and find value in a group of like-minded followers with whom they can share and even monetize their passion with. Consequently, the more an influencer perceives themselves as similar to their community of followers, the more this can also be expected to be reflected in their behavior on social media. Influencers that perceive their audience as a tight-knit group of peers may therefore engage more with their followers by sharing more personal information and trying to create likable content.

On the opposite, it could however also be argued that young adults on Instagram are looking for inspiration on personal growth and self-expansion as they transition into adulthood, and therefore turn towards those they can look up to, who may possess desirable characteristics or resources, and who may introduce them to new interests and undiscovered content. Simultaneously, influencers may be more motivated to engage in creating content for those they do not perceive as similar and whom they have little knowledge of.

In order to explore the influences of and motivation behind user-influencer engagement on social media, this study looks at what influencers disclose in their social media content and how much this content is interacted with, from a similarity-likability perspective. In order to do so, this study breaks down influencer disclosure into the following categories: information disclosure, personality disclosure, and commercial disclosure, which will be elaborated on in the subsequent sections. This study also explores how influencer size impacts not only user engagement but also influencer disclosure, as well as the relationship between influencer disclosure and user engagement. As a result, the following research model was constructed, aiming to measure the effect of similarity and likability on user engagement by factoring in influencer disclosure and influencer size.

Introduction of the Research Model

With the overall goal of exploring what influences follower engagement on influencer social media posts through the lens of the similarity-likability theory, *Influencer Disclosure* and *Influencer Size* are chosen as main factors assessed in this context. Influencer size refers to how large (or small) the influencer's following is, while influencer disclosure refers to the extent to which influencers divulge various attributes along with their social media posts. Essentially, influencer disclosure is broken down into three sub-categories: *Information Disclosure, Personality Disclosure,* and *Commercial Disclosure,* which, like influencer size, consist of various quantifiable measures used to conduct a quantitative analysis, which will be elaborated on in the subsequent section.

Within the study's main model hypotheses, H1 to H9, influencer disclosure and influencer size are examined regarding how they directly influence user engagement, which is measured through like rate, comment rate, and overall engagement rate (combining likes and comments), of Instagram users on influencer posts. Applying the similarity-likability theory to this context, H1-H8 of this study argue that the various influencer disclosure variables in a post determine the perceived similarity between ordinary users and influencers, and thus an influencer's likability. This is expected to lead to changes in engagement, where higher likability logically implies more likes and comments, which are a means to interact with and support others on Instagram, actual and parasocial friends alike. Simultaneously, H9 argues that influencer size also has a direct effect on similarity, likability, and consequently, engagement rates.

In line with the similarity-likability theory, sub-model hypotheses H1a-H8a simultaneously argue that the number of followers influencers have also bilaterally affects the amount of information disclosed by the influencers themselves.

Lastly, given the expected effects of influencer disclosure on engagement rates and influencer size on influencer disclosure, this research also assumes that and sets out to investigate if and how these effects interact with one another, that is, how influencer size impacts the relationship between influencer disclosure and user engagement. Thus, it is not only expected that influencer size directly impacts engagement rates due to perceived similarity and thus likability, but this paper also presents the argument that influencer size will affect the degree in which influencer disclosure impacts user engagement. This research dimension is represented within hypotheses H1b-H8b and will also be expanded on in the following sections.

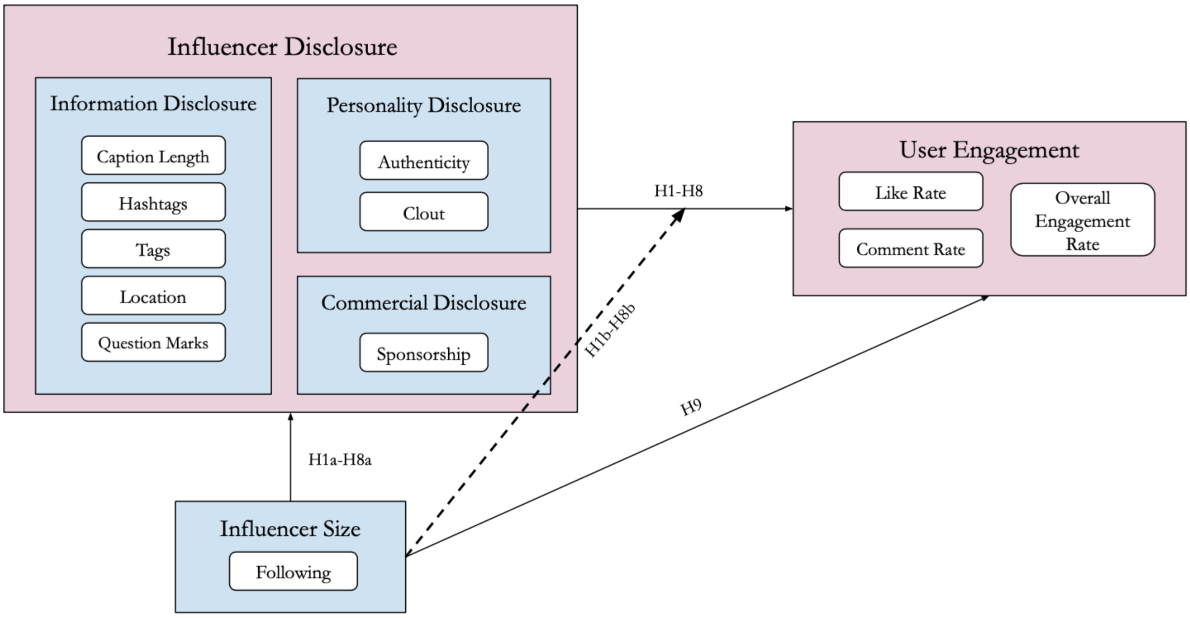

Figure 3: *Research Model*

Summing up, this study's underlying research model as presented in Figure 3 assumes an effect of similarity and likability on user engagement, based on influencer disclosure and influencer size, where users are expected to engage more with content that creates the perception that the influencer is either likable or similar. Additionally and as a result of these expected effects, the research model further presents the assumption that influencer size will impact influencer disclosure as well as the main relationship between influencer disclosure and user engagement.

Hypotheses

In the following sections, the four major explanatory categories, that is, information disclosure, personality disclosure, commercial disclosure, and influencer size, will be analyzed in more detail and explored in the context of the similarity-likability theory. Essentially, these four categories were chosen based on previous research findings suggesting that they are positively correlated with social media engagement (Jin et al., 2019; Abidin & Ots, 2016; Argyris et al., 2020; Jun & Yi, 2020; Breves et al., 2019; Campbell & Farrell, 2020; Folkvord et al., 2020). Aiming to expand previous findings, this paper argues that the components of influencer disclosure and influencer size relate to the *similarity breeds likability* and/or *likability breeds similarity* aspect of the similarity-likability theory. These assumptions will be explored in the following.

Influencer Disclosure Hypotheses

First, the main set of hypotheses looks at how influencer disclosure, consisting of information-, personality-, and commercial disclosure, as well as influencer size impact engagement rates, that is, the number of likes and comments, as well as the sum of the two, per follower on a post.

Information Disclosure

In this context, information disclosure refers to the amount and type of information communicated by influencers through their social media posts and represents how they interact with their community. It is suggested that the way in which influencers communicate matters and causes them to develop stronger connections and trust with their audience, as social media and especially visual platforms like Instagram - in contrast to traditional media channels such as television or print magazines - allow users greater insight into the lives of influencers who embrace sharing personal details and directly engage and interact with their followers through social media (Jin et al. 2019; Breves et al., 2019; Van Driel & Dumitrica, 2021).

As previously discussed, this amount of information disclosure and interactivity distinguishes influencers from traditional celebrities and makes their role unique (Jun & Yi, 2020). In this context, interactivity essentially refers to all active communication between influencers and their followings, which is possible - and the norm - 24/7 nowadays (Jun & Yi, 2020). Influencers that actively share relevant, engaging, personal content and engage their audience in ongoing conversations have been found to garner more active followers, as this may make both influencer and follower less anonymous and more present in their parasocial relationship (Jun & Yi, 2020). As a result, disclosing more information and sharing more with one's audience has been found to increase intimacy, strengthen loyalty, and encourage emotional attachment (Jun & Yi, 2020).

On the one hand, according to the similarity-likability theory, this paper, therefore, argues that influencers who disclose more information and engage more in conversations with their following will be perceived as more likable and/or relatable, which is expected to result in higher engagement rates.

On the other hand, according to the self-expansion theory, users would be more likely to interact with influencers who are dissimilar to themselves, that is, not relatable, in order to maximize personal growth. Therefore, this paper presents as a counter logic that influencers who disclose less information with their following would be perceived as more distant and less relatable, which would result in higher engagement rates.

In order to quantify information disclosure, the following measures were selected:

Caption Length

Captions are the texts that accompany visuals posted on social media and enable users to share additional content along with their posts, such as visual descriptions, personal details, related stories, or generally any kind of information they wish to share (see Figure 4 for explanatory post). Especially on visual-based platforms like Instagram, captions add an important, second layer to posts and enable influencers to be more interactive by sharing more personal information. Longer captions are hypothesized to grow the follower's perception of relatability and are thus expected to increase engagement according to the similarity-likability theory. Therefore, this paper puts forth the following hypothesis:

➢ **H1.** Posts with longer captions will have higher engagement rates.

While longer captions may contain a variety of information, posts with shorter captions, sometimes only a few words or emojis, share limited context with an audience, generating less interactivity and eliciting less relatability to the influencer posting them. As a result, shorter captions may be perceived as distant and secretive, and thus unfamiliar and dissimilar, which would increase engagement according to the self-expansion theory.

Hashtags

Hashtags are words or phrases following the '#' symbol and are used to connect and gather content about themes, topics, conversations, events, and more (Olafson, 2020). Users can thus easily explore related content by clicking any hashtag included in the caption or comments of a social media post (see Figure 4). Due to the connective nature of hashtags, posts using them are more visible and interactive in nature, as they invite others who are interested in the same topic. As a result, influencers that share posts with hashtags profit from disclosing more information, which increases their perceived relatability and leads to more engagement according to the similarity-likability theory. Therefore, this paper hypothesizes that:

➢ **H2.1.** Posts that use hashtags will have higher engagement rates.
➢ **H2.2.** Posts that use more hashtags will have higher engagement rates.

Contrarily, posts that do not use hashtags are less visible, less interactive, and disclosing less information, so influencers that share posts without hashtags do not attempt to connect with like-minded others, making them less relatable, or dissimilar, which would lead to more engagement according to the self-expansion theory.

Tags

Social media *tags* consist of the '@' symbol, followed by a user's or business's Instagram handle. Tags are clickable and included within post captions in order to connect a post to a certain account (see Figure 4). Like hashtags, tags are also connective in nature, and therefore, posts that use tags are more interactive, as they share additional information on the people or businesses associated with a post. From the similarity-likability theory, influencers that share posts with tags disclose more information and will therefore be perceived as more relatable, leading to increased likability and more engagement. Thus, this paper presents the following hypotheses:

➢ **H3.1.** Posts that use tags will have higher engagement rates.
➢ **H3.2.** Posts that use more tags will have higher engagement rates.

According to the self-expansion theory, due to the increased information disclosure of posts with tags, influencers would be perceived as too similar and not sufficiently distinctive enough to inspire self-expansion, which would consequently lead to lower engagement rates.

Figure 4: *Examplaratory Instagram Post Showing Caption, Hashtags, Tags, and Location*

Description. Fashion influencer, *@sincerelyjules,* posting about a recent designer collection, using hashtag *#DiorAW21* and tag *@dior* in the caption (highlighted in pink). Furthermore, she uses a location tag (highlighted in yellow), and actively engages with her following through asking a question.

Source: Sincerelyjules (2021)

Location

Similar to tags, a location can be added to a social media post, where it is seen at the top of the post, right below the username (see Figure 4). When a user searches for a specific location, all posts that have tagged this specific location will appear in the search results. By sharing a location, influencers are opening themselves and their content up to more users interested in the location, and also provide their followers with more personal details about their whereabouts, making their posts more inviting for conversations and thus more interactive. This means that, conforming to the similarity-likability theory, influencers that share posts with a location tag disclose more information and are deemed more relatable, which increases likability and likely elicits more follower engagement. In keeping with the former theory, this paper thus hypothesizes that:

➢ **H4.** Posts with a location will have higher engagement rates.

In consonance with the self-expansion theory, influencers that share posts with location tags would be perceived as too similar in comparison to ordinary users and not sufficiently distinctive, possibly failing to provide inspiration for self-expansion while becoming too approachable by giving away too much personal information, which would be expected to decrease follower engagement.

Question Marks

As mentioned previously, it is common for influencers to communicate directly with their followers through their content. This often means that, within captions, influencers ask questions to their audience in order to engage their following and start a conversation in the comments (see Figure 4). Therefore, posts with questions in their caption are more interactive, and, in accordance with the similarity-likability theory, influencers that share posts with questions are perceived as more relatable, equal, and peer-like, which encourages engagement. Conforming to the main theory, this paper thus puts forth the following hypothesis:

➢ **H5.** Posts with questions will have higher engagement rates.

Contrarily, from the self-expansion perspective, influencers that share engaging posts using questions may be perceived as not dissimilar enough, as they would create the impression of friendship rather than aspiration, leading to lower expected engagement.

Personality Disclosure

Moving on to the next category, one of the main practices used by influencers to build connections with users is through exhibiting an endearing, authentic personality (Abidin & Ots, 2016). Being an influencer first and foremost means curating a persona that comes across as warm, inviting, and authentic, through which influencers are able “to sustain their accessibility, believability, emulate-ability, and intimacy - in other words, their ‘relatability’ ” (Abidin & Ots, 2016, p.155). Preceding studies have discovered that in order to cultivate an authentic personality, influencers should try to make an impression of similarity as opposed to trying to stand out and create a distance (Argyris et al., 2020). Users’ preference towards authentic personalities has not only been found to establish intimacy and relatability (Argyris et al., 2020), but exposing a more authentic personality has also led to stronger emotional connections and trust between influencers and their following (Jun & Yi, 2020). In conclusion, the more influencers disclose their authentic personalities, the more likable they become. According to the similarity-likability theory, this leads to higher levels of relatability, which therefore is assumed to translate into higher engagement.

Contrarily, from the self-expansion theory, users would be more likely to interact with influencers who are dissimilar and less relatable to themselves in order to maximize self-expansion. Therefore, this paper presents a counter logic that influencers who disclose their personality will be perceived as too relatable and therefore insufficient for expanding the self, which would be reflected in lower engagement rates.

In order to quantify personality disclosure, the following measures were chosen:

Authenticity Scores

An authenticity score is a percentile given to a piece of text, which ranges from 0 to 100 (LIWC, n.d.a). This score refers to “when people reveal themselves in an authentic or honest way, [where] they are more personal, humble, and vulnerable” (LIWC, n.d.a, “Authenticity”). Authenticity scores of post captions were analyzed using the LIWC (Linguistic Inquiry and Word Count) software, which will be elaborated on in the Methodology section. In this study, authenticity scores indicate the personality disclosure of individual posts made by influencers, whereas higher authenticity scores are considered to disclose more of the influencers’ personality and therefore make them more likable, which is expected to positively influence perceived relatability and engagement according to the similarity-likability theory. In line with the main theory, this study presents the hypothesis that:

➢ **H6.** Posts with more authentic captions will have higher engagement rates.

In contrast, according to the self-expansion theory, high authenticity scores, leading to greater perceived relatability, would negatively impact engagement.

Clout Scores

Similar to the authenticity score, a clout score is a percentile given to a piece of text, which was also established using LIWC software (LIWC, n.d.a). This score refers to “the relative social status, confidence, or leadership that people display through their writing or talking” (LIWC, n.d.a, “Clout”). While clout represents another part of an influencer’s disclosed personality and thus the amount of personality disclosed, it is important to note that its original meaning is “power and influence over other people or events" (Cambridge Dictionary, n.d., “Clout”), and that influencers with high clout in their captions may be perceived as bragging and pretentious. Thus, in this study, clout scores represent the opposite of relatability and authenticity. From the similarity-likability perspective, influencers that share posts with *low* clout scores will be associated with greater authenticity and likability, and thus, increased perceived relatability and engagement is expected. In keeping with the similarity-likability theory, this study hypothesizes that:

➢ **H7.** Posts with less clout in their captions will have higher engagement rates.

According to the self-expansion theory, influencers that share posts with high clout scores would be less likely to be seen as real and relatable, which would increase engagement.

Commercial Disclosure

The next category, commercial disclosure, particularly comes into play when influencers become endorsers (Campbell & Farrell, 2020), and is based on the fact that studies have found that the way social media users process content differs depending on whether a commercial nature is disclosed or not (Evans et al., 2018). Importantly, previous research has shown that commercial intent on social media posts leads to a significant negative impact on “brand attitude, purchase intention, and intention to spread eW[ord]O[f]M[outh]” (Evans, Phua, Lim, & Jun, 2017, p.139).

Therefore, while users often look to influencers and their expertise when they are considering making purchases (Campbell & Farrell, 2020), influencers simultaneously need to be prudent about the type and amount of commercial partnerships within their feed, since both the wrong type of partnerships as well as a high frequency of commercial posts may lead to criticism, sellout-accusations, and may cause followers to abandon the influencer due to a lack of trust and realness (Argyris et al., 2020). Thus, influencers often face the dilemma of generating income versus hurting the relationship with their followers.

Generally, sponsorships are a key characteristic of social media influencers that may not only affect the follower-influencer relationship, but first and foremost separates influencers from ordinary social media users. Therefore, this study argues that disclosing commercial intent and the overall commercial nature of a post differentiates influencers from their audience, drawing a line beyond which the influencer is perceived as a business instead of a fellow consumer, and significantly reducing relatability, which in turn is expected to negatively impact likability and user engagement.

However, according to the self-expansion theory, users would be more likely to interact with influencers who provide opportunities for self-expansion and are an inspiration to their followers, which would be the case with frequent commercial posts. Therefore, this paper presents the counter logic that influencers who post and disclose commercial content will be perceived as dissimilar, showcasing the characteristics their followers would like to see to inspire self-expansion, which would result in higher engagement rates.

In order to quantify commercial disclosure, the following measure was chosen:

Sponsorship

Sponsored posts are posts that are made by an influencer in partnership with and in order to promote a brand and its products or services in exchange for compensation (Mudambi & Schuff, 2010). As previously discussed, influencer marketing can take many forms but is required by law to be marked as such, and previous studies have found that commercial content negatively affects the way social media users feel and the extent to which these users will engage (Evans et al., 2018). Due to this significant impact of commercial content on user engagement and attitudes, it is argued that influencers who share posts that are *not* commercially sponsored will be deemed more relatable and likable, which, according to this study’s central theory, will increase engagement. Therefore, this study hypothesizes that:

➢ **H8.** Unsponsored posts will have higher engagement rates.

From the self-expansion perspective, commercially sponsored content would signal success to inspiration-seeking followers looking to expand their horizons beyond brands and products they already know, which is why higher engagement rates on sponsored posts would be expected.

Influencer Size

In regards to the last category, influencer size, preceding studies have found that perceived similarity directly impacts the follower-influencer relationship (Jin et al., 2019; Abidin & Ots, 2016; Argyris et al., 2020; Campbell & Farrell; Folkvord et al., 2020). For example, social media influencers are perceived as more trustworthy than traditional celebrities due to the fact that they are perceived more like ordinary users (Jin et al., 2019). Moreover, research has shown that users respond positively to other users that share similar demographics, interests, or other characteristics (Jin et al., 2019; Argyris et al., 2020). In accordance with the similarity-likability theory, this similarity would breed liking, which would encourage follower engagement.

Contraversely, from the self-expansion perspective, influencers with large followings would be associated with dissimilarity and thus present greater opportunities for self-expansion, which in turn would lead to greater engagement.

In order to quantify influencer size, the following measure was chosen:

Following Size

Following size is essentially measured in the number of followers any profile has garnered. As mentioned previously, following size can vary greatly among influencers. Since the average, ordinary Instagram user has around 150 followers (Erin, 2020), this paper argues that influencers with a smaller following will be perceived as more peer-like and thus more similar than influencers with massive followings, creating a sense of intimacy, friendship, and shared perspective. Therefore, this study proposes the following hypothesis:

- ➢ **H9.** Posts from profiles with a smaller following will have higher engagement rates.

According to the self-expansion theory, this similarity would discourage user engagement, whereas according to the similarity-likability theory, smaller influencers are expected to record higher engagement rates.

Moving on from the main set of hypotheses, following size furthermore forms the basis of the next set of hypotheses.

Influencer Size Hypotheses

Having elaborated on the expected effects of the various influencer disclosure measures on engagement rates in the previous section, the following segment will focus on the effect that following size, in line with the similarity-likability theory, is expected to have on the influencer disclosure variables themselves.

Information Disclosure

While H9 argues that influencers with fewer followers will record higher engagement rates, based on the assumption that following size relates to similarity and influencers with a smaller following are perceived as more similar, peer-like, and likable, this study simultaneously argues that following size affects the way in which influencers choose to disclose information with their following in the first place. Thus, the feeling of similarity and relatability between influencer and follower is expected to be bilateral and reflected not only in follower engagement but also in the way that influencers themselves behave and interact. From an influencer perspective, fewer followers are expected to create a feeling of similarity, intimacy, friendship, and a tight-knit community, which is why this study, with the similarity-likability theory in mind, also hypothesizes that:

- ➢ **H1a.** Influencers with a smaller following will share posts with longer captions.
- ➢ **H2a.** Influencers with a smaller following will share posts with more hashtags.
- ➢ **H3a.** Influencers with a smaller following will share posts with more tags.
- ➢ **H4a.** Influencers with a smaller following will share posts with location more often.
- ➢ **H5a.** Influencers with a smaller following will share posts with questions more often.

From the contrasting, self-expansion theory perspective, following size would be expected to have the opposite impact, and smaller influencers would feel too similar to their followers if they shared a lot of information and interacted with them through engaging conversations to perceive themselves and construct their personal brands as objects of inspiration and self-expansion.

Personality Disclosure

While hypotheses H6 and H7 argue that the more influencers disclose their authentic personality the more likable they are, hypotheses H6a and H7a look at these propositions from an influencer perspective. Essentially, it is hypothesized that following size affects the perception of relatability and likability, leading to higher perceptions of intimacy and comfort in disclosing information and showing one's true, everyday personality with a smaller community of followers.

Therefore, this study presents the following hypotheses:

➢ **H6a.** Influencers with a smaller following will share posts with more authentic captions.
➢ **H7a.** Influencers with a smaller following will share posts with less clout in their captions.

Contrarily, from a self-expansion theory perspective, the opposite effect would be expected, that is, influencers attempting to appear less authentic and displaying more clout in order to be perceived as different and as an inspiration for self-expansion.

Commercial Disclosure

As hypothesis H8 argues that sponsorships are a key factor differentiating influencers from ordinary users and that followers will therefore relate to and engage significantly less with sponsored posts, the effect of following size on sponsorship is equally expected to entail a negative correlation. Based on the similarity-likability theory, it is assumed that influencers who consider themselves as more similar to their community will also choose a more relatable habit of minimizing the amount of commercial content they expose their audience to. Therefore, this study hypothesizes that:

➢ **H8a.** Influencers with a smaller following will share fewer sponsored posts.

From a self-expansion perspective, smaller influencers would be expected to share more sponsored content in order to differentiate themselves from their followers and thus potentially become an object of inspiration to others.

Interaction Hypotheses

Lastly, having elaborated on the expected relationship between influencer disclosure and engagement rates as well as the relationship between influencer size and influencer disclosure,

this research also expects and sets out to examine interaction effects between these two sets of hypotheses. As visualized in the research model (see Figure 3), it can logically be derived from the previously stated assumptions that influencer size not only affects engagement rates and disclosure individually, but that it also impacts the way in which the two are related. Thus, it is assumed that the number of followers acts as a mediator and that the effects examined in hypotheses H1 to H8 may be subject to changes contingent on the underlying follower number of an influencer.
Therefore, the following hypotheses are proposed:

- **H1b.** Influencer size impacts the relationship between caption length and engagement rates.
- **H2b.** Influencer size impacts the relationship between hashtag usage and engagement rates.
- **H3b.** Influencer size impacts the relationship between tag usage and engagement rates.
- **H4b.** Influencer size impacts the relationship between location usage and engagement rates.
- **H5b.** Influencer size impacts the relationship between question asking and engagement rates.
- **H6b.** Influencer size impacts the relationship between caption authenticity and engagement rates.
- **H7b.** Influencer size impacts the relationship between caption clout and engagement rates.
- **H8b.** Influencer size impacts the relationship between post sponsorship and engagement rates.

Methodology

With the underlying sets of hypotheses in mind as laid out and visualized in the previous section, this section will proceed to discuss the underlying methodology used to obtain the research results which will subsequently be presented and interpreted.

Research Paradigm, Logic, & Strategy

Premised on the existence of established behavior patterns within the context of user-influencer interaction, this study first and foremost followed a mostly positivist research philosophy (Schrag, 1992), with the overall goal of testing the existing theory that was presented in the previous section. Hereby, it was the study's objective to scientifically examine and subsequently generalize the causalities of the underlying phenomenon of user-influencer interaction on Instagram. Assuming that both general as well as platform-specific culture influences user behavior, the study set out to gain empathetic, representative in-depth insights expressed through real, cause-effect relationships between various variables. In doing so, the ultimate aim was to reduce complexity and contribute to research as well as practice by narrowing down explanations and predicting human behavior on trending social media platforms.

Therefore, this study followed a mostly deductive logic, in which hypotheses were based on existing theory and applied to a novel context. Subsequently, the underlying hypotheses were tested with the goal of confirming, rejecting, and expanding current knowledge. With this approach in mind, the data collection process and analysis steered towards evaluating the propositions introduced in the previous chapter and to ultimately deduct a refined concept about influencer interaction on Instagram.

In order to do so and in alignment with the research philosophy, a mono-method quantitative methodology was deemed suitable to conduct highly structured research, based on a large sample allowing for reliable and generalizable contributions.

In practice, this was realized through cross-sectional, non-experimental research, which was selected due to its unobtrusive, purely observational nature and the resulting lack of experimental influence on either of the variables, creating a natural, real-life foundation for the study (Thompson & Panacek, 2007; University of Minnesota, 2016). More specifically, this non-experimental research was conducted in the form of what may be called quantitative netnography. While a *netnography* is a new research method generally defined as ethnographic research to study cultures and communities emerging through computer-mediated communications (Kozinets, 2010; Bowler, 2010) that has become more relevant than ever, it has to date almost exclusively been defined and applied in a qualitative manner (Kozinets, 2006;

Heinonen & Medberg, 2018). This study, however, applied it in an exclusively quantitative manner in order to explore publicly observable data on the Internet that has the power to reveal findings of human behavior online.

Sampling Strategy & Data Collection

In order to answer the research question “What motivates social media users to interact with influencers, whose attention might be diffused and commercially motivated?”, primary data was gathered on Instagram, which was picked because the platform significantly coined if not launched the idea of influencers as they are known today, and still is the central environment of influencers and influencer marketing (Influencer Marketing Hub, 2021a).

Given the study’s mostly positivist research paradigm, the goal of the data collection process was to gather a sufficiently large amount of influencer profiles that were, to some extent, homogenous in order to increase the comparability and validity of the research. Thus, the niche of fashion was selected, since it was neither too narrow nor too general. Moreover, it was considered important that the selected profiles were native influencers, that is, no persons of public interest or tied to a community of followers not mainly garnered through Instagram in the first place. Since influencer size was presumed to be a key factor and thus to ensure valid and reliable results, the influencer sampling process was based on creating an equal representation of all five influencer categories presented in the Literature Review, that is, nano-, micro-, mid-tier-, macro-, and mega-influencers. Furthermore, since caption analysis was also identified as a focal point of interest, originating from or being based in an English-speaking country was added as another sampling criteria in order to enable a fruitful analysis. Therefore, a sufficiently large sample consisting of 250 fashion influencers from English-speaking countries worldwide was ultimately set out to gather, with 50 influencers representing each of the five influencer categories.

The applied sampling strategies included simple random sampling followed by snowball sampling. Initially, simple random sampling was chosen in order to reduce selection bias and because of its straightforward nature among the various probability sampling methods (Thompson, 2012). One issue faced during this initial stage of the sampling process was that, due to the existence of significantly more nano-, micro-, and mid-tier-influencers compared to macro- and mega-influencers, the vast majority of profiles resulting from simple random sampling belonged to the smaller influencer categories, whereas only a handful of large influencers were identified in this stage. Therefore, a subsequent, second sampling stage proceeded with snowball sampling, which is often used when recruiting subjects that are harder to gather than others (Thompson, 2012), and which eventually enabled the assembly of a sufficient sample of English-speaking mega- and macro-influencers from within the fashion niche. In the following paragraphs, the sampling methods used will be explained in greater detail.

In the first, simple random sampling stage, Instagram's search function was utilized in order to identify posts containing hashtags *#ad* and *#sponsored*, which are two of the most common indications used by influencers to mark sponsored content, which this study used to distinguish influencers from ordinary users. The search results were then manually scanned for posts and profiles meeting the previously defined criteria. This sampling method resulted in an initial list of approximately 60 randomly selected fashion influencers on Instagram.

The second sampling stage, snowball sampling, then commenced using Instagram's *Suggested Profiles* feature, where users can, when looking at a specific profile, receive suggestions of similar profiles compiled by the Instagram algorithm. This function was used on the profiles of previously selected influencers in order to expand the number of subjects within a particular category, especially for mega- and macro-influencers. In the end, 250 fashion influencers from English-speaking countries were identified, including 50 from each of the five influencer categories. The subsequent goal was then to gather and collect data from each of the selected influencer's 100 most recent posts.

In order to efficiently extract data from the chosen influencer profiles, third-party platform *picuki.com* was used due to the various barriers imposed by Instagram itself to prevent third parties from scraping larger amounts of data from the platform. By dynamically loading a profile's posts, picuki.com significantly facilitated the data gathering process and enabled the extraction of more than the last twelve posts, which is the limit imposed by Instagram in regards to immediately loading posts. This eventually allowed the export of source code JSON files containing the data of the 100 most recent posts for each of the 250 sampled influencers in late January 2020.

These JSON files were subsequently fed into a Java application, which was run on February 1st, 2021, and returned a .txt file for each individual influencer's profile. The data points collected were as follows. From a respective profile, the Java crawler gathered the influencer's Instagram handle, the profile's URL, the number of users that were following the profile, number of users the influencer was following, number of posts, and the profile's description. From each individual post, the program further gathered, among others: the number of likes and comments, the post caption, a list of any hashtags or tags found within the caption, the location included (if any), and whether it was tagged as commercial (true or false). In order to determine the latter, the application scanned each post caption for either of the following strings: “advertising”, “advertisement”, “partnership”, “collaboration”, “cooperation”, “sponsor”, “#ad”, “ambassador”, “% off”, “code”, “(ad)”, “gifted”, “collaboration”, and “giveaway”, as these terms were previously identified as reliable to indicate sponsored content. Hereby, the value *true* was returned by the application if any of the strings were found.

Data Analysis

In order to explore the collected data, the sample was then prepared for regressions within multiple steps and subsequently analyzed to test the research hypotheses.

Preparation of the Dataset

In order to prepare the data that was returned by the Java application for the subsequent statistical analyses, the text files were first converted into MS Excel and, after removing two faulty posts with null results, the remaining 24,998 observations were coded.

The coding process consisted of multiple steps, starting with the addition of country codes to all posts, depending on the respective influencer's main location or declared home base. The respective country codes in use can be retrieved from the description of variables as presented later on. Furthermore, the number of characters in each caption, that is, caption length, was counted, along with the number of hashtags and tags used if any, and it was determined whether questions were asked within the caption by checking for the use of question marks. Furthermore, since it was impossible to add an exhaustive list of all strings used by influencers to indicate advertisements to the list scanned for by the Java application without seriously falsifying results (e.g. string “ad” is one of the most frequently used sponsorship tags but also appears within many regular words), all post captions were checked manually for disclosed sponsorships undetected by the crawler. Afterward, sponsorships, as well as location usage, were coded as binary variables.

Furthermore, in order to put likes and comments into relation with influencer size and thus appropriately quantify the level of user-influencer interaction (Corporate Finance Institute, n.d.), interaction rates were calculated for each post, resulting in *like rate* (likes/followers*100), *comment rate* (comments/followers*100), and *engagement rate* ((likes+comments)/followers*100).

Lastly, caption texts were analyzed using Linguistic Inquiry and Word Count (LIWC) software and assigned scores between 0 and 100 for a multitude of themes detected by the tool through reading “a given text and count[ing] the percentage of words that reflect different emotions, thinking styles, social concerns, and even parts of speech” (LIWC, n.d.b., “How It Works”). In order to keep the number of variables manageable and relevant to the research hypotheses, merely the LIWC scores measuring caption authenticity and clout were included in the main analyses. Furthermore, all non-numerical variables such as the actual written captions, hashtags, and tags were removed from the dataset.

Log-Transformation

Next, when taking a closer look at the various dependent and independent variables emerging from the hypotheses, it was discovered that the variables across the dataset were mostly not normally distributed and seriously over-dispersed, indicating that they were not fit for linear regression (see Table 1 below) (Feng et al., 2014; Htoon, 2020). For example, Figure 5 shows the distribution histogram of variable *likerate* in comparison with a normal frequency curve overlay.

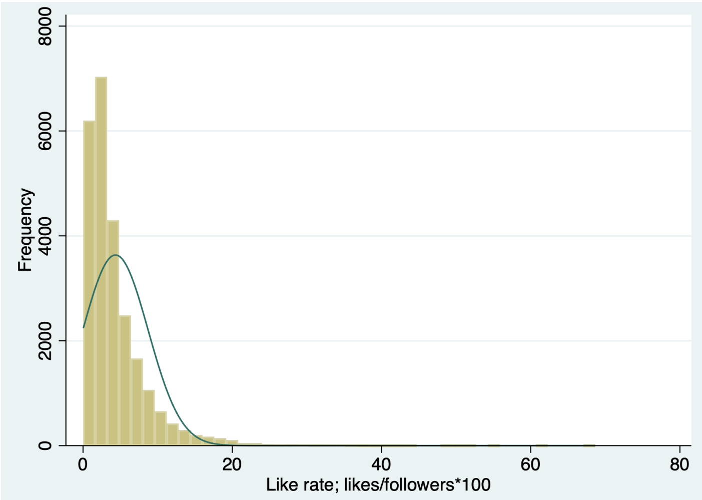

Figure 5: *Distribution of variable* likerate

In order to reduce the skewness of the dataset and convert it into a linear function, the commonly used approach of log-transformation was chosen to proceed (Feng et al., 2014). Herein, a natural log-transformation approach was selected given the underlying distribution, and a constant of one was added during the transformation in order to avoid zero-inflating the newly created variables named in the format "*ln_variablename*" to indicate their log-transformation (Feng et al., 2014).

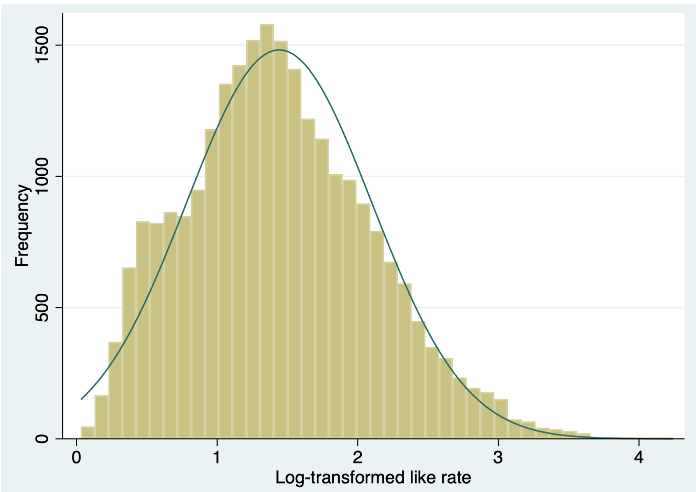

Figure 6: *Distribution of variable* ln_likerate

As can be seen in Figure 6 as well as Table 1, the log-transformation significantly decreased skewness and kurtosis values of the various dependent and independent variables and created near-normal distributions fit for linear regressions.

Descriptive Statistics

Variables	Obs	Mean	Std. Dev.	Skew.	Kurt.
likerate	24998	4.34	4.377	2.985	18.388
commentrate	24998	.413	1.416	28.166	1535.296
engagementrate	24998	4.754	5.093	3.927	37.934
captionlen	24998	222.729	267.929	2.394	11.15
hashtagcount	24998	2.553	6.023	3.31	13.617
tagcount	24998	.941	1.94	8.395	100.899
authentic	24998	41.117	36.599	.357	1.577
clout	24998	54.342	25.743	-.041	2.47
follower	24998	919000	2420000	6.122	51.838
ln likerate	24998	1.441	.659	.387	2.865
ln commentrate	24998	.231	.388	2.661	12.152
ln engagementrate	24998	1.495	.686	.391	2.915
ln captionlen	24998	4.723	1.281	-.338	2.79
ln hashtagcount	24998	.638	.932	1.471	4.322
ln tagcount	24998	.481	.53	1.164	5.49
ln authentic	24998	2.997	1.487	-.533	1.736
ln clout	24998	3.815	.805	-2.39	9.762
ln follower	24998	11.736	2.334	-.104	1.852

Table 1: *Descriptive Statistics (Skewness, Kurtosis)*

Control Variable Country

Despite the international, borderless nature of online social networks like Instagram, it was assumed that there nevertheless exist influences related to the country where an influencer is from, located in, or identifies with. Extensively covered in existing research, ethnocentrism, which is the existence of and human behavior based on in- and outgroups, has been found to greatly impact consumer behavior and to have a direct effect on willingness to buy (e.g. Shimp & Sharma, 1987; Josiassen, 2011). While ethnocentric consumers are likely to prefer products made in their home country and tend to consume what benefits their in-group, those ranking low in ethnocentrism are more likely to evaluate imported products on their own merits, regardless of their origin, or may even prefer products from other countries (Josiassen, 2011).

In the context of Instagram influencers and the similarity-likability theory that is equally premised on the existence of in- and outgroups, this means that users can be expected to interact more with posts from influencers originating from the same region or country as the user, since they would be more inclined to like them based on their similarity and vice versa. Contrarily, from the self-expansion theory perspective, users would be more likely to interact with influencers from countries beyond their immediate environment, in order to learn from them and to interact with different cultures, inspire travel, exploring, and self-expansion along with it.

Thus, while it was far beyond the scope of this research to determine every follower's country of origin in order to explore the exact effect of individual countries in more depth, the impact of influencer country on engagement rates was still expected and controlled for within the analyses.

Summing up, Table 2 below provides a list and description of all relevant variables used within the data analyses, which will be elaborated on shortly.

Description of Variables

name	label
likes	Likes
likerate	Like rate; likes/followers*100
ln_likerate	Log-transformed like rate
comments	Comments
commentrate	Comment rate; comments/followers*100
ln_commentrate	Log-transformed comment rate
engagementrate	Engagement rate; (likes+comments)/followers*100
ln_engagementrate	Log-transformed engagement rate
captionlen	Caption length (total, incl. spaces)
ln_captionlen	Log-transformed caption length
hashtagyesno	Hashtag usage, binary; 1 = hashtags used, 0 = no hashtags used
hashtagcount	Number of hashtags used
ln_hashtagcount	Log-transformed number of hashtags used
tagyesno	Tag usage, binary; 1 = tags used, 0 = no tags used
tagcount	Number of tags used
ln_tagcount	Log-transformed number of tags used
location	Location tag, binary; 1 = yes, 0 = no
qmarkyesno	Question mark usage, binary; 1 = question marks used, 0 = no question marks used
authentic	Caption authenticity score
ln_authentic	Log-transformed authenticity score
clout	Caption clout score
ln_clout	Log-transformed clout score
sponsored	Sponsored, binary; 1 = yes, 0 = no
follower	Number of followers
ln_follower	Log-transformed number of followers
country	Influencer country; 1 = US, 2 = AUS, 3 = UK, 4 = CAN, 5 = IE, 6 = NZ, 7 = SIN

Table 2: *Description of Variables*

Regressions

As a first step preceding the regressions, a correlation analysis of all independent variables was performed in order to identify potential multicollinearity problems within the data. The correlations matrix in Table 3 below shows that, unsurprisingly, variables *ln_hashtagcount* and *ln_tagcount* were highly positively correlated with their respective binary twin variables *hashtagyesno* and *tagyesno*, at correlation coefficients of 0.808 and 0.848, respectively. Given this multicollinearity, separate regressions were calculated for these variables throughout all analyses, resulting in two runs for each regression. A moderate positive correlation, that is, ranging between coefficients of 0.3-0.7 (Ratner, 2009), was detected among some independent variables, such as between sponsorship and tag usage (0.452 and 0.402), or caption length and hashtag usage (0.437 and 0.465). However, since the most commonly used threshold of 0.7 was not exceeded, no further variables were removed from the regression calculations (Ratner, 2009).

Matrix of Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) ln_captionlen	1.000										
(2) hashtagyesno	0.437	1.000									
(3) ln_hashtagcount	0.465	0.808	1.000								
(4) tagyesno	0.333	0.396	0.273	1.000							
(5) ln_tagcount	0.383	0.365	0.298	0.848	1.000						
(6) location	0.008	0.008	0.020	0.016	0.015	1.000					
(7) qmarkyesno	0.278	0.134	0.134	0.055	0.070	0.008	1.000				
(8) ln_authentic	0.316	0.036	-0.020	0.031	0.029	-0.002	0.093	1.000			
(9) ln_clout	0.239	0.136	0.137	0.128	0.126	-0.009	0.087	-0.147	1.000		
(10) sponsored	0.355	0.362	0.270	0.452	0.402	0.011	0.050	0.050	0.099	1.000	
(11) ln_follower	-0.178	-0.117	-0.215	-0.049	-0.073	-0.095	-0.052	-0.005	0.001	-0.033	1.000

Table 3: *Correlation Analysis*

As the three sets of hypotheses incorporate different angles of user-influencer interaction and thus various dependent variables, multiple econometric research techniques were applied. All models were estimated using Stata16.

In regards to the main model influencer disclosure hypotheses H1-H9, and enabled by the log-transformation of all continuous variables in the dataset, linear regression was selected to estimate the relationship between the continuous dependent variables like rate, comment rate, and engagement rate, and the various influencer disclosure measures as specified in the research model. Given the dependent variables' continuous nature as well as the data's (artificially created) linear distribution, homogeneity of variance, as well as low multicollinearity, simple OLS regression was applied to predict the model (Frost, n.d.). Herein, two runs were conducted for each dependent variable due to the correlation among *hashtagyesno* and *ln_hashtagcount*, as well as *tagyesno* and *ln_tagcount* (Ratner, 2009). Full model regressions including all independent variables and control variable *country* were chosen over stepwise regressions in order to observe the full, most accurate effect (Smith, 2018). Given the log-transformation of the data, Htoon (2020) and UCLA (n.d.) were followed to interpret the resulting coefficients of both continuous and binary independent variables. Following each run, the Variance Inflation Factor (VIF) was calculated in order to rule out further multicollinearity (Shrestha, 2020).

Analyzing the influencer size hypotheses H1a-H8a and aiming to estimate the relationship between the number of followers as an independent variable and the various influencer disclosure measures as dependent variables, OLS regressions and logistic regressions were deemed most suitable to predict the continuous respectively binary outcome variables (Menard, 2001). Where applicable, log-transformed variables were used. After the logistic regressions, the area under the curve (AUC) was calculated to estimate the quality of the regression as an addition to the *Pseudo R-squared* parameter (Faraggi & Reiser, 2002).

Lastly, in regards to interaction hypotheses H1b-H8b and aiming to identify whether the relationships estimated in H1-H9 were contingent on the number of followers, interaction analyses were performed by substituting one interaction term after the other into the full OLS models as was run before, keeping all other variables constant in order to observe the singled out interaction effects between each independent variable and the number of followers (Aguinis & Gottfredson, 2010).

Results of all regressions are illustrated within Tables 5-9 as well as Tables 3-6 in the Appendix and will be elaborated on in the following section. In order to check for robustness, a range of robustness checks were conducted, which will be discussed later on.

Quality Assessment

One of the factors contributing to this study's quality is its overall research design, which took into account the complex and multi-sided causes of human behavior. By not only examining one causal relationship but instead looking at two as well as the interaction effect among the two, the study design acknowledged the many ways in which user behavior on social media is influenced and thereby stands out from existing studies. Thus, a gradual development of the research model underlined the overall recognition that the interaction between users and influencers on Instagram is a highly complex construct. Together with the findings modeling the causal relationship(s) that were set out to analyze in the first place, it equipped this research with significant internal validity. Furthermore, by accounting for a possible lack of support for the hypotheses and introducing alternative theory as well as refining the analyses through robustness checks that will be discussed later on, the strength of the research conclusion in approximating rich explanations was further increased.

Another quality indicator of this research is constituted within its sample, which consisted of a large data pool of nearly 25.000 observations harvested in early 2021. Not only is the data's recency a crucial factor that clearly sets this study apart in terms of user behavior, platform functions, and usage patterns, but in combination with the large number of post data collected, which cannot easily be done due to protective measures implemented by Instagram, this made the research foundation truly unique. The mostly algorithm-powered process of data collection not only added a layer of dependability but moreover, the sample can be considered very representative as it - just like users consuming content on Instagram - was not restricted to one country or region but looked at English-speaking users worldwide. Additionally, sampling and coding bias was reduced to a minimum by implementing an inclusive and numbers-driven approach, leaving very little room for one-sidedness and thus supporting the external validity of the study.

Focusing the area of interest on one of the main niches on Instagram, that is, fashion blogging, furthermore provided the necessary detail to produce relevant findings, however at the same time left sufficient space to transfer the results of this study to other areas of interest. By choosing a niche that revolved around Instagram's main function, that is, sharing aesthetic visuals, findings became sufficiently generalizable and transferrable to most other forms of content and user-influencer interaction, and can even be extended to other visual-centered social media platforms.

Finally, not only the study's dependable and credible sample collection and the minimization of researcher bias ameliorated the overall research reliability, but so do the overall consistency and repeatability that can be expected from the study design due to a sufficient size of data being collected in a real user environment, balancing out major outlying influences tied to field experiments. Recording real responses without the subjects' awareness of the experiment further increased the reproducibility of the study, so that receiving inconsistent results would in fact rather indicate a shift in usage behavior and, given the fast-paced development of social networking platforms, even be a desirable outcome, making this research a seismograph for contemporary usage patterns.

Results

In the following chapter, first, the descriptive statistics of the dataset at hand will be presented, followed by the statistical results of the three sets of hypotheses that emerged from the regressions. Lastly, to confirm the robustness of the research model and to account for alternative explanations of some observed effects, a set of robustness checks was performed and will be elaborated on in the final part of this section.

Descriptive Statistics

Table 4 below summarizes the descriptive statistics for the underlying dataset, which consisted of 24'998 observations, among which influencers from the five previously identified influencer categories based on following size were equally distributed (see Table 1, Appendix). Approximately 48% of posts were made from US influencers, roughly 25% from Australian, and 21% from UK influencers, respectively. The remaining 6% were posted by Canadian, Irish, New Zealand-, and Singapore-based influencers (see Table 2, Appendix).

Caption length (*captionlen*) ranged from one to 2338 characters, with an average of 223 characters per caption. Binary variable *hashtagyesno* recorded a mean value of 0.42, meaning that less than half of the analyzed observations, that is, posts, used hashtags. The number of hashtags used ranged from 0 to 34, with an average of approximately three hashtags used per post. The mean value of binary variable *tagyesno*, indicating whether any accounts were tagged within a caption, was at 0.53, thus slightly higher than that of hashtags, whereas the number of tags ranged between 0 and 35, with an average of approximately one tag. Location tags accompanied only approximately 6% (0.056) of all observations, with a standard deviation of 0.23, and question marks were used in about 15% (0.146) of all post captions, with a standard deviation of 0.35.

Personality disclosure variables *authentic* and *clout* each ranged from 0 to 99 with means of approximately 41 and 54, respectively. The standard deviation for authenticity was close to 37 and nearly 26 for clout.

Next, binary variable *sponsored* recorded a mean of 23% (0.231), indicating that nearly every fourth analyzed post was marked as sponsored, with a standard deviation of 0.42. On an influencer level, the percentage of sponsorship within the recorded 100 posts ranged from 1% to 91%, indicating that every influencer posted commercial content at least once, and some within nearly every post. The average sponsorship percentage per influencer was recorded at approximately 23%.

The number of followers (*follower*) ranged from 1’036 to 25’776’259, with a mean of 918’984 and a standard deviation of 2’416’122.

The number of likes (*likes*) recorded ranged from 5 to 3,156,408, with the average post receiving approximately 40’724 likes at a standard deviation of 151’401. The number of comments (*comments*) per post ranged from a minimum of 0 to a maximum of 2’924’252, with a mean of approximately 538 comments per observation and a standard deviation of 19’651. Important to note, the maximum values of *likes* and *comments* may reflect the extrema of viral posts or giveaways aimed at triggering extreme engagement. Vice versa, comments may have been disabled for some posts, which may account for the minimum extreme of zero comments.

Lastly, values of *likerate* ranged from 0.03 to 68.66, with an average of 4.34 and a standard deviation of 4.38. *Commentrate* varied between 0 and nearly 104, which indicates that followers may have either commented more than once or that some posts, for example, giveaways, gained traction outside the usual community of followers. The mean comment rate was recorded at approximately 0.41, thus, significantly lower than the average like rate, with a standard deviation of 1.42. Lastly, *engagementrate* ranged from 0.03 to 129,5, with an average engagement rate of 4.75 and a standard deviation of 5.09.

Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
country	24998	1.924	1.189	1	7
category	24998	3	1.414	1	5
captionlen	24998	222.729	267.929	1	2338
hashtagyesno	24998	.417	.493	0	1
hashtagcount	24998	2.553	6.023	0	34
tagyesno	24998	.534	.499	0	1
tagcount	24998	.941	1.94	0	35
location	24998	.056	.229	0	1
qmarkyesno	24998	.146	.353	0	1
authentic	24998	41.117	36.599	0	99
clout	24998	54.342	25.743	0	99
sponsored	24998	.23	.421	0	1
follower	24998	918984.17	2416121.6	1036	25776259
likes	24998	40723.599	151400.64	5	3156408
comments	24998	538.285	19651.326	0	2924252
likerate	24998	4.34	4.377	.031	68.656
commentrate	24998	.413	1.416	0	103.992
engagementrate	24998	4.754	5.093	.031	129.495

Table 4: *Descriptive Statistics*

Due to high kurtosis among both dependent and independent variables, most were, as indicated in the Methodology section, naturally log-transformed, in order to become more suitable for linear regression (Feng et al., 2014). In doing so, skewness and kurtosis scores were significantly reduced and distributions were normalized (see Table 1, Methodology).

Hypothesis Testing

Next, in order to evaluate the hypotheses and research model, various regressions as presented within the Methodology section were performed. Empirical results are presented in Tables 5-9 as well as Tables 3-6 in the Appendix and are elaborated on in further detail in the following.

H1-H9: Influencer Disclosure Hypotheses

This section will present the regression results for main model hypotheses H1-H9 as seen in Table 5. The results for each hypothesis will be described and the assumptions on why social media users interact with influencers will be supported or not supported.

Linear Regression: H1-H9

Dep. Variable / Indep. Variables	ln_likerate (1)	ln_likerate (2)	ln_commentrate (1)	ln_commentrate (2)	ln_engagementrate (1)	ln_engagementrate (2)
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
ln_captionlen	-.030***	-.027***	.014***	.008***	-.023***	-.024***
hashtagyesno	-.241***	-	-.046***	-	-.246***	-
ln_hashtagcount	-	-.097***	-	-.010***	-	-.094***
tagyesno	-.132***	-	.	-	-.130***	-
ln_tagcount	-	-.137***	-	.	-	-.128***
location	-.061***	-.060***	.	.	-.055***	-.054***
qmarkyesno	-.052***	-.055***	.	.	-.047***	-.049***
ln_authentic	.012***	.008***	.	.	.010***	.007**
ln_clout	.011**	.	.	.	.011**	.
sponsored	.061***	.	.038***	.025***	.071***	.019*
ln_follower	-.084***	-.087***	-.106***	-.106***	-.106***	-.108***
1b.country						
2.country	-.090***	-.074***	-.013**	-.009*	-.091***	-.074***
3.country	.019*	.041***	.074***	.080***	.041***	.066***
4.country	.151***	.181***	.041***	.045***	.157***	.187***
5.country	.127***	.112***	-.236***	-.238***	.073**	.058*
6.country	-.326***	-.318***	-.293***	-.290***	-.368***	-.360***
7.country	-.546***	-.527***	-.201***	-.195***	-.577***	-.554***
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.135	0.124	0.414	0.412	0.163	0.150
Mean VIF	1.26	1.29	1.26	1.29	1.26	1.29

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant

Table 5: *Regression Results, H1-H9*

Information Disclosure

H1. Posts with longer captions will have higher engagement rates.

For posts with longer captions, a statistically significant negative effect was found for like rate (-.030***/-.027***) and engagement rate (-.023***/-.024***), meaning that longer captions were found to have lower like and engagement rates. A statistically significant positive effect was found for comment rate (.014***/.008***), meaning longer captions generally caused higher comment rates, whereas it should be mentioned that the positive effect for comment rate was not as strong as the negative effects for like and engagement rates. Since a statistically significant positive effect was found for only comment rate, this hypothesis cannot be supported.

H2.1. Posts that use hashtags will have higher engagement rates.

H2.2. Posts that use more hashtags will have higher engagement rates.

In regards to hashtags, for both posts that use hashtags and posts that use more hashtags, a consistent, statistically significant negative effect was found for like rate (-.241*** and -.097***), comment rate (-.046*** and -.010***), and engagement rate (-.246*** and -.094***). This means that posts that used hashtags were found to have overall lower engagement rates than posts that did not use hashtags and that posts that used more hashtags were found to have overall lower engagement rates than posts that used fewer hashtags. It should be noted that the effects of hashtag use on like rate and engagement rate are some of the strongest effects found throughout the results. Due to the consistent negative effects, both hypotheses are not supported.

H3.1. Posts that use tags will have higher engagement rates.
H3.2. Posts that use more tags will have higher engagement rates.
In regards to tags, for posts that use tags and posts that use more tags, a negative effect was found for both like rates (-.132*** and -.137***) and engagement rates (-.130*** and -.128***), while no statistically significant effect was found for comment rate in both cases. This means that posts that use tags were found to have lower like and engagement rates than posts that did not use tags and that posts that used more tags were found to have lower like and engagement rates than posts that used fewer tags. It should be noted that the negative effects of tag use and the number of tags on both like rate and engagement rate were very strong in comparison to other effects found. Taking all of these effects into consideration, neither hypothesis is supported.

H4. Posts with a location will have higher engagement rates.
For posts with a location, a statistically significant negative effect was found for both like rate (-.061***/-.060***) and engagement rate (-.055***/-.054***), meaning that posts that shared a location were found to have lower like and engagement rates in comparison to posts that did not share a location. No statistically significant effect was found for comment rate. Due to the negative and not statistically significant effects, this hypothesis is not supported by the data.

H5. Posts with questions will have higher engagement rates.
Looking at posts with question marks, a statistically significant negative effect was found for both like rate (-.052***/-.055***) and engagement rate (-.047***/-.049***), meaning that posts that included questions in the caption were found to have lower like and engagement rates in

comparison to posts that did not have questions in the caption. No statistically significant effect was found for comment rate. Therefore, this hypothesis is not supported.

Personality Disclosure

H6. Posts with more authentic captions will have higher engagement rates.

In regards to posts with more authentic captions, a statistically significant positive effect was found for both like rate (.012***/.008***) and engagement rate (.010***/.007***), meaning that posts with higher authenticity scores were found to have higher like and engagement rates. No statistically significant effect was found for comment rate. Due to the statistically significant positive effects for both like rate and engagement rate, this hypothesis is overall supported.

H7. Posts with less clout in their captions will have higher engagement rates.

For posts with higher clout scores, a statistically significant positive effect was found for both like rate (.011**) and engagement rate (.011**), meaning that posts with more clout in their captions were found to have higher engagement rates. Thus, less clout in a caption caused lower engagement rates. Again, no statistically significant effect was found for comment rate. Overall, due to the positive effects of higher clout scores on engagement, this hypothesis cannot be supported.

Commercial Disclosure

H8. Unsponsored posts will have higher engagement rates.

Concerning posts that were sponsored, a statistically significant positive effect was found for like rate (.061***), comment rate (.038***/.025***), and engagement rate (.071***/.019*), meaning that sponsored posts were found to be more likely to have higher overall engagement rates than posts that were not sponsored. Due to these results, hypothesis H8 cannot be supported.

Influencer Size

H9. Posts from profiles with a smaller following will have higher engagement rates.

Lastly, regarding posts from influencers with larger followings, a statistically significant negative effect was found for like rate (-.084***/-.087***), comment rate (-.106***/-.106***), and engagement rate (-.106***/-.108***). This means that posts from profiles with a larger following were found to have lower overall engagement rates than posts that came from profiles with smaller followings. Therefore, hypothesis H9 is supported.

H1a-H8a: Influencer Size Hypotheses

This section will present the regression results for H1a-H8a as displayed in Tables 6 and 7.

Linear Regression: H1a-H8a

Dep. Variable / Indep. Variable	ln_captionlen	ln_hashtagcount	ln_tagcount	ln_authentic	ln_clout
	Coef.	Coef.	Coef.	Coef.	Coef.
follower	-9.70e-08***	-5.97e-08***	-1.91e-08***	-2.60e-08***	-1.66e-08***
Number of obs.	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.033	0.024	0.008	0.002	0.002

*** p<.01, ** p<.05, * p<.1

Table 6: *Regression Results (Linear), H1a-H8a*

Logistic Regression: H1a-H8a

Dep. Variable / Indep. Variables	location	qmarkyesno	sponsored
	Coef.	Coef.	Coef.
follower	-4.67e-07***	-7.38e-08***	-8.56e-08***
Number of obs.	24‘998	24‘998	24‘998
Pseudo R-squared	0.019	0.003	0.004
AUC	0.6179	0.5470	0.5265

*** p<.01, ** p<.05, * p<.1

Table 7: *Regression Results (Logistic), H1a-H8a*

Information Disclosure

H1a. Influencers with a smaller following will share posts with longer captions.

In regards to longer captions, regression results convey that influencer size had a statistically significant negative effect on caption length (-9.70e-08***) that is among the strongest measured. This means that influencers with larger followings were found to have shorter captions and vice versa. Therefore, this hypothesis is supported.

H2a. Influencers with a smaller following will share posts with more hashtags.

Looking at hashtag count, influencer size was also found to have a statistically significant, moderate negative effect on the number of hashtags used (-5.97e-08***), indicating that influencers with larger followings tended to use fewer hashtags in their captions and that smaller influencers appeared to be using more hashtags. Thus, hypothesis H2a is supported.

H3a. Influencers with a smaller following will share posts with more tags.
In regards to tags, results indicate that influencer size had a statistically significant, rather weak negative effect on the number of tags used (-1.91e-08***). This means that influencers with larger followings were found to use slightly fewer tags in their captions, and that, similar to the results of hashtag usage, smaller influencers tended to tag more accounts in their captions. Due to these findings, hypothesis H3a is also supported.

H4a. Influencers with a smaller following will share posts with location more often.
Looking at location, regressions revealed that influencer size showed signs of a moderate, statistically significant negative effect on whether location was disclosed (-4.67e-07***). In other words, the more followers influencers had, the less likely they were to share the location of their posts in comparison to smaller influencers. These findings support hypothesis H4a.

H5a. Influencers with a smaller following will share posts with questions more often.
In regards to questions, results convey that influencer size had a statistically significant negative effect on the inclusion of questions (-7.38e-08***). This means that influencers with larger followings were found to include questions less often than influencers with smaller followings. Thus, this hypothesis is supported.

Personality Disclosure

H6a. Influencers with a smaller following will share posts with more authentic captions.
Concerning authenticity, linear regression showed that influencer size was found to also have a statistically significant negative effect on the authenticity score of posts (-2.60e-08***). This means that influencers with larger followings were found to share posts with less authentic captions than influencers with smaller followings. Due to this result, H6a is also supported.

H7a. Influencers with a smaller following will share posts with less clout in their captions.
Next, linear regression results for clout conveyed that the number of followers also slightly negatively affected the clout score of posts at the highest significance level (-1.66e-08***). Opposing the underlying assumptions of H7a, this means that influencers with larger followings were found to share posts with less clout in their captions than influencers with smaller

followings. Due to the statistically significant negative effect found, this hypothesis cannot be supported.

Commercial Disclosure

H8a. Influencers with a smaller following will share fewer sponsored posts.

Lastly, in regards to sponsorship, logistic regression revealed that influencer size had a strong, statistically significant negative effect on post sponsorship (-8.56e-08***), meaning that increasing following size negatively impacted the likelihood of influencers to share sponsored posts. Therefore, hypothesis H8a cannot be supported.

H1b-H8b: Interaction Hypotheses

This section will elaborate on the linear regression results for interaction hypotheses H1b-H8b as displayed in Tables 8 and 9 below, as well as Tables 3-6 in the Appendix.

Information Disclosure

H1b. Influencer size impacts the relationship between caption length and engagement rates.

In regards to like rate, results for H1b revealed major changes to the previously reported full model direct effect coefficient of *ln_captionlen*, shifting from a moderate negative to a strong positive effect (from -.030***/-.027*** to .240***/.328***). A moderate interaction effect between *ln_follower* and *ln_captionlen* was found (-.024***/-.030***). Similarly, significant increases to the full model coefficients were identified when looking at the interaction model predicting comment rate (.014***/.008*** to .140***/.148***), where a weaker interaction effect was identified (-.011***/-.012***). Looking at engagement rate, the full model coefficient also recorded drastic changes shifting from a moderate negative to a strong positive caption length coefficient (-.023***/-.024*** to .288***/.373***), with a moderate negative interaction term (-.027***/-.034***). Thus, influencer size not only affected the direct relationship between caption length and engagement rates in a way that the magnitude of this relationship decreased with an increased follower number, but including the interaction term in the full model also changed the nature of the overall individual coefficients quite drastically. Therefore, hypothesis H1b is supported.

H2b. Influencer size impacts the relationship between hashtag usage and engagement rates.

Looking at the interaction terms between influencer size and the relationship of hashtag usage and engagement rates, the following results were retrieved: in regards to like rate, *hashtagyesno* recorded drastic changes from its full, direct model coefficient (-.241*** to .329***/.341***), and a similar effect for *ln_hashtagcount* was found (-.097*** to .363***/.337***). Both hashtag usage variables recorded two of the strongest negative interaction terms (-.049***/-.051*** and -.044***/-.042***, respectively). In regards to comment rate, the *hashtagyesno* main model coefficient changed from moderately negative to strongly positive (-.046*** to .080***/.081***), with a slightly negative interaction term (-.011***). Similarly, the *ln_hashtagcount* main model coefficient also significantly changed from negative to positive (-.010*** to .108***/.104***), and an also rather weak interaction term of -.011*** was found. Looking at *ln_engagementrate*, similar effects were recorded, with the main model coefficients changing a lot and strong interaction terms, among the highest that was found (*hashtagyesno*: -.055***/-.057***; *ln_hashtagcount*: -.048***/-.046***). Thus, the overall negative interaction terms found at the highest significance level, as well as the drastic changes in main model coefficients support hypothesis H2b and its underlying assumption that influencer size impacts the relationship between hashtag usage and engagement rates.

H3b. Influencer size impacts the relationship between tag usage and engagement rates.

In regards to the two variables measuring tag usage, *tagyesno* and *ln_tagcount*, results show that the main model coefficients for both like and engagement rate also both recorded major changes from negative to positive values when substituting the respective interaction terms into the model. Moderate negative interaction effects were found for both *tagyesno* (-.029***/-.030***; -.034***/-.035***) as well as *ln_tagcount* (-.037***/-.032***; -.038***/-.043***). Rather weak but statistically significant negative interaction effects in regards to comment rate were identified (*tagyesno:* -.010***; *ln_tagcount:* -.010***/-.011***), and, while main model coefficients were almost entirely statistically insignificant in the main model, including the respective interaction terms lead to statistically significant, strongly positive coefficients of .109***/.105*** (*tagyesno)* and .118***/.121*** (*ln_tagcount*), respectively. These negative interaction effects throughout all engagement rates as well as the drastic changes that were registered among the main model coefficients suggest that hypothesis H3b can be supported and that influencer size does significantly impact the relationship between tag usage and engagement rates.

Linear Regression: H1b-H8b

Dep. Variable / Indep. Variables	ln_likerate											
	Main Model		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln_captionlen	-.030***	-.027***	**.240*** **	**.328*** **	-.036***	-.030***	-.027***	-.030***	-.029***	-.028***	-.023***	-.023***
hashtagyesno	-.241***	-	-.247***	-	**.329*** **	**.341*** **	-	-	-.241***	-	-	-.250***
ln_hashtagcount	-	-.097***	-	-.123***	-	-	**.363*** **	**.337*** **	-	-.098***	-.101***	-
tagyesno	-.132***	-	-.117***	-	-.124***	-	**-**	-.141***	**.206*** **	**.184*** **	-	-
ln_tagcount	-	-.137***	-	-.134***	-	-.114***	-.126***	-	-	-	**.277*** **	**.243*** **
location	-.061***	-.060***	-.061***	-.059***	-.064***	-.064***	-.058***	-.057***	-.063***	-.060***	-.061***	-.063***
qmarkyesno	-.052***	-.055***	-.059***	-.068***	-.049***	-.050***	-.052***	-.053***	-.054***	-.059***	-.061***	-.057***
ln_authentic	.012***	.008***	.016***	.010***	.015***	.014***	.010***	.011***	.013***	.009***	.008**	.011***
ln_clout	.011**	.	.009*	.	.011**	.009*	.	.009*	.011**	.009*	.	.009*
sponsored	.061***	.	.062***	.025**	.064***	.054***	.034***	.048***	.062***	.037***	.021**	.054***
ln_follower	-.084***	-.087***	.022***	.048***	-.067***	-.067***	-.066***	-.067***	-.071***	-.072***	-.071***	-.071***
ln_follower												
× ln_captionlen			-.024***	-.030***								
× hashtagyesno					-.049***	-.051***						
× ln_hashtagcount							-.044***	-.042***				
× tagyesno									-.029***	-.030***		
× ln_tagcount											-.037***	-.032***
× location												
× qmarkyesno												
× ln_authentic												
× ln_clout												
× sponsored												
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.135	0.124	0.147	0.142	0.142	0.142	0.140	0.140	0.138	0.129	0.128	0.138

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant
Control variable *country* was also included, however not reported due to brevity and relevance

Table 8: *Regression Results (Like Rate), H1b-H8b*

Linear Regression: H1b-H8b

Dep. Variable / Indep. Variables	ln_likerate (cont.)											
	Main Model		(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
ln_captionlen	-.030***	-.027***	-.030***	-.027***	-.031***	-.027***	-.030***	-.027***	-.032***	-.028*	-.030***	-.027***
hashtagyesno	-.241***	-	-.241***	-	-.241***	-	-.239***	-	-.242***	-	-.241***	-
ln_hashtagcount	-	-.097***	-	-.097***	-	-.098***	-	-.095***	-	-.100***	-	-.098***
tagyesno	-.132***	-	-.132***	-	-.132***	-	-.131***	-	-.131***	-	-.131***	-
ln_tagcount	-	-.137***	-	-.137***	-	-.138***	-	-.137***	-	-.136***	-	-.137***
location	-.061***	-.060***	**.269*** **	**.251*** **	-.062***	-.061***	-.062***	-.061***	-.061***	-.060***	-.062***	-.062***
qmarkyesno	-.052***	-.055***	-.052***	-.056***	**.**	**.105***	-.056***	-.059***	-.051***	-.054***	-.052***	-.056***
ln_authentic	.012***	.008***	.012***	.008**	.012***	.008**	**.118*** **	**.113*** **	.012***	.008**	.012***	.008***
ln_clout	.011**	.	.011**	.	.011**	.	.	.	**.180*** **	**.194*** **	.011**	.
sponsored	.061***	.	.060***	.	.061***	.	.	.	.060***	.	**.233*** **	**.260*** **
ln_follower	-.084***	-.087***	-.083***	-.085***	-.084***	-.085***	-.058***	-.061***	-.030***	-.028***	-.081***	-.083***
ln_follower												
× ln_captionlen												
× hashtagyesno												
× ln_hashtagcount												
× tagyesno												
× ln_tagcount												
× location			-.030***	-.029***								
× qmarkyesno					.	-.014***						
× ln_authentic							-.009***	-.009***				
× ln_clout									-.014***	-.016***		
× sponsored											-.015***	-.021***
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.135	0.124	0.136	0.124	0.136	0.124	0.138	0.126	0.137	0.126	0.136	0.125

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant
Control variable *country* was also included, however not reported due to brevity and relevance

Table 9: *Regression Results (Like Rate), H1b-H8b (continued)*

H4b. Influencer size impacts the relationship between location usage and engagement rates.

Analyses furthermore showed that there existed a moderate negative interaction effect (-.030***/-.029***) between the number of followers and location usage in regards to like rate, and that the overall main model coefficient also decidedly changed (from -.061***/-.060*** to .269***/.251***) when including the interaction term in the full model. In regards to comment rate, where the effect between *location* and *ln_commentrate* was not statistically significant before, a strong positive coefficient at the highest significance level (.338***/.336***) was found when considering the interactive term in the main model. A moderate negative interaction term of -.031***/-.030*** was recorded, which decreased when looking at the interaction effect on engagement rate (-.037***/-.036***). Similar to previous findings, main model coefficients indicating the effect of location on *ln_engagementrate* saw major changes (from -.055***/-.054*** to .350***/.333***). Thus, the moderate negative interaction effects across all engagement rates, as well as drastic coefficient changes, support hypothesis H4b and its underlying assumption that influencer size impacts the relationship between location tagging and engagement rates.

H5b. Influencer size impacts the relationship between question asking and engagement rates.

In regards to question mark usage, rather weak and the statistically least significant results among the hypothesis set were recorded. While full model coefficients moderately changed for like rate and slightly more for comment rate and engagement rate, a weak negative interaction effect was found across all three engagement measures (like rate: -0.014***; comment rate: -.009***/-.010***; engagement rate: -.012**/-.020***), overall supporting H5b and the assumption that influencer size affects the relationship between question mark usage and engagement rates.

Personality Disclosure

H6b. Influencer size impacts the relationship between caption authenticity and engagement rates.

Looking at authenticity levels and their interaction with influencer size, results show that *ln_authentic* main model coefficients moderately changed across all engagement rates when including the respective interaction term into the full model. Weak negative interaction effects were recorded (like rate: -.009***; comment rate: -.006***; engagement rate: -.010***), leading to the conclusion that a higher follower number leads to a slightly diminished effect of authenticity on engagement rates and that H6b can, despite recording rather weak effects, be supported.

H7b. Influencer size impacts the relationship between caption clout and engagement rates.

While in regards to like rate no statistically significant changes to the overall main model coefficients were found, a weak negative interaction effect of -.009*** was identified, indicating that the relationship between *ln_clout* and *ln_likerate* is indeed moderated by the number of followers. In regards to comment rate, an even weaker negative interaction term of -.002*/** was found, with slight changes to the main model coefficients (.023**/.025**). Overall engagement rate saw a similar but stronger change, with main model coefficients becoming statistically more significant and higher (changing from .011** to .191***/.204***), and another rather weak negative interaction effect of -.015***/-.017*** was found. Overall, despite ranking among the weakest interaction effects, the underlying assumption of influencer size affecting the relation between clout and engagement rates as suggested in H6b can be supported.

Commercial Disclosure

H8b. Influencer size impacts the relationship between post sponsorship and engagement rates.

Lastly, looking at H8b, analyses also revealed rather weak negative interaction effects in regards to both like rate (-.015***/-.021***), comment rate (-.008***/-.009***), as well as engagement rate (-.019***/-.026***). Along with the interaction effects, the main model coefficients also changed moderately when the respective interaction effects were inserted (like rate: .061*** to .233***/.260***; comment rate: .038***/.025*** to .128***/.125***; engagement rate: .071***/.019* to .294***/.317***). Thus, results showed that the larger the following, the weaker the effect of (positive) post sponsorship on engagement rates were, which supports H8b.

Robustness

In order to check for the general robustness of the above results and to test further explanations in an attempt to account for surprising findings, various additional and follow-up regressions were run.

First of all, in addition to the post-level analysis of 24,998 observations, the main model regressions were also run on an influencer level, taking into consideration all 100 posts made by each of the 250 influencers at once. Results overall persisted, indicating structural validity of the research. Furthermore, log-level, level-log, and level-level regressions were also performed, all of which exhibited similar results with slightly lower explanatory power than the log-log analysis.

Furthermore, hypotheses H1b-H8b also greatly contributed to the robustness of H1-H8 results by examining the latter closer in regards to different levels of influencer sizes. Examining how the core regression coefficients behaved when modifying the regressors towards interaction greatly enhanced the structural validity of this research.

Next, looking at the results of personality disclosure hypotheses H6 and H7, an additional linear regression was run to check for robustness. Herein, additional LIWC scores *posemo* and *negemo*, indicating caption scores of positive and negative emotions respectively, were used as a supplemental set of variables for personality disclosure, as it can be argued that showing emotions contributes to sharing one's personality. Showing no signs of multicollinearity with the other variables in the full model (see Table 7, Appendix), *posemo* and *negemo* were log-transformed and subsequently inserted into the full regression model. Results, displayed in Table 10 below, not only show that overall findings persisted but in particular that the disclosure of positive and negative emotions within captions had a very similar positive effect on engagement, just like authenticity and clout. Furthermore, looking towards the amount of positive and negative emotion disclosure in the context of hypotheses H6a and H7a, further linear robustness regressions examining the impact of *ln_follower* on *ln_posemo* and *ln_negemo*, respectively, were run. Herein, findings yet again showed that an increase in followers led to the disclosure of fewer positive and negative emotions (see Table 8, Appendix), which also supports the findings regarding initial personality disclosure variables authenticity and clout.

Linear Regression: H6-H7 Robustness

Dep. Variable / Indep. Variables	ln_likerate (1)	ln_likerate (2)	ln_commentrate (1)	ln_commentrate (2)	ln_engagementrate (1)	ln_engagementrate (2)
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
ln_captionlen	-.037***	-.035***	.014***	.007***	-.03***	-.032***
hashtagyesno	-.236***	-	-.046***	-	-.241***	-
ln_hashtagcount	-	-.094***	-	-.009***	-	-.09***
tagyesno	-.129***	-	.	-	-.127***	-
ln_tagcount	-	-.133***	-	.	-	-.124***
location	-.06***	-.059***	.	.	-.054***	-.053***
qmarkyesno	-.052***	-.055***	.	.	-.047***	-.049***
ln_authentic	.012***	.008***	.	.	.01***	.007**
ln_clout	.012**	.009*	.	.	.012**	.009*
sponsored	.061***	.	.038***	.025***	.071***	.02*
ln_follower	-.084***	-.087***	-.106***	-.106***	-.106***	-.108***
ln_posemo	.011***	.011***	.	.	.011***	.011***
ln_negemo	.035***	.032***	.	.	.034***	.032***
1b.country						
2.country	-.091***	-.075***	-.013**	-.01*	-.092***	-.075***
3.country	.02*	.043***	.074***	.08***	.043***	.067***
4.country	.15***	.18***	.041***	.045***	.157***	.186***
5.country	.125***	.111***	-.236***	-.238***	.071**	.057*
6.country	-.325***	-.317***	-.293***	-.289***	-.368***	-.358***
7.country	-.548***	-.528***	-.201***	-.195***	-.578***	-.555***
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.137	0.125	0.414	0.412	0.164	0.151

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant

Table 10: *Regression Results, H6-H7 Robustness*

Next, the research model was scanned for instrumental variables, that is, third variables without logical reason as to why they would predict the outcome variable but with an effect on the independent variable, that might therefore silently influence the dependent variable and dilute the results (Mehta, 2001). Given the two viability requirements for instrumental variables, that is, no logical relationship with the outcome variable and a significant correlation with an explanatory variable (Mehta, 2001), the follower-following ratio (*followingratio*) was calculated by dividing the number of accounts influencers themselves followed by the number of accounts that followed them. Quantifying the proportion of an influencer's binary role as both user and influencer, *followingratio* was identified as a potential instrumental variable. Looking at correlation, it was found to be moderately negatively correlated (-0.345) with *ln_follower*. Therefore proceeding, the endogeneity, or, put differently, the subjectivity, of variable *ln_follower* was examined in regards to the follower-following ratio, which may in practice be explained by the so-called *follow4follow* trend, where individuals mutually (often tacitly) agree to follow one another as long as the other follows back.

Running an instrumental variable regression on the full model (Table 9, Appendix) showed that overall results persisted, while post estimation endogeneity analyses of *ln_follower* resulted in very small Durbin- and Wu-Hausmann score p-values (0.000), leading to a rejection of H0 which suggested that *ln_follower* was an exogenous variable. Thus, results showed that follower size and, indirectly, engagement rates (as expressed in H9) were indeed affected by instrumental variable following ratio.

Lastly, further examining the somewhat surprising results regarding personality disclosure hypotheses H6 and H7, a follow-up regression was run to test an alternative explanation, which will be discussed in more detail later on. Hereby, a full model linear regression with influencer size as the dependent variable was performed in order to examine which post characteristics had positive effects on *ln_follower*. Controlling for *followingratio* and *country*, overall results persisted, and most information disclosure variables also had negative effects on this different type of user-influencer engagement, which underlines the robustness of previous results. Simultaneously, sponsorship, authenticity, and clout were yet again found to positively impact the user-influencer relationship. Regression results are displayed in Table 10, Appendix.

Discussion

Up until this point of the study, an overview of current literature revolving around influencers and social media platforms, Instagram in particular, was gathered and presented. Then, the study introduced the underlying research theories assumed to influence user-influencer interaction, that is, the similarity-likability theory and, from an opposite perspective, the self-expansion theory. Subsequently, the two theories were applied to the context of the study to construct and present the research model, along with the resulting three sets of hypotheses. Next, the research methodology was presented, followed by the results that emerged from the analyzed hypotheses construct, which are summarized in Table 11 below.

In the following, these findings will be linked back to the underlying theories and research domain, and interpretations of the findings will be presented regarding their contribution to knowledge, implications for practice, future research, and limitations.

Results Overview, by Hypothesis	
H1. Posts with longer captions will have higher engagement rates.	Not Supported
H2.1. Posts that use hashtags will have higher engagement rates.	Not Supported
H2.2. Posts that use more hashtags will have higher engagement rates.	Not Supported
H3.1. Posts that use tags will have higher engagement rates.	Not Supported
H3.2. Posts that use more tags will have higher engagement rates.	Not Supported
H4. Posts with a location will have higher engagement rates.	Not Supported
H5. Posts with questions will have higher engagement rates.	Not Supported
H6. Posts with more authentic captions will have higher engagement rates.	Supported
H7. Posts with less clout in their captions will have higher engagement rates.	Not Supported
H8. Unsponsored posts will have higher engagement rates.	Not Supported
H9. Posts from profiles with a smaller following will have higher engagement rates.	Supported
H1a. Influencers with a smaller following will share posts with longer captions.	Supported
H2a. Influencers with a smaller following will share posts with more hashtags.	Supported
H3a. Influencers with a smaller following will share posts with more tags.	Supported
H4a. Influencers with a smaller following will share posts with location more often.	Supported
H5a. Influencers with a smaller following will share posts with questions more often.	Supported
H6a. Influencers with a smaller following will share posts with more authentic captions.	Supported
H7a. Influencers with a smaller following will share posts with less clout in their captions.	Not Supported
H8a. Influencers with a smaller following will share fewer sponsored posts.	Not Supported
H1b. Influencer size impacts the relationship between caption length and engagement rates.	Supported
H2b. Influencer size impacts the relationship between hashtag usage and engagement rates.	Supported
H3b. Influencer size impacts the relationship between tag usage and engagement rates.	Supported
H4b. Influencer size impacts the relationship between location usage and engagement rates.	Supported
H5b. Influencer size impacts the relationship between question asking and engagement rates.	Supported
H6b. Influencer size impacts the relationship between caption authenticity and engagement rates.	Supported
H7b. Influencer size impacts the relationship between caption clout and engagement rates.	Supported
H8b. Influencer size impacts the relationship between post sponsorship and engagement rates.	Supported

Table 11: *Results Overview, by Hypothesis*

Contributions

In the following section, the findings from the Results section will be connected and analyzed, and interpretative arguments will be presented in order to explain the most interesting results, especially those that are associated with hypotheses that could not be supported. Results will be connected with the two main theories of this study, similarity-likability and self-expansion, and greater, broader contributions to knowledge that this study brings forth will be explored at the end.

Personality Disclosure

Regarding personality disclosure, while the logic behind authenticity and clout deemed that the two personality types are opposite, results of H6 and H7 convey that they both have in fact beneficial effects on engagement. These findings indicate that users like any type of personality disclosure, regardless of whether this may be authenticity, clout, or another kind of character display. This indication was further examined and its underlying assumption confirmed by the additional robustness tests performed to assess the effect of supplementary personality disclosure variables positive and negative emotion on engagement rates.

Interesting and important to note about this result is that, looking at the rather weak effect strengths, both personality disclosure variables, while being among the few beneficial influences, do not seem to be key determinants in regards to user engagement on posts. Information disclosure, commercial disclosure, and influencer size all play more important roles than how much personality is disclosed in a post when looking at engagement rates.

One possible explanation for this may be that personality disclosure plays a less important role when it comes to interaction with individual posts but instead a more important one when it comes to following an influencer. Through following, users may express approval or liking of the displayed personality of an influencer and may then be less inclined to judge every individual post based on this criterion. This may have the additional benefit of lowering cognitive effort when scrolling, thus making it more pleasant for users to view their feed, which consists of ‘approved’ influencers, making room for other criteria to determine whether a post is likable or not. Conducting a linear regression robustness test to explore the factors impacting influencer size, results indicate that influencers that typically disclose more of their personality are more likely to garner followers. From these findings, it may be concluded that, while personality disclosure does not play an important role in post engagement, its influence may have shifted to a different form of engagement, namely following.

When looking at the amount of personality disclosure in regards to influencer size (H6a-H7a), results show that smaller influencers - as expected in the context of the similarity-likability theory and rooted in the more personal relationships that smaller influencers have with their audience - disclose more personality compared to larger influencers. Checking this finding for robustness shows that, along with authenticity and clout, positive and negative emotion scores in captions also decrease with increased following size, leading to the confirmation that smaller influencers are more open to showing all facets of their personality, which confirms results of previous research and theoretical assumptions.

Furthermore, since results unanimously indicate that disclosing clout is in fact received positively, clout's perception and meaning must thus be reconsidered in the context of the self-branding economy that is Instagram. Instead of coming across as pretentious, clout seems to be recognized as self-confidence, which goes hand-in-hand with the inspirational aspect of the self-expansion theory, and may thus explain increased like and engagement rates with increased clout disclosure. This self-branding pattern may go so far as to what is often deemed *Fake It 'til You Make It*, a popular phrase especially within the social media culture, advising individuals to imitate confidence and competence in order to realize their goals and until they actually achieve the desired qualities in real-life (Nielsen, 2015; Powell-Brown, 2004). Thus, the fewer followers an influencer has, the higher the apparent need to fake power and influence on social media until it is actually realized, which may explain elevated clout scores for smaller influencers.

Drawing an overarching, triangular connection across all findings related to personality disclosure, it can be concluded that more personality disclosure within a post caption increases engagement rates slightly and that smaller influencers generally disclose more personality. At the same time, smaller influencers are, as expected from the Literature Review, confirmed to record higher engagement rates, meaning that personality disclosure results are, as displayed in Figure 7, overall and in themselves coherent across this study.

Commercial Disclosure

In contrast to the study's theoretical expectations, sponsorship was found to have a statistically significant, moderate positive effect on all engagement rates, leading to the assumption that decreased relatability to commercially motivated influencers is not a deterrent in regards to likability and engagement. Simultaneously, this can also be interpreted as sponsorship not negatively affecting relatability, but in fact, being perceived as a more likable post characteristic.

These findings thus show a connection to the alternative theory presented in this study, the self-expansion theory, where commercially sponsored posts constitute indicators of success and are therefore aspirational, which leads to higher engagement. Moreover, influencers are also often perceived as experts in their niche, and it is generally assumed that influencers strive to only advertise products that they truly believe in to avoid a negative backlash and to protect their credibility as experts. Thus, in this case, instead of interacting with influencers on the grounds of their likability and/or relatability, it is concluded that influencers are predominantly used for their expert knowledge and recommendations. In the social media world, this frequently takes the shape of sponsored posts which can be consumed in an inspirational manner that allows users to expand their horizons beyond the products and brands they already know.

Consistent with these surprising findings, the results of hypothesis H8a also contradict the similarity-likability assumption that influencers with larger followings share more commercial posts. In actuality, results convey that influencers with smaller followings are the ones to post more sponsored content, which can therefore be connected to a self-expansion perspective, from which smaller influencers perceive their ability to monetize their audience as a key distinguishing factor between themselves and their audience.

Another approach to explain this finding to consider is that smaller influencers typically earn significantly less than influencers with larger followings (Influencer Marketing Hub, 2021b), causing them to take on more commercial partnerships in order to generate the same income that larger influencers are able to collect from fewer sponsorships. Thus, influencers with a larger audience have more freedom in choosing which and how many brands to work with, but are simultaneously also often subject to much more scrutiny in regards to selling out and therefore have to watch out more to protect their integrity by keeping an acceptable balance between sponsored and unsponsored posts (BBC, 2020). Moreover, these findings also expose the trend for businesses and brands to partner with micro- or nano-influencers, who have shown to produce the most engagement and more clicks for a fraction of the cost of more popular influencers (Pusztai, 2019).

Summing up and connecting all commercial disclosure findings in an overarching, triangular approach shows that the study’s findings about post sponsorship are congruent: both sponsored content and smaller influencers generally record higher engagement rates, and thus, the fact smaller influencers share more sponsored posts is coherent with this argumentation.

Information Disclosure

Overall, in regards to information disclosure and engagement (H1-H5), none of this study's hypotheses are supported, and thus, it can be concluded that self-expansion theory plays a more vital role than similarity-likability theory in this context.

Constituting an exemption from the other information disclosure variables, caption length (H1) has the only statistically significant effect on comment rate, and it is also the only variable found to have a positive effect on any of the engagement rates. As longer captions are more likely to be interactive and open up a conversation with followers, the fact that longer captions lead to an increase in the number of comments makes sense and supports the previous logic using the similarity-likability theory.

The remaining variables, including caption length, however, have a negative effect on likes and overall engagement, indicating that too much information is disclosed. Since this study focuses on the fashion niche in particular, which first and foremost revolves around appearance, it can be interpreted that followers of fashion influencers primarily look for visual, more superficial inspiration over engaging conversations or deeper, more serious topics, reading and pondering of which would disrupt Instagram's crucial ease of scrolling (Juhász, 2020), which is characterized as a "sensational, instantaneous 'televisual glance'" and a moment of "fleeting interest" before the next image already appears (Carah & Shaul, 2016, p.72). In addition, younger users were found to scroll feeds 2.5 times faster than older users (Sloane, 2015). Given Instagram's young user base (Perrin & Anderson, 2019), the argument that shorter captions are easier to consume and that Instagram users are less likely to pause their ephemeral flow of images to read long captions before they approve of the post by giving their likes is further supported. While comment rate tends to increase with longer captions, this positive effect is the smallest of the three effects, meaning that like and engagement rates decrease more with longer captions than comment rate increases. Thus, the self-expansion theory outranks the similarity-likability theory in the context of caption length.

In summary, since results concerning the remaining information disclosure variables (i.e., hashtags, tags, location, and question marks) coherently convey that engagement rates decrease as information disclosure increases, it can be concluded that the self-expansion theory generally plays a more important role in this context and that users are more likely to engage when bite-sized information is disclosed. It is worth mentioning that in regards to hashtags, an additional explanation for this behavior may be that using hashtags is nowadays often perceived as desperation for visibility (Shiller, 2020), which may seriously harm one's inspirational value and perceived potential for self-expansion, as well as undermine one's authenticity, which, as previously discussed, has a positive effect on engagement. Furthermore, in regards to location, an additional argument may be that sharing location information may make influencers too approachable, and instead of standing on a pedestal of aesthetics and beauty, sharing their location makes them stand in a specific city just like the rest of their community, conveying a too real and tangible image to still constitute a flawless source of inspiration. Thus, overall, the more information an influencer discloses, the more this perception is distracted from or entirely taken away, leading to less engagement.

Summarizing and connecting information disclosure findings, increased information disclosure thus generally decreases engagement, and smaller influencers typically disclose more information. Adding the fact that smaller influencers record higher engagement rates to this logic unveils a major contradiction visualized in Figure 7 below, which indicates that the previous inferences about information disclosure representing a key factor as to why smaller influencers record more engagement is either flawed or incomplete.

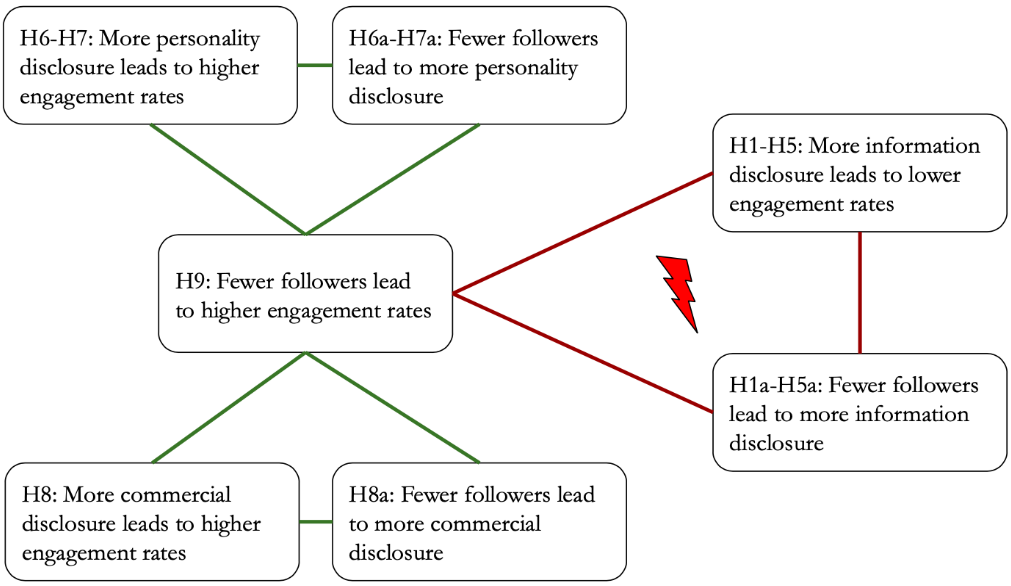

Figure 7: *Summary of Findings and Contradiction*

Influencer Size as a Moderator

Looking to account for this contradiction, the interaction effects measured within hypotheses H1b-H8b are further looked at. As previous results already indicate, influencer size seems to play a more central role in the research construct, and may potentially be an important pointer for the depth of the parasocial follower-influencer relationship.

And, in fact, results for H1b-H8b do convey that influencer size is a moderator in this overarching logic and that interaction effects are overall strong enough to undermine the full model findings in a way that the direct effect magnitude among disclosure and engagement decreases with increased following size. Put differently, when influencer size increases, relationship strength between disclosure variables and engagement decreases, which means that followers of larger influencers are less sensitive to the amount of post characteristics disclosed.

Importantly, these results mirror the findings of the Literature Review, which suggest that especially nano- and micro-influencers have a highly personal relationship with their audience, which to some extent consists of friends, acquaintances, and other local users (Izea, 2020; Mediakix, n.d.; Campbell & Farrell, 2020) who are more likely to take the time to read captions thoroughly, judge its content, and only then potentially interact with a post, in comparison to followers of big influencers who would usually not hope to see any content directly relating to their location, environment, or friends, and thus consume the content as a less personal, more visual-focused part of their Instagram feeds and rather interact with a post if it is simply aesthetically pleasing.

Most interestingly, while statistically significant however weak interaction effects in regards to personality and commercial disclosure exist, the by far strongest interaction effects were recorded particularly in regards to information disclosure hypotheses, which is where the direct effect logic seems to be faulty or incomplete. Thus, the interaction effect seems to be the missing puzzle piece of the overarching logic, an effect powerful enough to confound the direct effect model logic in regards to information disclosure, making it indispensable to take into account influencer size when looking at the relationship between information disclosure and engagement rates.

This important finding not only contributes to existing knowledge but also has far-reaching implications for various stakeholders in practice, which will be elaborated on shortly.

Similarity-Likability Versus Self-Expansion

Returning to the theoretical background of this study, results support the similarity-likability theory in regards to authenticity disclosure (H6 & H6a) and influencer size (H9 & H1a-H6a): more authentic posts and posts from smaller influencers are engaged with more, and smaller influencers disclose more information with their communities. In regards to the self-expansion theory, results support the hypotheses concerning information disclosure (H1-H5), clout (H7 & H7a), and sponsorship (H8 & H8a), putting forward that posts with little shared information, more clout and sponsorships are received better, and that the latter were more frequent among smaller influencers.

Firstly, looking at which of the theories predominantly affects the way in which influencers post in the first place as analyzed within hypotheses H1a-H8a, the assumption that smaller influencers disclose more information as they feel more similar to their followers, creating an intimate and likable atmosphere, can be confirmed. With six out of eight hypotheses supported (as displayed in Table 11 above), results clearly indicate that the way in which influencers post is predominantly shaped by the similarity-likability theory.

Nevertheless, turning to the main model, that is, the relationship between the various post characteristics disclosed and the resulting follower engagement, it is found that the self-expansion theory clearly outweighs the similarity-likability theory. Only two out of 11 hypotheses are weakly supported for the latter, while an overwhelming majority of nine hypotheses point towards self-expansion motives for primary user-follower interaction.

To conclude, while the underlying hypotheses are based on the similarity-likability theory supported in previous literature, this study’s findings convey that, overall, the self-expansion theory plays a more significant role in exploring why social media users interact with influencers online. For the first set of hypotheses, the vast majority of findings, in particular regarding information disclosure, clout, and commercial disclosure, convey that users are more interested in influencers who differentiate themselves from their following and play a more aspirational role, as opposed to being a peer. It is important to mention that following size plays a moderating role in this finding, meaning that the vast communities of large influencers are much less sensitive about what is disclosed in particular and rather seem to look at the underlying visual in a self-expanding manner, while followers of smaller influencers, having deeper social connections to the influencer, might be more prone to the similarity-likability theory in their engagement motives.

For the second set of hypotheses, the fact that smaller influencers are more likely to exhibit clout in their captions and produce more sponsored content additionally conveys that, sometimes, smaller influencers try to differentiate themselves from the typical, ordinary user, and thus also partly embrace the self-expansion theory.

For the final set of hypotheses, results indicate that as influencers grow in following size, users become less and less sensitive about what is disclosed in particular, be it information, personality, or commercial background, and appear to focus first and foremost on the aesthetics of the visual itself. This supports the self-expansion theory in that as influencers grow, users see them less as actual people and more so as resources of inspiration in their niche.

Taking all of these findings into consideration, it, therefore, can be concluded that there exists a significantly stronger argument that users interact with influencers for self-expansion purposes as opposed to because the influencer is similar or likable.

Broader Contributions to Knowledge

Briefly summing up broader research contributions, it was found that smaller influencers generally disclose more attributes with their posts on Instagram and that more personality and commercial disclosure increases engagement, while more information disclosure decreases engagement. Further, it was found that these effects are all, to varying extent, moderated by influencer size, and that especially information disclosure is perceived very differently by followers of smaller versus larger influencers, a finding that was linked to the superficiality or depth of the personal, parasocial relationship between follower and influencer. Further, the similarity-likability theory was found to influence the way in which influencers post in the first place, which can be understood as a baseline, but beyond this, the self-expansion theory was identified as the predominant theory explaining why users interact with influencers on social media.

Especially the latter finding contributes interesting knowledge and great insight for future studies regarding both of the underlying theories. While there exists a plethora of previous research on the similarity-likability theory (e.g. Alves et al., 2016; Collisson & Howell, 2014; Berscheid & Hatfield-Walster, 1969; Byrne, 1971; Hampton et al., 2019; Montoya & Horton, 2012) and the self-expansion theory (Aron et al., 2006; Aronson & Worchel, 1966; Jones et al., 1972; Izard & Smith, 1960; Goldstein & Rosenfeld, 1969; Nahemow & Lawton, 1975), and studies have identified the relevance of perceived similarity when it comes to follower-influencer relationships (Jin et al., 2019; Abidin & Ots, 2016; Argyris et al., 2020; Breves et al., 2019; Folkvord et al., 2020), no preceding study, to our knowledge, has connected users' interactions on social media specifically with either the similarity-likability theory or the self-expansion theory, making this study the first of its kind.

Moreover, most preceding research focuses on the influencers' perspective, dissecting different aspects of the role, such as how influencer capitalize on selfies (Abidin, 2016), how they use self-branding (Duffy & Hund, 2015; Duffy & Pooley, 2019; Khamis et al., 2017) or emojis (Ge & Gretzel, 2018), how they differentiate from celebrities (Jin et al., 2019), how they capitalize on intimacy (Raun, 2018), and how they manage authenticity (Audrezet et al., 2020). The works that have focused on the user perspective either have researched how users are or are not able to recognize influencer advertisements (Van Reijmersdal & Van Dam, 2020; Evans et al., 2018) or how users' mental health is affected by influencers (Chae, 2018). The remaining literature focuses on and examines how influencers are used as a marketing mechanism (Argyris et al., 2020; Breves et al., 2019; Campbell & Farrell, 2020; Evans et al., 2018; Folkvord et al., 2020; Kádeková & Holienčinová, 2018; Haenlein et al, 2020; Konstantopoulou et al., 2019; Stubb & Colliander, 2019). While one previous study has researched how followers become loyal to influencers through interactivity (Jun & Yi, 2020), no preceding study has tried to model why users interact with influencers, and therefore brings a new perspective to influencer literature and provides novel and interesting results for future research to dig deeper into.

This study is also special in that it contributes to pioneering a quantitative application of the netnographic research method. While, to our knowledge, netnographic research has previously

mostly been conducted in a qualitative way (Kozinets, 2010), for example through interviews, this study uses netnography in an exclusively quantitative way, showcasing that the vast amount of data observable online is sufficient to generate knowledge about cultures and communities online.
Moreover, this study also especially contributes to very current knowledge due to the recency of the collected data. In addition to that, the fact that the data was collected from Instagram, which, given the visual nature of the platform and strict limitations to its API, is not easy to come by.

Furthermore, while this study specifically focuses on the niche of fashion blogging on Instagram, findings can generally be transferred to other niches and social media platforms. Like fashion influencers, most other niches like fitness, beauty, and interior heavily rely on aesthetic visuals, allowing this study's results to be generalized broadly. Similarly, while social media platforms such as Facebook, Twitter, and LinkedIn are not as heavily reliant on visual content as Instagram is, results can also be horizontally generalized to other visually-centered social media platforms, such as YouTube and Tik Tok.

Lastly, this study's findings also indicate that online interactions, despite increasingly replacing interactions in real-life and even more so during the global Covid-19 pandemic, are distinctly different from them. While the similarity-likability theory may play a larger role in real-life, everyday interactions, social media presents the opportunity for users to have one-sided interactions with people they likely would not meet or get to know in real-life. While social media allows users to connect with friends and family, it also provides users with a virtual window into the personal lives of strangers and role models on a daily basis, while in real-life, in order to get access to a person's private life, one would usually have to befriend said individual and earn their trust over time. Putting aside any geographical differences, this may become especially difficult if these two people perceive one another as very different or not necessarily likable. However, online, users are able to get lots of personal information about influencers without any of these obstacles, which is a very major distinction between real-life and online interactions and could explain why the self-expansion theory proved to be the more dominant force online even though the similarity-likability theory may play a more important role in real-life. These insights contribute to the literature of both theories and would be an interesting jumping-off point for future research.

Future Research

There are many more areas of interest that could be explored further, which will be discussed in the following. Generally, influencer-user interactions and relationships are still a highly complex phenomenon, and thus any future attempt to reduce this complexity would contribute greatly.

First of all, while this study deduces high levels of generalizability from its focus on visual-centered fashion blogging and thus expects homogenous results for other visual-focused niches across Instagram and similar social media platforms, either a direct benchmark study or a study revolving around a lesser visual-focused niche, such as education or illnesses/health, would be highly interesting.

Furthermore, one limitation of this study, which will be explored further in the succeeding section, is its inability to collect data on individual comments. Future researchers could thus set out to analyze the specific content of comments under a post in order to gain a richer understanding of the users' opinion of a specific influencer or a specific type of post. Furthermore, it would be insightful to look into the differences between influencers that interact with users in the comments through replying to and liking comments versus those that do not. The underlying motivation behind commenting would also represent a highly interesting topic to further explore, aiming to learn more about whether and how users respond to only the picture, the caption, or whether they may be commenting just to gain more visibility for their own profile. Conducting an even richer content analysis, future researchers could also conduct an artificial-intelligence- or machine-learning-powered analysis of Instagram visuals.

Another aspect of this study that could be delved further into is how different types of sponsorships or brand partnerships influence engagement in the context of how well they fit into an influencer's usual area of interest. Future research could analyze sponsorships for brand-influencer fit and analyze how engagement changes depending on how suitable the brand is to the niche, for example comparing a fashion influencer promoting a clothing brand versus an insurance brand. This would, of course, require a more detailed and case-based data collection or could be complemented through the comment analysis discussed previously.

Finally, another opportunity for future studies would be to research how this study would turn out on entirely different social media platforms. With Instagram's distinct focus on visuals, it would be interesting to see how results would differ on a platform with a different area of content, such as Twitter. In the same context, it would also be interesting to study how users interact with the same influencer on different platforms since influencers are usually present on more than one social media platform.

Implications for Practice

Unique to the context of social media influencers, the relationship between users and influencers in their simultaneous role as peer and promoter not only has important implications for influencers themselves, but importantly also for the brands/advertisers they work with, the social media platforms themselves, and even for regulators of the industry. In the following section, this study's findings will be connected to this variety of stakeholders, presenting both specific and broader implications.

Naturally, the results of this study are most directly applicable to influencers since influencers rely on users interacting with their content, which is the focus of this study. First of all, due to the strong, negative effect that information disclosure has on user engagement, influencers should, on the one hand, be more conservative in regards to sharing too much information in their captions. On the other hand, overall, authenticity, clout, and sponsorship have been shown to lead to increased engagement. Taking these findings into consideration, influencers can develop strategies to improve their engagement in order to increase visibility, since the Instagram algorithm rewards content with higher user engagement. Furthermore, findings also indicate that influencer size matters, so as influencers grow, they can reach different stages in their influencer careers, where users' motivations and behavior shift. It is therefore vital for influencers to recognize what stage of their influencer career they are in, and use these specific findings to manage their content and partnerships - and partnership terms - accordingly. For example, smaller influencers would be well-advised to base their pricing on the engagement rates they generate, as this is where they perform especially well compared to larger influencers.

As for businesses/advertisers and social media agencies, these findings provide crucial insights into what type of influencer makes an optimal partner for their underlying marketing goals. On the one hand, followers of smaller influencers are more likely to pay attention to the additional information disclosed within a post, which constitutes a deciding factor if the goal is to convey a specific marketing message or if a firm wants to specifically educate users about their product. Smaller influencers also have higher engagement rates and are more likely to garner an audience similar to themselves, so if a company would like to access a very specific type of audience, using smaller influencers would be the most effective marketing method. On the other hand, if a company's goal is to increase awareness and/or generate as many impressions of their product as possible, they would, based on the knowledge generated by this study, be advised to partner with larger influencers, since they have greater reach and are generally perceived as inspirational guides on the top of their niche. Thus, these findings can also be used by companies and social media agencies in optimizing their marketing or entire business strategy and enhance the overall performance of their business.

Furthermore, this study's findings are also relevant to the makers of social media platforms themselves, especially Instagram, who can use these insights to drive and adapt their product design and development strategies to make them most fit for user behavior in an empirically-driven problem/solution coevolution (Gleasure, 2015). For example, this could take the shape of implementing functionalities that encourage positive user engagement and therefore create fulfilling user experiences on their platforms. As the results suggest that increased follower size negatively affects engagement, these platforms could for example test removing the follower count from influencers' profiles in order to minimize this inclination. In regards to the findings surrounding information disclosure, platforms could also provide influencers with an option to hide post details, such as hashtags, tags, and location, in order not to deter users and still provide the interactive qualities of these attributes. Additionally, effects on comment rate were, for large parts of this study, not statistically significant, which may be related to a small number of comments on any given post. Commenting requires a lot of effort, especially in comparison to liking, however it contributes to bonding with the influencer and other users, and thus to a fulfilling user experience, so adding further functionalities to ease or simplify the process of commenting would also be beneficial for platforms to consider.

Lastly, looking at the way in which the findings contain implications for regulating authorities, which play an increasingly big role in shaping online social networks, significant relevance was identified especially in the context of influencer marketing, which is for the most part still a grey area when it comes to the monitoring and enforcement of advertising legislation (Bartels & Terstiege, 2020). Especially smaller influencers can easily advertise under policymakers' radars, and, as the results indicate that small influencers advertise more than large ones, this study may aid regulators in identifying which users are most vulnerable and may therefore require more protection. While followers of large influencers, on one hand, fall victim to undisclosed ads through their decreased interaction sensitivity, followers of small influencers, on the other hand, are also especially at risk of being influenced by hidden marketing campaigns due to their more personal relationship and higher disclosure sensitivity. Ultimately, this may encourage a tightening of the rules for fair influencer ad disclosure or a gradual increase of monitoring and enforcement requirements across all influencer categories, which would ultimately aid regulators in carrying out their role as consumer protection institutions.

Limitations

There are various limitations in this study worth noting. Firstly, in regards to this study's use of the LIWC software tool to analyze captions, the analysis quality generally decreases with shorter caption length. Thus, instead of analyzing each individual caption, this study could have analyzed captions in bulk on an influencer level in order to derive more accurate results. The main model data points affected by this limitation are authenticity and clout scores, as well as positive and negative emotion scores within the robustness checks. This study however decided not to conduct these analyses on an influencer level because all other analyses conducted for the first set of hypotheses were analyzed at post-level. Therefore, for consistency, captions were analyzed individually.

Furthermore, a major limitation was faced in regards to data collection, which is generally limited as certain types of Instagram content are impossible to scrape. For example, the data scraper was able to collect data from posts, however not from Instagram Stories, certain Reels, and IGTV videos. These types of content, especially Stories, are also vital parts of the entire content package that Instagram influencers provide to their communities. However, out of all content types available on Instagram, posts, which were able to be scraped, are most significant to this study, because they are the most common format used by influencers on Instagram (Bailis, n.d.).

Additionally, the data crawler this study used was limited in what data points could be collected from a given Instagram post. For example, a data point that could not be collected was tags within a picture itself, which is commonplace for influencers to do instead of using caption tags and may again potentially affect the accuracy of the findings regarding tags. Another drawback in this regard was that the individual comments on an Instagram post could not be collected, only the number of comments. Not only does this affect the accuracy of some of the results, for example, sometimes tags and hashtags are hidden in the comments, but collecting comments would have also enabled greater insight into user sentiment and influencer interaction with their audience, that is, how often they reply or like comments. Furthermore, it required this study to interpret comments as a single-leveled, overall positive form of interaction with influencers, which may not always be accurate. However, a large number of comments also improves overall post visibility within the algorithm (Warren, 2021), leading to more traffic for the influencer. Therefore, even negative comments can have a positive effect on the attention economy (Marwick, 2015) of Instagram, as they increase impressions, and generate even more engagement.

Conclusion

The rise of influencers on social media within recent years has been steep and indisputable. Today, there are millions of aspiring online trendsetters that attract increasing numbers of followers by sharing their personalities and private lives with ordinary users as if they were real-life friends. While conveying this image of friendship, influencers however also utilize their channels to broadcast paid content to their communities, and in doing so obscure the lines between friend and salesman, peer and promoter.

Motivated by this paradoxical setup and the fact that influencers are as popular as ever before, this study's objective was to investigate what incentive social media users find in the interaction with influencers, who neither constitute a real friendship nor a genuine source for product recommendations.

Setting out to answer the research question "What motivates social media users to interact with influencers, whose attention might be diffused and commercially motivated?", this study first reviews existing literature about influencers, that is, individual who have gained a large following and the status of trusted tastemakers on social media (e.g. Jun & Yi, 2020). Having emerged gradually from various trends in media and society, influencers are clearly distinct from traditional persons of public interest, as they independently self-brand themselves as micro-celebrities on social media. Ranging from small to large in size, influencers are particularly drawn to visual-focused social media platform Instagram, which is almost single-handedly responsible for shaping the term influencer and propelling influencer marketing to a multi-billion dollar industry (Haenlein et al., 2020).

Following the Literature Review, this positivist and mostly deductive study reviews two existing theories and applies them to the context of social media influencers. The main theory, the similarity-likability theory, assumes that influencer interaction appeals to users because of any real or perceived connection they may share with influencers. Rooted in social identity theory, this approach argues that similarity breeds liking and vice versa. An alternative explanation for user-influencer interaction is found in the self-expansion theory, which suggests that users have a tendency to interact with those who are dissimilar to themselves, and who thus present an opportunity to grow through interaction. Allowing to expand horizons and discover new things, the self-expansion theory represents the aspirational facet of the user-influencer relationship.

Limiting the scope of the research to fashion influencers on Instagram, the research model and three sets of hypotheses are then presented, aiming at scientifically examining the topic from an influencer disclosure perspective, an influencer size perspective, as well as from an interaction term perspective. Having sampled, collected, and prepared a sufficiently large dataset of recent influencer posts, statistical regressions are then conducted to examine interaction, quantified through like rate, comment rate, and engagement rate.

Results show that both personality and commercial disclosure are overall positive factors in terms of engagement and that smaller influencers tend to disclose more of each. Looking at information disclosure, a major contradiction is identified, suggesting that the underlying research logic was flawed or incomplete. Looking to account for this contradiction, influencer size is found to exert a moderating interaction effect that leads to different levels of disclosure sensitivity and affects user-influencer interaction across the different influencer categories.

Overall, it is concluded that the way in which influencers post is predominantly shaped by the similarity-likability theory, while main model results however clearly indicate that social media users are mostly drawn to aspirational content and personal growth opportunities, and thus predominantly interact with influencers, whose attention might be diffused and commercially motivated, out of self-expansion motives.

Besides filling an existing research gap, these findings also entail significant contributions to existing research as well as far-reaching implications for various stakeholders in practice, providing rare evidence for the future strategic setup of influencers, businesses, platform designers, and regulating authorities.

References

Abidin, C. (2015). Communicative ❤ Intimacies: Influencers and Perceived Interconnectedness. *ADA: A Journal of Gender, New Media, & Technology, 8.* doi:10.7264/N3MW2FFG

Abidin, C. (2016). “Aren’t These Just Young, Rich Women Doing Vain Things Online?”: Influencer Selfies as Subversive Frivolity. *Social Media + Society, 2*(2), 1-17. doi:10.1177/2056305116641342

Abidin, C., & Ots, M. (2016). Influencers tell all? Unravelling Authenticity and Credibility in a Brand Scandal. In M. Edström, A. T. Kenyon, & E.-M. Svensson (Eds.), *Blurring the Lines: Market-driven and Democracy-driven Freedom of Expression* (pp. 153-161). Göteborg, SE: Nordic Information Centre for Media and Communication Research.

Abrams, D., & Hogg, M. A. (1990). Social Identification, Self-Categorization and Social Influence. *European Review of Social Psychology, 1*(1), 195–228. doi:10.1080/14792779108401862

Advertising Standards Authority. (2020, July 31). *Recognising ads: Social media and influencer marketing*. Retrieved from https://www.asa.org.uk/advice-online/recognising-ads-social-media.html

Agrawal, A. J. (2016, December 27). Why Influencer Marketing Will Explode In 2017. *Forbes*. Retrieved from https://www.forbes.com/sites/ajagrawal/2016/12/27/why-influencer-marketing-will-explode-in-2017/

Aguinis, H., & Gottfredson, R. K. (2010). Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Journal of Organizational Behavior, 31*(6), 776–786. doi:10.1002/job.686

Alves, H., Koch, A., & Unkelbach, C. (2016). My friends are all alike — the relation between liking and perceived similarity in person perception. *Journal of Experimental Social Psychology, 62*, 103-117. doi:10.1016/j.jesp.2015.10.011

Andrejevic, M. (2004). *Reality TV: The Work of Being Watched*. Lanham, MD: Rowman & Littlefield Publishers.

Argyris, Y. A., Wang, Z., Kim, Y., & Yin, Z. (2020). The effects of visual congruence on increasing consumers' brand engagement: An empirical investigation of influencer marketing on Instagram using deep-learning algorithms for automatic image classification. *Computers in Human Behavior, 112*. doi:10.1016/j.chb.2020.106443

Arnold, A. (2017, June 20). Why YouTube Stars Influence Millennials More Than Traditional Celebrities. *Forbes*. Retrieved from https://www.forbes.com/sites/under30network/2017/06/20/why-youtube-stars-influence-millennials-more-than-traditional-celebrities/

Aron, A., Steele, J. L., Kashdan, T. B., & Perez, M. (2006). When similars do not attract: Tests of a prediction from the self-expansion model. *Personal Relationships, 13(4),* 387-396. doi:10.1111/j.1475-6811.2006.00125.x

Aronson, E., & Worchel, P. (1966). Similarity versus liking as determinants of interpersonal attractiveness. *Psychonomic Science, 5(4)*, 157-158. doi:10.3758/BF03328329

Audrezet, A., de Kerviler, G., & Moulard, J. G. (2020). Authenticity under threat: When social media influencers need to go beyond self-presentation. *Journal of Business Research, 117*, 557-569. doi:10.1016/j.jbusres.2018.07.008

Bailis, R. (n.d.). *The State of Influencer Marketing: 10 Influencer Marketing Statistics to Inform Where You Invest*. Retrieved from https://www.bigcommerce.com/blog/influencer-marketing-statistics/#what-is-influencer-marketing

Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: quantifying influence on twitter. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining - WSDM '11,* 65-74. doi:10.1145/1935826.1935845

Barassi, V., & Treré, E. (2012). Does Web 3.0 come after Web 2.0? Deconstructing theoretical assumptions through practice. *New Media & Society, 14*(8), 1259-1285. doi: 10.1177/1461444812445878

Barlow, A. (2007). *Blogging America: The New Public Sphere.* Westport, CT: Praeger Publishers.

Bartels, L. & Terstiege, M. (2020). Influencer-Marketing – Digitale Einflussnahme auf Zielgruppen. In M. Terstiege (Ed.), *Digitales Marketing – Erfolgsmodelle aus der Praxis* (pp. 61-78). Wiesbaden, Germany: Springer Gabler. doi: 10.1007/978-3-658-26195-5_4

BBC. (2020, October 16). Instagram targets rule-breaking influencers. *BBC*. Retrieved from https://www.bbc.com/news/technology-54571630

Beals, J. S. (2010). *Self Marketing Power: Branding Yourself as a Business of One*. Omaha, NE: Keynote Publishing.

Bernstein, B. (2019, May 24). A Brief History Of The Influencer. *Medium*. Retrieved from https://medium.com/@bhbern/a-brief-history-of-the-influencer-1a0ef2b36c6e

Berscheid, E., & Hatfield-Walster, E. (1969). Rewards Others Provide: Similarity. In E. Berscheid & E. Hatfield-Walster (Authors), *Interpersonal Attraction* (1st ed., pp.69-91). Menlo Park, CA: Addison-Wesley.

Bourdieu, P. (1986). The Forms of Capital. In J. G. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). New York, NY: Greenwood Press.

Bowler, G. M. (2010). Netnography: A Method Specifically Designed to Study Cultures and Communities Online. *The Qualitative Report, 15*(5), 1270-1275. doi:10.46743/2160-3715/2010.1341

Breves, P. L., Liebers, N., Abt, M., & Kunze, A. (2019). The Perceived Fit between Instagram Influencers and the Endorsed Brand. *Journal of Advertising Research, 59*(4), 440-454. doi:10.2501/jar-2019-030

Britt, R. K., Hayes, J. L., Britt, B. C., & 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. doi:10.1080/15252019.2020.1763873

Brody, E. W. (2001). The “Attention” Economy. *Public Relations Quarterly, 46*(3), 18-21.

Brown, D., & Hayes, N. (2008). *Influencer Marketing: Who Really Influences Your Customers?* New York, NY: Taylor & Francis.

Burns, K. S. (2020). The History of Social Media Influencers. In B. Watkins (Ed.), *Research Perspectives on Social Media Influencers and Brand Communication* (pp. 1-21). London, UK: Lexington Books.

Burrow, A. L., & Rainone, N. (2017). How many likes did I get?: Purpose moderates links between positive social media feedback and self-esteem. *Journal of Experimental Social Psychology, 69,* 232-236. doi:10.1016/j.jesp.2016.09.005

Business Insider. (2021, January 6). *Influencer Marketing: Social media influencer market stats and research for 2021*. Retrieved from https://www.businessinsider.com/influencer-marketing-report?r=DE&IR=T

Byrne, D. E. (1971). *The Attraction Paradigm.* New York, NY: Academic Press.

Cambridge Dictionary. (n.d.). *Clout*. Retrieved from https://dictionary.cambridge.org/dictionary/english/clout

Campbell, C., & Farrell, J. R. (2020). More than meets the eye: The functional components underlying influencer marketing. *Business Horizons, 63*(4), 469-479. doi:10.1016/j.bushor.2020.03.003

Campbell, C., & Grimm, P. E. (2019). The Challenges Native Advertising Poses: Exploring Potential Federal Trade Commission Responses and Identifying Research Needs. *Journal of Public Policy and Marketing, 38*(1), 110-123. doi:10.1177/0743915618818576

Campbell, J. D., Trapnell, P. D., Heine, S. J., Katz, I. M., Lavallee, L. F., & Lehman, D. R. (1996). Self-Concept clarity: Measurement, personality correlates, and cultural boundaries. *Journal of Personality and Social Psychology, 70*(1), 141-156. doi:10.1037/0022-3514.70.1.141

Carah, N. & Shaul, M. (2016). Brands and Instagram: Point, tap, swipe, glance. *Mobile Media & Communication, 4*(1), 69-84. doi:10.1177/2050157915598180

Chaiken, S., & Manis, M. (1980). Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion. *Journal of Personality and Social Psychology, 39(*5), 752-766. doi:10.1037/0022-3514.39.5.752

Chae, J. (2018). Explaining Females' Envy Toward Social Media Influencers. *Media Psychology, 21*(2), 246-262. doi:10.1080/15213269.2017.1328312

Chen, C.-P. (2013). Exploring Personal Branding on YouTube. *Journal of Internet Commerce, 12*(4), 332-347. doi:10.1080/15332861.2013.859041

Christensen, C. (2007). YouTube: The Evolution of Media? In D. Moton & G. Dumler (Eds.), *Navigating America: Information Competency and Research for the Twenty-First Century* (pp.36-40). New York: McGraw-Hill Higher Education.

Cialdini, R. B., Trost, M. R., Newsom, J. T., & Geen, R. (1995). Preference for Consistency: The Development of a Valid Measure and the Discovery of Surprising Behavioral Implications. *Journal of Personality and Social Psychology, 69(2)*, 318-328. doi:10.1037/0022-3514.69.2.318

Collisson, B., & Howell, J. L. (2014). The Liking-Similarity Effect: Perceptions of Similarity as a Function of Liking. *The Journal of Social Psychology, 154(5)*, 384-400. doi:10.1080/00224545.2014.914882

Competition & Markets Authority. (2019, January 23). *Social media endorsements: being transparent with your followers.* Retrieved from https://www.gov.uk/government/publications/social-media-endorsements-guide-for-influencers/social-media-endorsements-being-transparent-with-your-followers

Contestabile, G. (2018, January 15). Influencer Marketing in 2018: Becoming an Efficient Marketplace. *Adweek*. Retrieved from https://www.adweek.com/performance-marketing/giordano-contestabile-activate-by-bloglovin-guest-post-influencer-marketing-in-2018/

Cooper, P. (2021, February 10). How the Facebook Algorithm Works in 2021 and How to Make it Work for You. *Hootsuite*. Retrieved from https://blog.hootsuite.com/facebook-algorithm/

Corporate Finance Institute. (n.d.). *Engagement Rate - The level of engagement generated from a created content or a brand campaign*. Retrieved from https://corporatefinanceinstitute.com/resources/knowledge/ecommerce-saas/engagement-rate/

De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. *International Journal of Advertising, 36*(5), 798-828. doi:10.1080/02650487.2017.1348035

Diehl, M. (1990). The Minimal Group Paradigm: Theoretical Explanations and Empirical Findings. *European Review of Social Psychology, 1*(1), 263–292. doi:10.1080/14792779108401864

Digital Marketing Institute. (2018, October 25). 20 Surprising Influencer Marketing Statistics. *Digital Marketing Institute*. Retrieved from https://digitalmarketinginstitute.com/blog/20-influencer-marketing-statistics-that-will-surprise-you

Driessens, O. (2012). The Celebritization of Society and Culture: Understanding the Structural Dynamics of Celebrity Culture. *International Journal of Cultural Studies, 16*(6), 641–657. doi:10.1177/1367877912459140

Driessens, O. (2013). Celebrity Capital: Redefining Celebrity Using Field Theory. *Theory and Society, 42*(5), 543–560. doi:10.1007/s11186-013-9202-3

Droesch, B. (2019, April 26). More Than Half of US Social Network Users Will Be Mobile-Only in 2019. *Insider Intelligence.* Retrieved from https://www.emarketer.com/content/more-than-half-of-social-network-users-will-be-mobile-only-in-2019

Duermyer, R. (2021, January 8). *What Is Blogging?* Retrieved from https://www.thebalancesmb.com/blogging-what-is-it-1794405

Duffy, B. E., & Hund, E. (2015). “Having it All” on Social Media: Entrepreneurial Femininity and Self-Branding Among Fashion Bloggers. *Social Media + Society, 1*(2), 1-11. doi:10.1177/2056305115604337

Duffy, B. E., & Pooley, J. (2019). Idols of Promotion: The Triumph of Self-Branding in an Age of Precarity. *Journal of Communication, 69*(1), 26-48. doi:10.1093/joc/jqy063

Durante, K. M., & Griskevicius, V. (2016). Evolution and Consumer Behavior. *Current Opinion in Psychology, 10*, 27-32. doi:10.1016/j.copsyc.2015.10.025

Entwistle, J. (2017, October). *Influencers vs. Ambassadors. Which is right for your campaign?* Retrieved from https://www.originoutside.com/insights/influencers-vs-ambassadors-which-is-right-for-your-campaign

Erin. (2020, January 16). Instagram Followers: How Many Does the Average Person Have? Retrieved from https://www.hashtagsforlikes.co/blog/instagram-followers-how-many-does-the-average-person-have/

Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology, 59*, 255-278. doi:10.1146/annurev.psych.59.103006.093629

Evans, D. S. (2009). The Online Advertising Industry: Economics, Evolution, and Privacy. *Journal of Economic Perspectives, 23*(3), 37-60. doi:10.1257/jep.23.3.37

Evans, N. J., Phua, J., Lim, J., & Jun, H. (2017). Disclosing Instagram Influencer Advertising: The Effects of Disclosure Language on Advertising Recognition, Attitudes, and Behavioral Intent. *Journal of Interactive Advertising, 17*(2), 138-149. doi:10.1080/15252019.2017.1366885

Evans, N. J., Hoy, M. G., & Childers, C. C. (2018). Parenting “YouTube Natives”: The Impact of Pre-Roll Advertising and Text Disclosures on Parental Responses to Sponsored Child Influencer Videos. *Journal of Advertising, 47*(4), 326-346. doi:10.1080/00913367.2018.1544952

Facebook. (n.d.). *About Facebook*. Retrieved from https://about.fb.com/company-info/

Facebook for Business. (n.d.a.). *About Instagram Insights*. Retrieved from https://www.facebook.com/business/help/788388387972460

Facebook for Business. (n.d.b.). *Instagram Shopping*. Retrieved from https://business.instagram.com/shopping

Fairchild, C. (2007). Building the Authentic Celebrity: The “Idol” Phenomenon in the Attention Economy. *Popular Music and Society, 30*(3), 355–375. doi:10.1080/03007760600835306

Faraggi, D., & Reiser, B. (2002). Estimation of the Area Under the ROC Curve. *Statistics in Medicine, 21*(20), 3093–3106. doi:10.1002/sim.1228

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry, 26*(2), 105-109. doi:10.3969/j.issn.1002-0829.2014.02.009

Folkvord, F., Roes, E., & Bevelander, K. (2020). Promoting healthy foods in the new digital era on Instagram: An experimental study on the effect of a popular real fit influencer versus a fictitious fit influencer on brand attitude and purchase intentions. *BMC Public Health*, *20*(1), 1677-1685. doi:0.1186/s12889-020-09779-y

Frost, J. (n.d.). *7 Classical Assumptions of Ordinary Least Squares (OLS) Linear Regression*. Retrieved from https://statisticsbyjim.com/regression/ols-linear-regression-assumptions/

Fry, R. (2018, April 11). Millennials are the largest generation in the U.S. labor force. *Pew Research Center.* Retrieved from https://www.pewresearch.org/fact-tank/2018/04/11/millennials-largest-generation-us-labor-force/

Federal Trade Commission. (n.d.). *Advertising FAQ's: A Guide For Small Business*. Retrieved from https://www.ftc.gov/tips-advice/business-center/guidance/advertising-faqs-guide-small-business

Gandini, A. (2016). Digital work: Self-branding and Social Capital in the Freelance Knowledge Economy. *Marketing Theory*, *16*(1), 123-141. doi:10.1177/1470593115607942

Ge, J., & Gretzel, U. (2018). Emoji rhetoric: A social media influencer perspective. *Journal of Marketing Management, 34*(15-16), 1272-1295. doi:10.1080/0267257x.2018.1483960

Gleasure, R. (2015). Conceptual Design Science Research? How and Why Untested Meta-Artifacts Have a Place in IS. In B. Donnellan, M. Helfert, J. Kenneally, D. VanderMeer, M. Rothenberger, & R. Winter (Eds.), *New Horizons in Design Science: Broadening the Research Agenda* (pp. 99-114). Cham, Switzerland: Springer. doi:10.1007/978-3-319-06701-8_7

Gleasure, R. (2019). Circadian Rhythms and Social Media Information-Sharing. In F. D. Davis, R. Riedl, J. vom Brocke, P.-M. Léger, A. Randolph, & T. Fischer (Eds.), *Information*

Gleasure, R. & Morgan, L. (2017). The Pastoral Crowd: Exploring Self-hosted Crowdfunding Using Activity Theory and Social Capital. *Information Systems Journal, 28*(3), 489–515. doi:10.1111/isj.12143

Systems and Neuroscience - NeuroIS Retreat 2019 (pp. 1-11). Cham, Switzerland: Springer. doi:10.1007/978-3-030-28144-1_1.

Goldstein, J. W., & Rosenfeld, H. M. (1969). Insecurity and preference for persons similar to oneself. *Journal of Personality, 37(2)*, 253-268. doi:10.1111/j.1467-6494.1969.tb01743.x

Government of Canada. (2019, December 19). Influencer marketing: businesses and influencers must be transparent when advertising on social media. *Competition Bureau Canada*. Retrieved from https://www.canada.ca/en/competition-bureau/news/2019/12/influencer-marketing-businesses-and-influencers-must-be-transparent-when-advertising-on-social-media.html

Haenlein, M., Anadol, E., Farnsworth, T., Hugo, H., Hunichen, J., & Welte, D. (2020). Navigating the New Era of Influencer Marketing: How to be Successful on Instagram, TikTok, & Co., *California Management Review, 63*(1), 5-25. doi:10.1177/0008125620958166

Hale, J. (2020, February 11). Canadian Government Officially Warns 100 Brands And Ad Agencies To Comply With Influencer Marketing Laws. *Tubefilter.* Retrieved from https://www.tubefilter.com/2020/02/10/canada-competition-bureau-influencer-marketing-laws-in-canada-ad-standards/

Hampton, A. J., Fisher Boyd, A. N., & Sprecher, S. (2019). You're like me and I like you: Mediators of the similarity–liking link assessed before and after a getting-acquainted social interaction. *Journal of Social and Personal Relationships, 36*(7), 2221-2244. doi:10.1177/0265407518790411

Hayes, D. (2018, April 8). A Rousing History of Influencer Marketing (We Promise). *The Shelf.* Retrieved from https://www.theshelf.com/the-blog/influencer-marketing-timeline

Hearn, A. (2008). Meat, Mask, Burden. Probing the Contours of the Branded Self. *Journal of Consumer Culture 8(*2), 197-217. doi:10.1177/1469540508090086

Hearn, A. (2011). Confessions of a Radical Eclectic: Reality Television, Self-Branding, Social Media, and Autonomist Marxism. *Journal of Communication Inquiry, 35*(4), 313-321. doi:10.1177/0196859911417438

Hearn, A., & Schoenhoff, S. (2016). From Celebrity to Influencer: Tracing the Diffusion of Celebrity Value across the Data Stream. In P. D. Marshall and S. Redmond (Eds.), *A Companion to Celebrity* (pp. 194-212). Chichester, UK: John Wiley & Sons.

Heinonen, K., & Medberg, G. (2018). Netnography as a tool for understanding customers: implications for service research and practice. *Journal of Services Marketing, 32*(6), 657-679. doi:10.1108/JSM-08-2017-0294

Hogg, M. A. (2016). Social Identity Theory. In S. McKeown, R. Haji, & N. Ferguson (Eds.), *Understanding Peace and Conflict Through Social Identity Theory* (pp. 3-17). Switzerland: Springer International Publishing. doi:10.1007/978-3-319-29869-6_1

Htoon, K. S. (2020, February 29). Log Transformation: Purpose and Interpretation. *Medium*. Retrieved from https://medium.com/@kyawsawhtoon/log-transformation-purpose-and-interpretation-9444b4b049c9

Influencer Marketing Hub. (2021a, February 14). *The State of Influencer Marketing 2020: Benchmark Report.* Retrieved from https://influencermarketinghub.com/influencer-marketing-benchmark-report-2020/

Influencer Marketing Hub. (2021b, January 21). *Influencer Rates: How Much do Influencers Really Cost in 2021?* Retrieved from https://influencermarketinghub.com/influencer-rates/

Instagram. (n.d.a.). *About Us*. Retrieved from https://www.instagram.com/about/us/

Instagram. (n.d.b.). *Features*. Retrieved from https://about.instagram.com/features

Instagram. (n.d.c.). *Search & Explore*. Retrieved from https://about.instagram.com/features/search-and-explore

Instagram. (n.d.d.). *Shopping*. Retrieved from https://about.instagram.com/features/shopping

Ismail, K. (2018, December 10). Social Media Influencers: Mega, Macro, Micro or Nano. *CMS Wire.* Retrieved from https://www.cmswire.com/digital-marketing/social-media-influencers-mega-macro-micro-or-nano/

Izard, C. E., & Smith, M. B. (1960). Personality similarity and friendship. *Journal of Abnormal and Social Psychology, 61*(1), 47-51. doi:10.1037/h0040056

Izea. (2020, September 4). *Influencer tier guide for marketers: What different levels of influencers can achieve for your brand.* Retrieved from https://izea.com/2020/09/04/influencer-tier-guide-for-marketers-what-different-levels-of-influencers-can-achieve-for-your-brand/

Jin, S. V., Muqaddam, A., & Ryu, E. (2019). Instafamous and social media influencer marketing. *Marketing Intelligence & Planning, 37*(5), 567-579. doi:10.1108/mip-09-2018-0375

Jones, E. E., Bell, L., & Aronson, E. (1972). The reciprocation of attraction from similar and dissimilar others: A study in person perception and evaluation. In C. G. McClintock (Ed.), *Experimental Social Psychology* (pp. 142-179). New York, NY: Holt, Rinehart.

Josiassen, A. (2011). Consumer Disidentification and Its Effects on Domestic Product Purchases: An Empirical Investigation in the Netherlands. *Journal of Marketing, 75*(2), 124-140. doi:10.1509/jmkg.75.2.124

Juhász, A. (2020, September 19). Is Infinite Scroll the Slot Machine of the New Generation? *Medium*. Retrieved from https://medium.com/swlh/is-infinite-scroll-the-slot-machine-of-the-new-generation-eb6acd8fefa9

Jun, S., & Yi, J. (2020). What makes followers loyal? The role of influencer interactivity in building influencer brand equity. *Journal of Product & Brand Management, 29*(6), 803-814. doi:10.1108/jpbm-02-2019-2280

Kádeková, Z., & Holienčinová, M. (2018). Influencer Marketing As A Modern Phenomenon - Creating A New Frontier of Virtual Opportunities. *Communication Today, 9*(2), 90-105.

Ketay, S., Beck, L. A., & Welker, K. M. (2020). Self-expansion: Intergroup and Sociocultural Factors. In B. Mattingly, K. McIntyre, & J. G. Lewandowski (Eds.), *Interpersonal Relationships and the Self-Concept* (pp. 177-193). Cham, Switzerland: Springer. doi:10.1007/978-3-030-43747-3_11

Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, 'micro-celebrity' and the rise of Social Media Influencers. *Celebrity Studies, 8*(2), 191-208. doi:10.1080/19392397.2016.1218292

Khanafiah, D., & Situngkir, H. (2004). Social Balance Theory - Revisiting Heider's Balance Theory for many agents. *Industrial Organization.* Retrieved from https://arxiv.org/pdf/nlin/0405041.pdf

Kircher, M. M. (2020, June 03). Instagram Should Give Everyone Swipe Up. *Vulture*. Retrieved from https://www.vulture.com/2020/06/how-can-i-get-instagram-swipe-up.html

Koernig, S. K., & Boyd, T. C. (2009). To Catch a Tiger or Let Him Go: The Match-Up Effect and Athlete Endorsers for Sport and Non-Sport Brands. *Sport Marketing Quarterly, 18*(1), 25-37.

Konstantopoulou, A., Rizomyliotis, I., Konstantoulaki, K., & Badahdah, R. (2019). Improving SMEs' competitiveness with the use of Instagram influencer advertising and eWOM. *International Journal of Organizational Analysis, 27*(2), 308–321. doi:10.1108/ijoa-04-2018-1406

Kowalczyk, C. M., & Pounders, K. R. (2016). Transforming celebrities through social media: The role of authenticity and emotional attachment. *Journal of Product & Brand Management, 25*(4), 345–356. doi:10.1108/JPBM-09-2015-0969

Kozinets, R. V. (2006). Click to Connect: Netnography and Tribal Advertising. *Journal of Advertising Research, 46*(3), 279-288. doi:10.2501/S0021849906060338

Kozinets, R. V. (2010). *Netnography: Doing Ethnographic Research Online*. Thousand Oaks, CA: Sage.

Kozinets, R. V., de Valck, K., Wojnicki, A. C., & Wilner, S. J. S. (2010). Networked Narratives: Understanding Word-of-Mouth Marketing in Online Communities. *Journal of Marketing, 74*(2), 71-89. doi:10.1509/jmkg.74.2.71

Kumpumäki, S. (n.d.). *AI Influencers: Is This The Future of Influencer Marketing?* Retrieved from https://blog.beatly.com/en/blog/ai-influencers-future-influencer-marketing

Laudon, K. C., & Traver, C. G. (2019). *E-commerce 2018* (14th ed.). Harlow, England: Pearson.

Lee, D. (2020, February 12). The FTC is cracking down on influencer marketing on YouTube, Instagram, and TikTok. *The Verge.* Retrieved from: https://www.theverge.com/2020/2/12/21135183/ftc-influencer-ad-sponsored-tiktok-youtube-instagram-review

Lewandowski, G. W., Nardone, N., & Raines, A. J. (2010). The Role of Self-concept Clarity in Relationship Quality. *Self and Identity, 9(4)*, 416-433. doi:10.1080/15298860903332191

LIWC. (n.d.a). *Interpreting LIWC Output*. Retrieved from https://liwc.wpengine.com/interpreting-liwc-output/

LIWC. (n.d.b.). *How it Works*. Retrieved from http://liwc.wpengine.com/how-it-works/

Locke, T. (2019, November 11). 86% of young people say they want to post social media content for money. *CNBC.* Retrieved from https://www.cnbc.com/2019/11/08/study-young-people-want-to-be-paid-influencers.html

Lu, L.-C., Chang, W.-P., & Chang, H.-H. (2014). Consumer attitudes toward blogger's sponsored recommendations and purchase intention: The effect of sponsorship type, product type, and brand awareness. *Computers in Human Behavior, 34*, 258-266. doi:10.1016/j.chb.2014.02.007

Lugo, J. (2019, May 9). *The word 'influencer' is now in the Mirriam-Webster dictionary.* Retrieved from https://www.influencerupdate.biz/news/68099/the-word-influencer-is-now-in-the-dictionary/#:~:text=Merriam%2DWebster%20recently%20announced%20it,individuals%20with%20a%20large%20following

Marshall, P. D. (2016). Exposure: The Public Self Explored. In P. D. Marshall and S. Redmond (Eds.), A *Companion to Celebrity* (pp. 497-517). Chichester, UK: John Wiley & Sons.

Martineau, P. (2019, April 22). Inside the Weird, and Booming, Industry of Online Influence. *Wired*. Retrieved from https://www.wired.com/story/inside-the-industry-social-media-influence/

Marwick, A. E. (2013). *Status Update: Celebrity, Publicity, and Branding in the Social Media Age*. New Haven, CT: Yale University Press.

Marwick, A. E. (2015). Instafame: Luxury Selfies in the Attention Economy. *Public Culture, 27*(1-75), 137-160. doi:10.1215/08992363-2798379

Mathew, J. (2018, July 30). Understanding Influencer Marketing And Why It Is So Effective. *Forbes.* Retrieved from https://www.forbes.com/sites/theyec/2018/07/30/understanding-influencer-marketing-and-why-it-is-so-effective/?sh=640ac27f71a9

McCoole, V. (2018, July 29). Behind the Scenes of Instagram's Million-Dollar Influencer Brand Deals. *Forbes*. Retrieved from https://www.forbes.com/sites/veenamccoole/2018/07/29/behind-the-scenes-of-instagrams-million-dollar-influencer-brand-deals/

McQuarrie, E. F., & Phillips, B. J. (2014). The Megaphone Effect in Social Media: How Ordinary Consumers Become Style Leaders. *GfK Marketing Intelligence Review, 6*(2), 16–20. doi:10.2478/gfkmir-2014-0092

Mediakix. (n.d.). Influencer tiers for the influencer marketing industry. *Mediakix.* Retrieved from https://mediakix.com/influencer-marketing-resources/influencer-tiers/

Mediakix. (2021, February 23). How Many Influencers Are There? New Data, Calculations & Charts. *Mediakix.* Retrieved from https://mediakix.com/blog/how-many-influencers-are-there/

Mehta, P. D. (2001). Control Variable in Research. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (1st ed., pp. 2727-2730). USA: Elsevier.

Menard, S. (2001). *Applied Logistic Regression Analysis* (1st ed.). Los Angeles, CA: Sage.

Montoya, R. M., & Horton, R. S. (2004). On the Importance of Cognitive Evaluation as a Determinant of Interpersonal Attraction. J*ournal of Personality and Social Psychology, 86*(5), 696-712. doi:10.1037/0022-3514.86.5.696

Montoya, R. M., & Horton, R. S. (2012). A meta-analytic investigation of the processes underlying the similarity-attraction effect. *Journal of Social and Personal Relationships*, *30*(1), 64-94. doi:10.1177/0265407512452989

Morning Consult. (2019, November). The Influencer Report Engaging Gen Z and Millennials. *Morning Consult.* Retrieved from https://morningconsult.com/wp-content/uploads/2019/11/The-Influencer-Report-Engaging-Gen-Z-and-Millennials.pdf

Mudambi, S. M., & Schuff, D. (2010). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly, 34*(1), 185-200. doi: 10.2307/20721420

Nahemow, L., & Lawton, M. P. (1975). Similarity and propinquity in friendship formation. *Journal of Personality and Social Psychology, 32*(2)*,* 205-213. doi:10.1037/0022-3514.32.2.205

Nielsen, K. (2015). ‘‘Fake It ’til You Make It’’: Why Community College Students’ Aspirations ‘‘Hold Steady’’. *Sociology of Education, 88*(4), 265-283. doi:10.1177/0038040715601889

Olafson, K. (2020, October 07). How to Use Hashtags: A Quick and Simple Guide for Every Network. *Hootsuite*. Retrieved from https://blog.hootsuite.com/how-to-use-hashtags/

O’Reilly, T. (2005). What is Web 2.0. *Design Patterns and Business Models for the Next Generation of Software*. Retrieved from https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

Park, J., Ciampaglia, G. L., & Ferrara, E. (2016). Style in the Age of Instagram: Predicting Success within the Fashion Industry using Social Media. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 64-73. doi:10.1145/2818048.2820065

Park, L. E., & Crocker, J. (2008). Contingencies of self-worth and responses to negative interpersonal feedback. *Self and Identity, 7(2),* 184-203. doi:10.1080/15298860701398808

Pathak, S. (2015, August 11). Why fashion and beauty brands love Instagram. *Digiday*. Retrieved from https://digiday.com/marketing/fashion-beauty-brands-instagram/

Pedroni, M. (2015). “Stumbling on the Heels of My Blog”: Career, Forms of Capital, and Strategies in the (Sub)Field of Fashion Blogging. *Fashion Theory, 19*(2), 179-199. doi:10.2752/175174115X14168357992355

Petroni, S. (2019). How Social Media Shape Identities and Discourses in Professional Digital Settings: Self-Communication or Self-Branding? In P. Bou-Franch & P. Garcés-Conejos Blitvich (Eds.), *Analyzing Digital Discourse - New Insights and Future Directions,* (pp. 251-281). London, UK: Palgrave Macmillan. doi:10.1007/978-3-319-92663-6_9

Perlmutter, D. D. (2005). The Ascent of Blogging. *Nieman Foundation at Harvard*. Retrieved from https://nieman.harvard.edu/articles/the-ascent-of-blogging/

Perrin, A., & Anderson, M. (2019, April 10). Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. *Pew Research Center*. Retrieved from https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/

Peters, T. (1997). The Brand Called You. *Fast Company*. Retrieved from https://www.fastcompany.com/28905/brand-called-you

Podlegal. (n.d.). *Disclosure Requirements for Bloggers and Internet Advertisers: A lesson from Australia Post*. Retrieved from https://www.podlegal.com.au/disclosure-requirements-for-bloggers-and-internet-advertisers-a-lesson-from-australia-post/

Powell-Brown, A. (2004). Can You Be a Teacher of Literacy If You Don't Love to Read? *Journal of Adolescent & Adult Literacy, 47*(4), 284-288.

Purtill, J. (2017, March 02). Instafamous must reveal #ads under new transparency rules. *ABC.* Retrieved from https://www.abc.net.au/triplej/programs/hack/social-influencers-must-reveal-ad-under-new-transparency-rules/8315962

Pusztai, H.-M. (2019, July 23). *Why Micro-Influencers May Be the Most Effective Influencer Marketing Strategy*. Retrieved from https://buffer.com/resources/micro-influencers/

Ratner, B. (2009). The correlation coefficient: Its values range between +1/−1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing, 17*(2), 139-142. doi:10.1057/jt.2009.5

Raun, T. (2018). Capitalizing intimacy: New subcultural forms of micro- celebrity strategies and affective labour on YouTube. *Convergence: The International Journal of Research into New Media Technologies, 24*(1), 99-113. doi:10.1177/1354856517736983

Russell, C. A., & Stern, B. B. (2006). CONSUMERS, CHARACTERS, AND PRODUCTS: A Balance Model of Sitcom Product Placement Effects. *Journal of Advertising, 35*(1), 7-21. doi:10.2753/JOA0091-3367350101

Schrag, F. (1992). In Defense of Positivist Research Paradigms. *Educational Researcher, 21*(5), 5-8. doi:10.3102/0013189X021005005

Sehl, K. (2021, January 4). 10 Unique Ways to Use Instagram Guides for Marketing. *Hootsuite.* Retrieved from https://blog.hootsuite.com/instagram-guides/

Senft, T. (2013). Microcelebrity and the Branded Self. In J. Hartley, J. Burgess, & A. Bruns (Authors), *A Companion to New Media Dynamics* (1st ed., pp. 346-354). Chichester, UK: Wiley-Blackwell.

Shepherd, I. D. H. (2005). From Cattle and Coke to Charlie: Meeting the Challenge of Self Marketing and Personal Branding. *Journal of Marketing Management, 21*(5-6), 589–606. doi:10.1362/0267257054307381

Shiller, C. (2020, January 9). Hashtags: Are they a thing of the past or are they still relevant in social media in 2020? *Columbus*. Retrieved from https://www.advertisecolumbus.com/blog/hashtags_2020

Shimp, T. A., & Sharma, S. (1987). Consumer Ethnocentrism: Construction and Validation of the CETSCALE. *Journal of Marketing Research, 24*(3), 280-289. doi:10.2307/3151638

Shrestha, N. (2020). Detecting Multicollinearity in Regression Analysis. *American Journal of Applied Mathematics and Statistics, 8*(2), 39-42. doi:10.12691/ajams-8-2-1

Sigall, H. (1970). Effects of competence and consensual validation on a communicator's liking for the audience. *Journal of Personality and Social Psychology, 16*(2), 251-258. doi:10.1037/h0030603

Siles, I. (2011). The rise of blogging: Articulation as a dynamic of technological stabilization. *New Media & Society, 14*(5), 781-797. doi:10.1177/1461444811425222

Sincerelyjules (2021, March 8). *Instagram.* Retrieved from https://www.instagram.com/p/CMLDpSrJf-R/?igshid=wqzpvj78iqb3

Singh, R., Ho, L., Tan, H. L., & Bell, P. A. (2007). Attitudes, personal evaluations, cognitive evaluation and interpersonal attraction: On the direct, indirect and reverse-causal effects. *British Journal of Social Psychology, 46(1)*, 19-42. doi:10.1348/014466606x104417

Sloane, G. (2015, September 30). Facebook: Fast-scrolling millennials consume ads 2.5 times faster. *Digiday*. Retrieved from https://digiday.com/media/advertisingweek2015-facebook-finds-that-fast-scrolling-millennials-consume-ads-2-5-times-faster/

Smith, G. (2018). Step away from stepwise. *Journal of Big Data, 5*(1), 1-12. doi:10.1186/s40537-018-0143-6

Sprecher, S., Treger, S., Hilaire, N., Fisher Boyd, A., & Hatfield, E. (2013). You Validate Me, You Like Me, You're Fun, You Expand Me: "I'm Yours!". *Current Research in Social Psychology, 21(5)*, 22-34.

Sprecher, S., Treger, S., Fisher Boyd, A., Hilaire, N., & Grzybowski, M. (2015). Associations Between Self-Expansion and Actual and Perceived (Dis)Similarity and Their Joint Effects on Attraction in Initial Interactions. *Self and Identity, 14*(4), 369–389. doi:10.1080/15298868.2014.1003592

Stubb, C., & Colliander, J. (2019). “This is not sponsored content” – The effects of impartiality disclosure and e- commerce landing pages on consumer responses to social media influencer posts. *Computers in Human Behavior, 98*, 210-222. doi:10.1016/j.chb.2019.04.024

Stubb, C., Nyström, A.-G., & Colliander, J. (2019). Influencer marketing: The impact of disclosing sponsorship compensation justification on sponsored content effectiveness. *Journal of Communication Management, 23*(2), 109–122. doi:10.1108/jcom-11-2018-0119

Suciu, P. (2019, December 26). Is Instagram The Social Media Service For Business In 2020? *Forbes.* Retrieved from https://www.forbes.com/sites/petersuciu/2019/12/26/is-instagram-the-social-media-service-for-business-in-2020/?sh=6ac5ea9d3bdf

Tankovska, H. (2021a). Number of social network users worldwide from 2017 to 2025. *Statista.* Retrieved from https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/#:~:text=Social%20media%20usage%20is%20one,almost%204.41%20billion%20in%202025

Tankovska, H. (2021b). Distribution of Instagram users worldwide as of January 2021, by age group. *Statista.* Retrieved from https://www.statista.com/statistics/325587/instagram-global-age-group/

Tajfel, H. (1972). Social Categorization. English Manuscript of “La catégorisation sociale”. In S. Moscovici (Ed.), *Introduction à la Psychologie Sociale* (1st ed., pp. 272–302). Paris, France: Larousse.

Tajfel, H., & Turner, J. C. (1979). An Integrative Theory of Inter-Group Conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Inter-Group Relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Thompson, C. B., & Panacek, E. A. (2007). Research study designs: Non-experimental. *Air Medical Journal, 26*(1), 18-22. doi:10.1016/j.amj.2006.10.003

Thomson, M. (2006). Human Brands: Investigating Antecedents to Consumers' Strong Attachments to Celebrities. *Journal of Marketing, 70*(3), 104-119. doi:10.1509/jmkg.70.3.104

Thompson, S. K. (2012). *Sampling* (3rd ed.). Hoboken, N.J.: Wiley.

Trammell, K. D., & Keshelashvili, A. (2005). Examining the New Influencers: A Self-Presentation Study of A-List Blogs. *Journalism & Mass Communication Quarterly, 82*(4), 968-982. doi:10.1177/107769900508200413

Triwidisari, A., Nurkhin, A., & Muhsin, M. (2017). The Relationships Between Instagram Social Media Usage, Hedonic Shopping Motives and Financial Literacy on Impulse Buying. *Dinamika Pendidikan, 12*(2), 170-181. doi:10.15294/DP.V12I2.13565

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford, England: Blackwell.

UCLA. (n.d.). FAQ How Do I Interpret a Regression Model When Some Variables Are Log Transformed? *UCLA Institute for Digital Research & Education*. Retrieved from https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faqhow-do-i-interpret-a-regression-model-when-some-variables-are-log-transformed/

University of Minnesota. (2016). *Research Methods in Psychology*. Twin Cities, MN: University of Minnesota Libraries Publishing. doi:10.24926/866

Van Driel, L., & Dumitrica, D. (2021). Selling brands while staying "Authentic": The professionalization of Instagram influencers. *Convergence: The International Journal of Research into New Media Technologies, 27*(1), 66-84. doi:10.1177/1354856520902136

Van Reijmersdal, E. A., & Van Dam, S. (2020). How Age and Disclosures of Sponsored Influencer Videos Affect Adolescents' Knowledge of Persuasion and Persuasion. *Journal of Youth and Adolescence, 49*(2), 1531-1544. doi:10.1007/s10964-019-01191-z

Vogt, W. P. (2005). *Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences* (3rd ed.). Thousand Oaks, CA: Sage.

Warren, J. (2020a, November 17). Everything You Need to Know About Instagram Guides. *Later.* Retrieved from https://later.com/blog/instagram-guides/

Warren, J. (2020b, April 27). Affiliate Marketing for Influencers: How to Make Money on Instagram. *Later*. Retrieved from https://later.com/blog/affiliate-marketing-for-influencers/

Warren, J. (2021, January 4). This is How the Instagram Algorithm Works in 2021. *Later.* Retrieved from https://later.com/blog/how-instagram-algorithm-works/

Yuan, C. L., Moon, H., Kim, K. H., & Wang, S. (2019). The influence of parasocial relationship in fashion web on customer equity. *Journal of Business Research*, 1-8. doi:10.1016/j.jbusres.2019.08.039

Appendix

Tables 1 & 2

Tabulation, Influencer Category & Influencer Country

Tabulation of Category

Influencer Category	Freq.	Percent
1 (Mega)	4999	20.00
2 (Macro)	4999	20.00
3 (Mid-tier)	5000	20.00
4 (Micro)	5000	20.00
5 (Nano)	5000	20.00
Total	24998	100.00

Tabulation of Country

Influencer Country	Freq.	Percent
1 (US)	12098	48.40
2 (AUS)	6200	24.80
3 (UK)	5200	20.80
4 (CAN)	400	1.60
5 (IE)	400	1.60
6 (NZ)	500	2.00
7 (SIN)	200	0.80
Total	24998	100.00

Table 3

Regression Results (Comment Rate), H1b-H8b

Linear Regression: H1b-H8b

Dep. Variable Indep. Variables	ln_commentrate											
	Main Model		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln_captionlen	.014***	.008***	**.140*****	**.148*****	.013***	.011***	.008***	.010***	.015***	.011***	.013***	.009***
hashtagyesno	-.046***	-	-.049***	-	**.080*****	**.081*****	-	-	-.046***	-	-.047***	-
ln_hashtagcount	-	-.010***	-	-.020***	-	-	**.108*****	**.104*****	-	-.010***	-	-.011***
tagyesno	.	-	.	-	.	-	-	.	**.109*****	**.105*****	-	-
ln_tagcount	-	.	-	.	-	.007*	.	-	-	-	**.118*****	**.121*****
location	.	.	.	.	.	.	.	.	.	.	.	.
qmarkyesno	.	.	.	.	.	.	.	.	.	.	.	.
ln_authentic	.	.	.	.	.	.	.	.	.	.	.	.
ln_clout	.	.	.	.	.	.	.	.	.	.	.	.
sponsored	.038***	.025***	.038***	.029***	.038***	.034***	.030***	.034***	.038***	.032***	.035***	.027***
ln_follower	-.106***	-.106***	-.056***	-.053***	-.102***	-.102***	-.101***	-.101***	-.101***	-.101***	-.102***	-.101***
ln_follower												
× ln_captionlen			-.011***	-.012***								
× hashtagyesno					-.011***	-.011***						
× ln_hashtagcount							-.011***	-.011***				
× tagyesno									-.010***	-.010***		
× ln_tagcount											-.010***	-.011***
× location												
× qmarkyesno												
× ln_authentic												
× ln_clout												
× sponsored												
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.414	0.412	0.422	0.420	0.415	0.415	0.415	0.415	0.415	0.414	0.415	0.414

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant
Control variable *country* was also included, however not reported due to brevity and relevance

Table 4

Regression Results (Comment Rate), H1b-H8b (continued)

Linear Regression: H1b-H8b

Dep. Variable / Indep. Variables	ln_commentrate (cont.)											
	Main Model		(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
ln_captionlen	.014***	.008***	.015***	.008***	.014***	.008***	.015***	.008***	.014***	.008***	.014***	.008***
hashtagyesno	-.046***	-	-.046***	-	-.045***	-	-.045***	-	-.046***	-	-.046***	-
ln_hashtagcount	-	-.010***	-	-.010***	-	-.010***	-	-.008***	-	-.010***	-	-.010***
tagyesno	.	-	.	-	.	-	.	-	.	-	.	-
ln_tagcount	-	.	-	.	-	.	-	.	-	.	-	.
location	.	.	**.338*** **	**.336*** **	.	.	.	.	.	.	.	.
qmarkyesno	.	.	.	.	**.104*** **	**.115*** **	.	.	.	.	.	.
ln_authentic	.	.	.	.	.	.	**.070*** **	**.072*** **	.	.	.	.
ln_clout	.	.	.	.	.	.	.	.	**.023** **	**.025** **	.	.
sponsored	.038***	.025***	.037***	.024***	.037***	.025***	.038***	.025***	.038***	.025***	**.128*** **	**.125*** **
ln_follower	-.106***	-.106***	-.104***	-.104***	-.105***	-.105***	-.088***	-.088***	-.099***	-.099***	-.104***	-.104***
ln_follower												
× ln_captionlen												
× hashtagyesno												
× ln_hashtagcount												
× tagyesno												
× ln_tagcount												
× location			-.031***	-.030***								
× qmarkyesno					-.009***	-.010***						
× ln_authentic							-.006***	-.006***				
× ln_clout									-.002*	-.002**		
× sponsored											-.008***	-.009***
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.414	0.412	0.416	0.414	0.415	0.413	0.417	0.415	0.415	0.413	0.415	0.413

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant
Control variable *country* was also included, however not reported due to brevity and relevance

Table 5

Regression Results (Engagement Rate), H1b-H8b

Linear Regression: H1b-H8b

Dep. Variable / Indep. Variables	ln_engagementrate											
	Main Model		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln_captionlen	-.023***	-.024***	**.288*** **	**.373*** **	-.029***	-.025***	-.023***	-.026***	-.022***	-.022***	-.017***	-.019***
hashtagyesno	-.246***	-	-.253***	-	**.396*** **	**.408*** **	-	-	-.246***	-	-.255***	-
ln_hashtagcount	-	-.094***	-	-.122***	-	-	**.410*** **	**.384*** **	-	-.095***	-	-.098***
tagyesno	-.130***	-	-.112***	-	-.120***	-	-	-.137***	**.268*** **	**.246*** **	-	-
ln_tagcount	-	-.128***	-	-.125***	-	-.104***	-.115***	-	-	-	**.316*** **	**.349*** **
location	-.055***	-.054***	-.055***	-.053***	-.058***	-.059***	-.052***	-.051***	-.057***	-.055***	-.057***	-.055***
qmarkyesno	-.047***	-.049***	-.056***	-.063***	-.044***	-.044***	-.046***	-.047***	-.050***	-.054***	-.052***	-.057***
ln_authentic	.010***	.007**	.015***	.009***	.013***	.012***	.010***	.010***	.011***	.008***	.009***	.006**
ln_clout	.011**	.	.010**	.	.011**	.009*	.	.009*	.010**	.010*	.009*	.
sponsored	.071***	.019*	.072***	.032***	.075***	.063***	.042***	.058***	.072***	.046***	.063***	.029***
ln_follower	-.106***	-.108***	.017***	.042***	-.087***	-.087***	-.085***	-.087***	-.090***	-.091***	-.090***	-.090***
ln_follower												
× ln_captionlen			-.027***	-.034***								
× hashtagyesno					-.055***	-.057***						
× ln_hashtagcount							-.048***	-.046***				
× tagyesno									-.034***	-.035***		
× ln_tagcount											-.038***	-.043***
× location												
× qmarkyesno												
× ln_authentic												
× ln_clout												
× sponsored												
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.163	0.150	0.177	0.171	0.171	0.170	0.168	0.169	0.166	0.157	0.166	0.156

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant
Control variable *country* was also included, however not reported due to brevity and relevance

Table 6

Regression Results (Engagement Rate), H1b-H8b (continued)

Linear Regression: H1b-H8b

Dep. Variable Indep. Variables	**ln_engagementrate (cont.)**											
	Main Model		(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
ln_captionlen	-.023***	-.024***	-.023***	-.023***	-.023***	-.024***	-.022***	-.024***	-.025***	-.025***	-.023***	-.023***
hashtagyesno	-.246***	-	-.246***	-	-.245***	-	-.244***	-	-.247***	-	-.245***	-
ln_hashtagcount	-	-.094***	-	-.094***	-	-.094***	-	-.091***	-	-.096***	-	-.095***
tagyesno	-.130***	-	-.130***	-	-.130***	-	-.129***	-	-.128***	-	-.128***	-
ln_tagcount	-	-.128***	-	-.128***	-	-.129***	-	-.128***	-	-.127***	-	-.128***
location	-.055***	-.054***	**.350***	**.333***	-.056***	-.055***	-.056***	-.055***	-.055***	-.054***	-.057***	-.056***
qmarkyesno	-.047***	-.049***	-.048***	-.050***	.	**.175***	-.051***	-.054***	-.046***	-.048***	-.048***	-.051***
ln_authentic	.010***	.007**	.010***	.007**	.010***	.007**	**.128***	**.124***	.010***	.007**	.010***	.007**
ln_clout	.011**	.	.011**	.	.011**	.	.010*	.	**.191***	**.204***	.011**	.
sponsored	.071***	.019*	.070***	.019*	.071***	.020*	.071***	.020*	.071***	.020*	**.294***	**.317***
ln_follower	-.106***	-.108***	-.104***	-.106***	-.104***	-.106***	-.077***	-.079***	-.049***	-.046***	-.102***	-.103***
ln_follower												
× ln_captionlen												
× hashtagyesno												
× ln_hashtagcount												
× tagyesno												
× ln_tagcount												
× location			-.037***	-.036***								
× qmarkyesno					-.012**	-.020***						
× ln_authentic							-.010***	-.010***				
× ln_clout									-.015***	-.017***		
× sponsored											-.019***	-.026***
N	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998	24‘998
R-squared	0.163	0.150	0.164	0.151	0.163	0.151	0.166	0.153	0.165	0.152	0.164	0.151

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation “.”: not statistically significant

Control variable *country* was also included, however not reported due to brevity and relevance

Table 7

Correlations Matrix, Robustness

Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) ln_captionlen	1.000												
(2) hashtagyesno	0.437	1.000											
(3) ln_hashtagcount	0.465	0.808	1.000										
(4) tagyesno	0.333	0.396	0.273	1.000									
(5) ln_tagcount	0.383	0.365	0.298	0.848	1.000								
(6) location	0.008	0.008	0.020	0.016	0.015	1.000							
(7) qmarkyesno	0.278	0.134	0.134	0.055	0.070	0.008	1.000						
(8) ln_authentic	0.316	0.036	-0.020	0.031	0.029	-0.002	0.093	1.000					
(9) ln_clout	0.239	0.136	0.137	0.128	0.126	-0.009	0.087	-0.147	1.000				
(10) sponsored	0.355	0.362	0.270	0.452	0.402	0.011	0.050	0.050	0.099	1.000			
(11) ln_follower	-0.178	-0.117	-0.215	-0.049	-0.073	-0.095	-0.052	-0.005	0.001	-0.033	1.000		
(12) ln_posemo	0.366	0.137	0.136	0.107	0.122	-0.004	0.085	0.100	0.161	0.104	-0.053	1.000	
(13) ln_negemo	0.119	0.010	-0.004	0.006	0.001	-0.004	0.046	0.055	-0.039	0.001	-0.035	-0.003	1.000

Table 8

Regression Results, H6a-H7a Robustness

Linear Regression: H6a-H7a Robustness

Dep. Variable / Indep. Variable	ln_posemo	ln_negemo
	Coef.	Coef.
ln_follower	-.025***	-.009***
N	24'998	24'998
R-squared	0.003	0.001

Table 9

Regression Results, Instrumental Variable

Instrumental Variables (2SLS) Regression: Robustness

Dep. Variable / Indep. Variables	ln_likerate	ln_commentrate	ln_engagementrate
	Coef.	Coef.	Coef.
ln_captionlen	-.075***	-.018***	-.069***
hashtagyesno	-.247***	-.065***	-.256***
tagyesno	-.12***	.	-.119***
location	-.172***	-.069***	-.17***
qmarkyesno	-.044***	.	-.039***
ln_authentic	.026***	.01***	.025***
ln_clout	.031***	.015***	.031***
sponsored	.084***	.061***	.098***
ln_follower	-.19***	- .182***	-.215***
N	24'998	24'998	24'998
R-squared	.	0.169	.
Tests of endogeneity: H0: variables are exogenous			
Durbin (score) chi2(1)	602.246 (p = 0.0000)	1252.95 (p = 0.0000)	617.786 (p = 0.0000)
Wu-Hausman F(1,24987)	616.841 (p = 0.0000)	1318.48 (p = 0.0000)	633.162 (p = 0.0000)

*** p<.01, ** p<.05, * p<.1 "."": not statistically significant

Table 10

Regression Results, Robustness

Linear Regression: Robustness

Dep. Variable / **Indep. Variables**	ln_follower (1)	ln_follower (2)
	Coef.	Coef.
ln_captionlen	-.414***	-.286***
hashtagyesno	-.405***	-
ln_hashtagcount	-	-.424***
tagyesno	.237***	-
ln_tagcount	-	.104***
location	-.743***	-.721***
qmarkyesno	-.194***	-.221***
ln_authentic	.079***	.035***
ln_clout	.116***	.1***
sponsored	.185***	.222***
followingratio	-1.347***	-1.328***
1b.country		
2.country	-1.675***	-1.646***
3.country	-1.062***	-1.095***
4.country	-.439***	-.383***
5.country	-2.181***	-2.188***
6.country	-2.938***	-2.976***
7.country	-1.277***	-1.413***
N	24’998	24’998
R-squared	0.278	0.291

*** p<.01, ** p<.05, * p<.1 “-“: excluded due to correlation