FACULTY OF ARTS AND PHILOSOPHY

How do physical features affect the online credibility of Instagram influencers?

AN EXPERIMENTAL STUDY

Word count: 19.553

Lotte Remue
Student number: 01807990

Supervisor: Prof. dr. Bernard De Clerck

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Arts in Interpreting (Dutch – English – French)

Academic year: 2021 – 2022

UNIVERSITEIT GENT

ABSTRACT

In today's digitised society, electronic word of mouth (eWOM) is no longer limited to traditional review communities. Social media platforms, such as Instagram, are increasingly employed to share opinions and experiences on. As such, brands and businesses approach Instagram users with a considerable number of followers, labeled 'influencers', to post reviews promoting their products. The present study examines which features of these sources increase the perceived credibility of such disguised advertisements, as well as how they affect users' attitude towards, and purchase intention of the product. More in particular, we employ a 2 (attractive vs. less attractive) x 2 (glasses vs. no glasses) experimental design to study the impact of perceived attractiveness and intelligence. First of all, the pre-test showed that a distinction should be made between social and physical attractiveness. Secondly, blemishes and weight gain increased social attractiveness, whereas glasses positively influenced both social attractiveness and perceived intelligence. Perceived intelligence and social attractiveness boosted message credibility, consumer attitude and purchase intention. Physical attractiveness, on the other hand, affected none of the constructs. Limitations and suggestions for future research are discussed.

(181 words)

ACKNOWLEDGEMENTS

Although a challenging task, this paper is without any doubt the most rewarding achievement in my four years at Ghent University. I thoroughly enjoyed conducting this study, and hope that it can be a useful contribution, albeit small, to the literature on online credibility.

First of all, I would like to express my profound gratitude to Prof. Dr. Bernard De Clerck, who guided me through both my Bachelor's paper and Master's thesis, and thanks to whom I discovered my love for academia. At each stage of the writing process, Mr. De Clerck provided me with insightful feedback, leading to many new insights. He supervised the design of the experiment and introduced me to the at first overwhelming, but also very valuable world of statistics. Without his support and supervision, I could have never achieved the same results.

Secondly, I want to thank my friends who stayed the course with me. I especially want to thank Ellen W., my help and stay, Ellen V.L., Judith, Emma, and all of the fellow interpreter students whom I have come to hold very dear this final year. I would like to thank my family as well, who were always curious for updates and who happily shared my survey with their friends and acquaintances.

Last but not least, I am deeply grateful to all of the respondents for their time and effort. Without them, I could have never accomplished this study.

TABLE OF CONTENTS

Word count: 19.553

LIST OF ABBREVIATIONS

ABBREVIATION	MEANING
EWOM	Electronic word of mouth
SM	Social media
MANG	Male, Attractive, No Glasses
MAG	Male, Attractive, Glasses
MLANG	Male, Less Attractive, No Glasses
MLAG	Male, Less Attractive, Glasses
FANG	Female, Attractive, No Glasses
FAG	Female, Attractive, Glasses
FLANG	Female, Less Attractive, No Glasses
FLAG	Female, Less Attractive, Glasses
PA	Physical attractiveness
SA	Social attractiveness
PI	Perceived intelligence
BRA	Brand attitude
PRA	Product attitude
MC	Message credibility
CA	Consumer attitude
PUI	Purchase intention
ADDINFO	Wish for additional information

LIST OF TABLES AND FIGURES

1. INTRODUCTION

In today's digital age, people are increasingly turning to the Internet for advice. Online review communities, such as TripAdvisor or Yelp, have become well established and continue to grow in popularity. Such sites are a viable source of information, as they enable consumers to browse through other people's experiences with the products or services they are interested in.

Social media (SM) are the latest channel for such electronic word of mouth (eWOM). A review left at a restaurant's Facebook page, or an Instagram blogger recommending a book, for example, are no longer exceptions. Similar to experiences shared on the traditional review communities, reviews posted on SM platforms mainly serve to diminish buyers' feelings of uncertainty, such as when they want to visit a new restaurant or purchase from 'anonymous' sellers on the Internet (Cheung, Luo, Sia & Chen, 2009). More often than not, however, the authors of these reviews are strangers themselves. Moreover, cues which might help to form an impression of the source in face-to-face interaction, such as tone of voice or gesticulation, are absent in an online environment. Consumers and SM users therefore tend to search for other cues which might reveal something about how credible these sources are (Walther, 1993; Xu, 2014). Such cues can be found in the review message itself (Cheung et al.; Lybaert, De Clerck, Plevoets, Loete & Depovere, 2020), but also in the appearance of the source (Ozanne, Liu & Mattila, 2019; Krämer & Winter, 2008).

Given the increasingly visual nature of social networking sites, a growing number of studies have been conducted on the impact of the profile pictures accompanying such reviews, concentrating on the specific elements within the photos influencing source credibility (Ozanne et al., 2019; Wang, 2016). As such, previous research has revealed that attractiveness is one of the factors which can improve source credibility (Ozanne et al.; Peng, Cui, Chung & Zheng, 2020). It can produce a 'halo effect', leading people to attribute unrelated characteristics, such as intelligence, sociability or credibility to good-looking people (Batres, 2020, p. 3). Thanks to these higher levels of credibility, attractive reviewers can positively influence consumers' attitude towards a brand or product (Lim, Radzol, Cheah & Wong, 2017; Ozanne et al.), which in turn typically leads to higher purchase intention (Lim et al.).

Fuelled by the need to excel in sales and financial performance, brands and businesses have turned these potentially lucrative reviews to their advantage. Especially on the visually engaging platform of Instagram, they have started approaching users with an important number

of followers, called ‘influencers’, to promote their products disguised as reviews. Although influencers are now obliged to signal sponsored content (Lee & Kim, 2020; Naderer, Matthes & Schäfer, 2021), brands and businesses successfully continue to employ this marketing strategy.

As such, Instagram has changed how brands and businesses interact with their customers. They either engage with followers directly, or target potential buyers through influencers promoting their products. In this way, influencers play an important role in users’ decision-making process, and are generally even more effective than traditional celebrities (Djafarova & Rushworth, 2017; Tafesse & Wood, 2021). Furthermore, the use of influencers in advertising has been linked to higher purchase intention (Balaban & Mustățea, 2019; Djafarova et al.; Lim et al., 2017). However, the role which credibility plays in this particular process remains unclear, as research has yielded ambiguous results (Balaban et al.; Lee et al., 2020; Lim et al.; Rebelo, 2017).

In light of the above, the specific review setting of the ‘sponsored influencer advertisement’ provides fertile soil for future research. Although multiple studies have already been conducted on the impact of physical appearance on consumer behaviour (Ozanne et al., 2019; Peng et al., 2020), there are various reasons why further research is necessary on reviews coinciding with visual stimuli.

First, the impact of physical attractiveness is not without its limits. Research in the eWOM environment, for instance, has revealed that the attractiveness halo effect stays away when reviews are negative and processed more profoundly (Ozanne et al., 2019). Although sponsored reviews are generally positive, the explicit reference to the post being an advertisement might raise caution. This research therefore aims to explore whether the halo effect that is present in ‘regular’ positive reviews also applies to the ‘hybrid’ Instagram reviews that are simultaneously advertisements.

Secondly, we will hone in on other aspects of physical appearance that have not been thoroughly discussed within the setting, but may be of relevance as well. More specifically, studies within the setting of job application procedures have shown glasses, as indicators of perceived intelligence, to positively impact credibility as well (Fetscherin, Tantleff-Dunn, & Klumb, 2020, p.341; van der Land, Willemsen & Unkel, 2015, p.181). These findings, however, have received little attention so far in contexts of online persuasive texts written by

influencers. This study therefore aims to expand the applicability of both types of findings in new communicative settings.

In short, the present study aims to shed light on how perceived attractiveness and the presence or absence of glasses (and their combined effect) affects influencers' perceived credibility, and how this credibility in turn influences purchase intention. This leads us to formulate the following two research questions:

RQ1: To what extent do physical features (attractiveness and glasses) affect the perceived online credibility of influencers' sponsored reviews on Instagram?

RQ 2: How does the perceived credibility of the influencer affect purchase intention?

The remaining part of this dissertation first provides an overview of the existing literature on Instagram and credibility. The section on credibility zooms in on both message and source credibility, and includes a discussion of the related concept 'halo effect' and previous research on the influence of glasses. As the literature review in Remue (2021) also centred on online credibility, several paragraphs, although many to a large extent adjusted or reworked, have been incorporated in this section. The theoretical framework concludes with a discussion of consumer attitude and purchase intention. Section 3 subsequently addresses methodology, discussing the study's design, stimulus material, pre-test, conceptual model and scales used. Section 4 provides the results of the study, which are discussed in depth in section 5. Finally, we formulate the study's limitations and make recommendations for future research.

2. LITERATURE REVIEW

2.1. The platform at use: Instagram

With over a billion active users worldwide, Instagram is one of the most popular SM platforms on the market (Statista, 2021a). Although Instagram was launched as just a photo-sharing app in 2010, it has been refined with many new features throughout the years, such as the possibility to share videos and 'Instagram Stories', which allow users to post photos and videos that automatically disappear after 24 hours. The introduction of Instagram Stories has proved to be very conducive to the app's popularity. It has increased the average time spent on Instagram by seven to ten minutes per day (Sharma, s.d., para. 25), and over 500 million accounts publish such stories on a daily basis (ibid.). Professional profiles, too, readily make use of (and benefit from) the feature; no fewer than a third of the most popular Instagram stories are communicated by business accounts (Mohsin, 2021, para. 7).

Facebook, the absolute market leader with 2.85 billion active users (Statista, 2021b), introduced an almost identical feature, 'Facebook Stories', in 2017. Still, many brands and businesses prefer Instagram to promote their products and services because of its younger audience and the extensive use of hashtags on the platform, which facilitate reaching potential new customers (Mohsin, 2021, para. 2-3). In the next section, we focus on how businesses approach and employ Instagram as a marketing tool.

2.1.1. The use of Instagram for commercial purposes

Over recent years, brands and businesses have been increasingly turning to SM as a marketing medium (Dwivedi, Kapoor & Chen, 2015). Apart from the ability to communicate to an audience, SM also allow brands and businesses to build, sustain, and strengthen relationships with (potential) customers (Jacobson, Gruzd & Hernández-García, 2020, p. 2), which is almost impossible in traditional one-way marketer-to-consumer communication. Moreover, Dwivedi et al. describe SM as "a low cost information exchange platform available for one and all" (p. 291), which enables brands and businesses to reach a more comprehensive audience than by using conventional marketing platforms (ibid.).

From the perspective of the consumer, the preference for SM marketing might be explained by its non-intrusive character. Intrusive advertising occurs when a consumer is involuntarily exposed to an irrelevant advertisement, such as a pop up. This is opposed to non-intrusive advertising, where a consumer receives advertisements which fit his or her needs (Bóveda-Lambie & Hair, 2012). Bóveda-Lambie et al. believe SM foster the latter, as these platforms can be used to "invite advertising into people's lives" (p. 20). Through a like or a follow, consumers can allow companies to become part of their daily lives. This facilitates access to the brand for the consumer, but also promotional activity on the part of the company (Bóveda-Lambie et al., p. 18). This type of non-intrusive advertising – labeled "invertising" – is generally evaluated more positively than regular SM advertising such as banners or button advertisements (Bóveda-Lambie et al., p. 30).

Although there is a wide range of SM platforms suitable for "invertising"[1], Instagram seems especially valuable to marketers: according to Statista, about 70% of Instagram users are aged 34 years or younger, with the largest group between 25-34 years old (2021a). This demographic share of the population – also known as Millennials (1981-1995) and Generation Z (1996-2012) – is especially receptive to SM marketing (Djafarova & Bowes, 2021) and prefers online over physical shopping (Kibo, 2020, para. 2). A 2019 study even revealed that US Millennials did 36% of their shopping via mobile devices (Melton, 2019, para. 4). This seems all the more mentionable as Instagram appears to considerably encourage impulse buying among females (Djafarova et al., p. 7). In this same study, Djafarova et al. reported that Instagram marketing does not inspire impulse buying among male users (ibid.). However, this could also be ascribed to the fashion industry context that was used, as the male participants indicated not being into fashion and mainly using Instagram for other purposes (Djafarova et al., p. 5).

Impulse buying left aside, Instagram definitely impacts consumers' shopping experience: users of the platform tend to base their purchase decisions on the content they come across (Mohsin, 2021, para. 10), and half of them report that Instagram advertisements boost their interest in a brand (Sharma, s.d., para. 5). In addition, 90% of users follow at least one business account, and over 80% of users say they discover new products and services thanks to Instagram (ibid.). Of all SM platforms, Instagram is also most likely to "convert" users from viewers into

[1] Facebook and Twitter, for example, allow users to like or follow fan pages or brand accounts.

shoppers (Adegbola, Gearhart & Skarda-Mitchell, 2018, p. 233), which, at the end of the day, is what businesses want to achieve.

Another factor which probably contributes to the widespread use of Instagram by businesses is the use of hashtags (#) and tags (@). That is, hashtags and tags create direct links to the brand's account. This allows users to discover the brand, but also encourages them to engage with it, creating a potential "multiplier effect" on audience engagement (Adegbola et al., 2018, p. 234). If a brand, for instance, launches a contest in which users are asked to post a photo, or to tag friends in order to enter, this greatly increases exposure. The visibility of the brand is, so to speak, 'multiplied' to the followers of the brand's followers. Hashtags, too, have been shown to extend visibility. Adding relevant hashtags approximately leads to three times more exposure (Sharma, s.d., para. 29). Consequently, businesses have been quick to use this feature to their advantage; in 2018, seven out of ten Instagram hashtags were branded (Mohsin, 2021, para. 5).

In order to consolidate its position as a branding platform, Instagram has created a guide and blog for businesses using the medium (https://business.instagram.com/blog). Instagram updates its blog regularly, providing brands and businesses with success stories and best practices to maximize their user engagement. In 2016, for instance, 'Business Tools' was introduced (Instagram Business Blog, 2016), enabling insights on the behaviour and demographics of a business' audience and making it possible to turn popular posts into advertisements. It did not stop there, however. Throughout the years, Instagram has kept adding new features which facilitate marketing on the platform. In 2017, for example, Instagram shopping was rolled out, allowing users to click on items pictured in photos and to shop directly from a product catalog on the brand's page. Since 2019, users can even do so without ever having to visit the brand's website (Instagram Business Blog, 2019).

These possibilities, together with other features such as growth insights and interactive stories, have convinced millions of businesses to join the platform (Instagram Business Blog, 2017). Instagram's latest report dates back to November 2017, when it celebrated a community of 25 million active businesses. Today, however, that figure must be much higher. Especially given that just six months prior to the report, Instagram counted only 15 million businesses (ibid.). Recent updates from the Instagram Business Blog mostly concern tools for, and advice on branded content from creators, following the rising popularity of social media influencers. In

the next section, we zoom in on these personalities, and consider how their online presence may affect consumer behaviour.

2.1.2. Instagram influencers

Social media influencers are online personalities who use social media to shape other people's opinions (Balaban et al., 2019; Djafarova et al., 2017; Rebelo, 2017; Tafesse et al., 2021). They are often labeled "microcelebrities" (Balaban et al., p. 33; Rebelo, p. 23) and can be found on platforms such as Youtube, Snapchat, Twitter, Instagram, TikTok or personal blogs (Balaban et al.; Lim et al., 2017; Rebelo). Social media users are generally considered influencers once they have accumulated a certain number of followers (Campbell & Farrell, 2020; Djafarova et al., p. 4).

Depending on their follower count, influencers can be divided into five categories: nanoinfluencers, microinfluencers, macroinfluencers, megainfluencers and celebrity influencers (Campbell et al., 2020). Nanoinfluencers are 'newcomers'. They have fewer than 10,000 followers, and many of them are acquaintances. Nanoinfluencers are considered highly authentic and have considerably higher engagement rates than other influencers. Microinfluencers count between 10,000 and 100,000 followers. Their audience is still rather localised, but they are generally able to make a living out of their 'influencer career'. Influencers with a follower count between 100,000 and 1 million are called macroinfluencers. This category of influencers is highly successful and represents the most popular choice among businesses to promote their brands. Megainfluencers, then, have surpassed the 1 million follower mark and have acquired a real celebrity status thanks to their online presence. They only differ from celebrity influencers in that they were not known prior to their influencer career (Campbell et al., pp. 470-472).

Brands and businesses typically prefer using microinfluencers over celebrities in their SM campaigns. Celebrity endorsement, too, can raise brand awareness, but influencers generally generate more engagement (Balaban et al., 2019; Campbell et al., 2020; Lim et al., 2017). There are multiple factors which can explain these higher engagement rates, the first being that influencers are "more capable of communicating to a niche segment" (Lim et al., p.20), which means that their followers typically enjoy a certain type of content. This enables brands to specifically look for an endorser who 'matches' their products. This is not insignificant, as a

product compatible with the influencer's content generally boosts product attitude and purchase intention (Lim et al., p.30).

In addition, influencers are regarded as more credible and trustworthy than traditional celebrities (Djafarova et al., 2017; Lim et al., 2017). They are more relatable, accessible, and their content is considered more authentic (Djafarova et al.; Naderer et al., 2021; Tafesse et al., 2021). They appeal to their audience by interacting and creating an apparent exclusive relationship with them (Balaban et al., 2019; Lim et al.; Rebelo, 2017). Consequently, SM users tend to be more influenced by influencer endorsements or reviews (Djafarova et al.; Tafesse et al.).

Of all SM platforms, Instagram is unmistakably the favourite among influencers (Chadha, 2018). Not only does it generate the highest interaction (Sharma, s.d., para. 5), it also allows 'instafamous' profiles to weave promotional posts into their regular content, reducing audience resistance (Balaban et al., 2019, p. 34; Djafarova et al., 2017, p. 4; Naderer et al., 2021) (cf. 2.1.1.). People are generally aware that brands approach influencers to promote their products (Djafarova et al., p. 6). Sponsored posts must even disclose an explicit reference to the post being an advertisement (Lee et al., 2020; Naderer et al.). Still, this does not seem to harm the credibility of the influencer or brand (Lee et al., p. 244). In fact, this transparency can even enhance the influencer's trustworthiness when the SM user feels similar to them (Naderer et al., p. 703). The fact that these content creators are paid or gifted the products for free (Campbell et al., 2020, p. 472) does not seem to harm the advertisement either, as SM users feel that posting deceitful reviews would damage the influencers' reputation (Djafarova et al., p.6).

One of the main reasons influencer reviews are so effective is because they provide insights on products from the influencers' domains of interest that are usually new and fairly unknown (Tafesse et al., 2021, p.2). Given that influencers generally have a 'niche' of followers with the same interests, this makes their reviews all the more convincing (Lim et al., 2017). Influencers also use their reputation to add value and positive associations to the product, which often results in a higher purchase intention (Balaban et al., 2019; Djafarova et al., 2017; Lim et al.).

The effect of influencer credibility itself, however, remains unclear, as previous studies have generated ambiguous results. While some research suggests that highly credible influencers do

not inspire more message credibility or a higher purchase intention than lowly credible influencers (Lee et al., 2020, p. 244), other research seems to indicate that purchase intention increases when perceived credibility is high (Rebelo, 2017, p. 51). Rebelo also found that purchase intention is affected by the source credibility dimensions 'attractiveness' and 'trustworthiness', but not 'expertise' (p. 52). This influence from attractiveness on purchase intention, however, could not be corroborated in Lim et al. (2017, p. 28).

Results put forward by the three aforementioned quantitative studies are further complicated by findings from qualitative research: while interviews conducted by Balaban et al. confirmed the influence of 'trustworthiness' and 'attractiveness' in influencers, they also underlined the importance of 'expertise' and 'similarity to the influencer' (2019, p. 41). The study did reveal, however, that the relevance of the 'expertise' dimension may depend on the domain of interest (ibid.). Furthermore, research seems to indicate that female users are more likely to be swayed by influencers than male users (Rebelo, 2017, p. 52).

The literature thus gives no conclusive answers on which credibility dimensions are most valuable to influencers, but instead underscores their complex nature. The impact of influencer credibility on purchase intention is even further problematized by seemingly disparate factors influencing the credibility assessment process as well. Apart from perceived attractiveness, items such as glasses, demographic characteristics, as well as the type of product or service being reviewed have been shown to impact perceived credibility and purchase intention (cf. 2.2.). This emphasizes the need to contextualize the effects of influencer credibility while taking into account other variables affecting the credibility assessment process. This study therefore not only aims to provide more clarity on previous findings concerning influencer credibility, but also wants to contextualize by incorporating findings from different settings, and in this way possibly extend their applicability.

The following section therefore serves to elaborate on the variables in this study, providing an overview of how they were approached and conceptualized in previous studies, and how they affected credibility perceptions in their respective contexts.

2.2. Credibility

Credibility refers to the extent to which a person believes information to be truthful, and allows the receiver to rate the source in relation to that information (Eisend, 2006, p. 2). It is seen as a multidimensional construct, and can be divided into three types: medium credibility, message credibility and source credibility.

Medium credibility refers to the means of communication used to provide the information (Armstrong & McAdams, 2009). The credibility that is attributed will differ according to the medium (e.g. a website vs. a newspaper), but also within the medium (e.g. a government website vs. a private website) (Armstrong et al., p. 436). Message credibility can be understood as the extent to which the content of the message is perceived to be true and depends on the perceived reliability of the source and the perceived quality of the message itself (Cheung et al., 2009). Features affecting message credibility include factors such as argument strength, valence, confirmation of prior belief, etc. (ibid.). Finally, source credibility has been described as "a communicator's positive characteristics that affect the receiver's acceptance of a message" (Ohanian, 1990, p. 41). Although various dimensions have been used to define source credibility, four of the most established are expertise, trustworthiness, attractiveness and homophily (Eisend; Ohanian; Ismagilova, Slade, Rana & Dwivedi, 2020).

As the medium in the present study remains constant (Instagram), medium credibility is not elaborated on. More information on Instagram as a branding platform can be found in section 2.1.1. Source and message credibility, however, receive more attention as they are key variables in the study.

2.2.1. Message credibility

As mentioned before, message credibility depends both upon the perceived reliability of the source and the quality of the message itself (Cheung et al., 2009). However, in an online environment, and more particularly in an eWOM context, users do not know each other personally and many exchanges take place anonymously. As a result, source and message credibility are heavily intertwined. A reliable source might have an impact on the perceived credibility of the message, but, conversely, certain textual cues might also alter the perception of the source.

This section discusses a number of text-internal features that can alter such perceptions of source and message. A brief discussion on cognitive processing is provided as well. The present study opts for a design with only one review message so as to focus on the manipulations of the source (cf. 3.2.), but the overview is included both to contextualize the characteristics of the review used, and to demonstrate the complex nature of the credibility assessment process. We feel that some of these aspects need to be addressed as they may prove important in interpreting the results. Moreover, these factors should be taken into account, should future research want to further explore this study's findings.

2.2.1.1. Text-internal features

Language errors and register

Research conducted on language errors and register has shown that the wording of a review does matter. Reviews containing bad grammar and misspellings are generally perceived as less valuable (Schindler & Bickart, 2012, p. 239) and reduce the perceived intelligence and credibility of respectively source and message for male reviewers (Lybaert et al., 2020, p. 14). Lybaert et al. also included the variable of register in their study. They specifically looked at the impact of *tussentaal*, a colloquial variety of Dutch that is seen as more informal than Standard Dutch, and found that the informal variety generated nearly the same effect as language errors (Lybaert et al., p. 25). However, a cumulative effect of language errors and register could not be found (ibid.).

Although only to a certain degree, reviews that contain more informal elements (i.e. expressive slang and humour) have also been perceived as more valuable (Schindler et al., 2012, p. 240). The influence of an informal register could thus differ, depending on the context, the extent to which it is used or the actual features integrated. While some informality might generate a feeling of connection between source and receiver, too much of it could make the source look less competent or less intelligent (Lybaert et al., 2020; Schindler et al.).

Message valence and sidedness

Message valence is a factor that highlights either the strengths or the weaknesses of a product or service (Lim & Van Der Heide, 2014). It is also referred to as recommendation framing, as the message is framed either positively or negatively (Cheung et al., 2009). Positively framed reviews will lead to positive evaluations, whereas the opposite is true for negative reviews (Lim

et al., p. 71). Results on the impact of these negative reviews, however, have been inconsistent: whereas some research states that consumers perceive negative reviews to be more credible than positive reviews (Lee & Koo, 2012, p. 1982), other research suggests that they do have a stronger impact, but are nonetheless less credible than positive reviews (Lim et al., p.79). That negative reviews should carry more weight than positive reviews is referred to as the 'negativity bias' (Lee et al., p.1975), and might result from negative reviews being scarcer, or consumers being extremely aversive to risks (ibid.).

Message sidedness, on the other hand, refers to whether a message contains one-sided or two-sided information. One-sided information will highlight only positive or negative aspects of the subject, whereas two-sided information will point out both, and therefore appear more complete in the eyes of the receiver (Cheung et al., 2009, p. 16). Research has shown that two-sided reviews are generally perceived as more credible (Cheung, Sia, & Kuan, 2012; Flanagin, Winter & Metzger, 2018), especially when the receiver is high in flexible thinking (Flanagin et al., p. 14), well informed about the product (i.e. high expertise) (Cheung et al., p.628) or not motivated to consider the review in detail (i.e. low involvement) (ibid.). Furthermore, two-sided reviews have been shown be more helpful than positive one-sided reviews, with a stronger effect for product reviews than service reviews (Li, Lee & Wu, 2020).

Argument quality

Argument quality can be defined as "the extent to which the message receiver views the argument as convincing or valid in supporting its position" (Cheung et al., 2009, p. 15). Multiple studies have already demonstrated that argument quality significantly influences perceived credibility (Cheung et al., 2009; Cheung et al., 2012; Shan, 2016). Findings by Cheung et al. even showed that argument quality might be the most influential factor when consumers evaluate online reviews (2012, p. 627). Research by Shan seems to corroborate this, as she found argument quality, engendered by relevance, objectivity and verifiability, to be the only factor that is positively related to both expertise and trustworthiness (p. 639), two well-established source credibility dimensions (Ohanian, 1990).

An overview of the aforementioned features integrated in the review used for the experiment can be found in section 3.2.

2.2.1.2. Type of cognitive processing

The extent to which the aforementioned factors influence the perceived credibility of the source or message depends on the cognitive process at work. The Elaboration Likelihood Model (ELM) states that persuasion can be attained via a central and a peripheral route (Petty & Cacioppo, 1986). While the central route requires careful and rational processing of the message, the peripheral route relies on heuristic cues, such as the attractiveness of the reviewer (Ozanne et al., 2019) (cf. 2.2.3.).

How likely a person is to consider a message thoroughly thus depends on individual and situational factors (Petty et al., 1986). One of these situational factors is message valence: negative valence in reviews leads consumers to pay closer attention to other cues as well (Xu, 2014, p. 143), as the negative emotions in the review require deeper processing (Ozanne et al., 2019). When information has to be processed more thoroughly (negative valence), consumers tend to look for additional cues which might help in their impression formation, whereas they rely on surface characteristics when only shallow processing is required (positive valence) (Ozanne et al.; Xu). However, when consumers are confronted with multiple reviews of a mixed valence, the high cognitive load will trigger them to rely on stereotype-consistent information (heuristic cues) instead of deeper processing (Ozanne et al., p. 733).

2.2.2. Source credibility

According to Ohanian (1990), source credibility can be measured as a three-dimensional construct consisting of trustworthiness, expertise and attractiveness. Some studies also make a distinction between different categories of attractiveness: social attractiveness, which is based on personality and how pleasant someone seems to be around (Edwards, Stoll, Faculak & Karman, 2015), task attractiveness**,** which refers to abilities and our desire to work with someone (ibid.) and physical attractiveness**,** which is merely based on looks (Wilton, 2015). More recent studies have also incorporated ‘homophily’ as an important factor influencing source credibility (Ismagilova et al., 2020).

When reading an online review, however, consumers may have a hard time forming an impression of the source as they lack the non-verbal cues which are normally present in face-to-face interaction. They therefore look for other cues, both within and outside of the review,

to decide whether they deem the source credible or not (Walther, 1993). The presence of some of these cues may depend on the platform at issue, such as gender, age, ethnicity, number of friends or followers, residence, reputation, social distance, etc. (Armstrong et al., 2009; Craciun & Moore, 2009; Lim et al., 2014; Lin & Xu, 2017; Shan, 2016; Xu, 2014; Yoo, Lee & Gretzel, 2007), but many can also be found in profile pictures.

Profile pictures not only allow people to form an impression of unknown individuals (Xu, 2014), they also help those displayed to manage and improve their online self-presentation (Krämer et al., 2008). That is, certain elements in profile pictures can serve as visual cues that increase perceived credibility. When the source is aware of these cues, they can engage in impression formation by manipulating their photo and presenting those characteristics which they believe to be beneficial to their perceived credibility (Krämer et al.). This section therefore presents the three source credibility dimensions as defined by Ohanian (1990) together with a number of visual cues they are affected by. The final subsection briefly discusses homophily.

2.2.2.1. Trustworthiness

Trustworthiness can be defined as the ability to be relied on as honest or truthful (Lexico, s.d.), and can be evoked through a number of physical features. Eye contact, for example, has been shown to be an important indicator of trustworthiness; recruiters tend to judge job applicants avoiding eye contact in their profile pictures as less trustworthy and distant (Wilton, 2015, p. 38-40). Consumers, too, seem to attach great importance to eye contact in peer-to-peer (P2P) e-commerce. A study conducted on the P2P platform Airbnb, "where pictures of the renter are included to diminish consumers' feelings of uncertainty" (Broeder & Remers, 2018, p. 336), revealed that pictures with eye contact generate more trust than those without, which in turn also leads to a higher booking intention (Broeder et al., p. 338).

However, there is no direct effect of eye contact on booking intention (Broeder et al., 2018, p. 338). An explanation put forward by Broeder et al. themselves is that people might base their opinion on other facial features as well, such as mouth shape (ibid.). Results from other research seem to corroborate this, as trustworthiness increases when the lack of eye contact is combined with a smile (Wilton, 2015, p. 39). In fact, smiling in general is seen as an influential indicator of trustworthiness (Wilton, p.44).

2.2.2.2. Expertise

Expertise is described as skill or knowledge in a particular field (Lexico, s.d.). Like trustworthiness, perceived expertise can be influenced by external features, such as hair colour. Redheads, for instance, tend to be considered cold and competent, while blondes are typically judged incompetent but more likeable (Takeda, Helms & Romanova, 2006, p. 89). This 'hair colour bias' is, for instance, visible in CEO positions in the UK: redhaired people are overrepresented in these leadership positions, whereas blondes are underrepresented in comparison to their share of the entire UK population (Takeda et al., p. 93).

The influence of attire cannot be underestimated either. Job applicants who wear formal attire in their LinkedIn profile pictures tend to be judged as serious, professional and experienced (Wilton, 2015, pp. 32-33). Photos taken in a holiday setting, with the candidates wearing more informal clothing, on the other hand, are evaluated more negatively and make those displayed appear less professional (Wilton, pp. 36-38). In addition, clothing style has been linked to education level: the more formal the clothing, the higher the perceived level of education (Johnston, Jackson & Kelley, 2009, p. 17). Still, caution is at hand regarding the generalizability of these findings. Formal attire, for instance, may be beneficial on LinkedIn, but this need not be the case for Instagram. Cues which signal expertise likely differ from context to context.

Finally, job applicants who smile in their LinkedIn profile pictures are generally perceived as having more expertise (Wilton, 2015, p. 44). Male candidates wearing a beard, too, are regarded as having more expertise than their clean-shaven counterparts (van der Land & Muntinga, 2014, p. 261). For female applicants, it is glasses that increase perceived expertise (van der Land, Willemsen & Unkel, 2015, p. 181). While bearded men may appear older and therefore be thought of as more experienced, bespectacled females may be rated higher in expertise because they look more intelligent (Wei & Stillwell, 2017). This could be the case because, although this perception is based on an inaccurate stereotype, glasses tend to be seen as an indicator of intelligence for both men and women (Wei et al., p. 37). As glasses play an important role in the present study, and as their impact on people's perceptions reaches further than only expertise, a separate subsection is devoted to this topic.

Glasses

As mentioned above, glasses are typically seen as an indicator of intelligence. Wei et al., for example, revealed that people use the visual cue of glasses to judge intelligence from Facebook profile images (2017, p. 37). Glasses have also been shown to increase perceptions of competence (Fetscherin et al., 2020, p. 341) and trustworthiness (Leder, Forster, & Gerger, 2011, p. 220). In their study, Leder et al. distinguished between full-rim and rimless glasses. While both increased perceived intelligence, only full-rim glasses harmed physical attractiveness (ibid.). Later research by van der Land et al. did not entirely confirm these results; women with (full-rim) glasses were perceived as having more expertise, but were also rated higher in attractiveness (2015, p. 181). Finally, glasses have been shown to positively influence warmth (Fetscherin et al., p. 341).

Although the literature reveals some contradictory results, all studies do seem to confirm that there is a positive link between glasses and perceived intelligence and expertise. If people wearing glasses are attributed higher levels of intelligence and expertise, and therefore also credibility (cf. 2.2.2.2.), and message credibility is partly dependent upon the message source (cf. 2.2.1.), we can formulate the following hypothesis answering RQ1:

H1: Reviews by influencers wearing glasses are deemed more credible than reviews by influencers without glasses.

Given the conflicting results from previous studies, it seems sensible to also look whether our data can verify or refute previous findings regarding the link between glasses and attractiveness, even if this is not related to answering the study's central questions.

2.2.2.3. Attractiveness

The attractiveness dimension has long been left out in studies on online credibility. Instead, this type of credibility research focused mainly on textual information. Cheung et al., for instance, justified the absence of the attractiveness dimension by stating that "the nature of the virtual discussion may not permit the conveyance of such cues" (2009, p. 16), suggesting that it is not possible to include attractiveness cues in a study conducted on an online forum. Today, most review platforms, as well as other social networking sites, allow for profile pictures, and

the attractiveness dimension has been increasingly incorporated in studies on online credibility (Ozanne et al., 2019; Peng et al., 2020).

Ozanne et al. explored how attractiveness cues in reviewers' profile pictures affected brand evaluations of the products reviewed and additionally checked for the influence of positioning and stereotypes. The results indicated that brand evaluations are higher when positive reviews are written by attractive (as opposed to less attractive) people (Ozanne et al., 2019, p. 733), indicating an attractiveness halo effect (see below). However, the attractiveness cue has no impact when reviews are negative (ibid.) (cf. 2.2.1.2.). The study also showed brand evaluations to be higher when the first of multiple reviews is stereotype consistent (a positive review written by an attractive person) (Ozanne et al., p. 733). This was anticipated not only because the first review stays in the short-term memory the longest, but also because people tend to rely on stereotypes when exposed to a high cognitive load (i.e. multiple reviews with mixed valence) (Ozanne et al., p. 730).

Peng et al. (2020) fine-grained this research by differentiating between attractive, less attractive and unattractive people. They found a U-shaped relationship between attractiveness and perceived credibility that was moderated by product relevance (Peng et al., p. 81). More in particular, attractive and unattractive people are both rated higher in credibility than their plain-looking counterparts, suggesting a "plainness penalty" (Peng et al., p. 81). Whereas attractive people are considered more credible mainly because of their apparent sociability, unattractive people enjoy greater perceived credibility because they appear to be more competent (Peng et al., p. 81). In comparison to plain-looking people, attractive individuals are also considered more competent, but to a lesser extent than unattractive individuals (Peng et al., p. 79).

Attractiveness, too, can be increased through (non-permanent) visual cues. Research in a recruitment context, for example, has revealed that smiling is beneficial to attractiveness (Wilton, 2015, p. 44). Female candidates with glasses have also been rated higher in attractiveness (van der Land et al., 2015, p. 181), although this finding was unexpected and seems to contradict existing literature (van der Land et al., 2015, p. 177).

Halo effect

The aforementioned halo effect is a type of cognitive bias which leads people to unconsciously rely on their first impression when attributing specific characteristics to an (unknown) individual (Nisbett & Wilson, 1977). A global characteristic, such as attractiveness, is then used to attribute unrelated personality traits, such as intelligence, responsibility, confidence, sociability, etc. (Batres, 2020, p. 3). The halo metaphor stems from the fact that one general characteristic "outshines" other traits (Long-Crowell, 2013, para. 6), resulting in a biased perception. However, the characteristics attributed can also be negative. This 'negative halo', which unfavourably influences one's judgement, is often labeled as the 'horn effect' (Oxford Reference, 2009).

Although the attractiveness halo effect has been demonstrated in multiple studies, research has also shown that attractiveness is no linear concept (Peng et al., 2020) (cf. 2.2.2.3.), which means that one cannot simply deduce a horn effect for unattractiveness. In addition, the impact of attractiveness is subject to moderating factors such as gender and product relevance (Peng et al., p. 81) or job type in the context of recruitment (van der Land et al., 2014, p. 259). In the eWOM environment, attractiveness seems to trigger a positive halo effect when reviews are positive, but not when they are negative (Ozanne et al., 2019, p. 733).

Sponsored reviews posted by Instagram influencers are usually of positive nature. If attractive reviewers generate a positive halo effect in such contexts, we can formulate a second hypothesis answering RQ1:

H2: Reviews by attractive influencers are deemed more credible than reviews by less attractive influencers.

Although the sponsored nature of the review does not seem to harm the credibility of the message or influencer (Lee et al., 2020, p. 244), it is still possible that a '#sponsored' reference raises caution and hereby triggers deeper processing, reducing the impact of source attractiveness. However, research has yet to examine this specific issue in the fairly new context of Instagram reviews. Results from this study may already give an indication.

2.2.2.4. Homophily

Homophily can be described as the "degree to which two or more individuals who interact are similar in certain attributes" (Ismagilova et al., 2020, p. 3). In a review setting, this corresponds to the perceived similarity between consumer and reviewer. Increased similarity generally increases the persuasiveness of the message, and positively influences source credibility and consumer behaviour (Ismagilova et al., p. 6). Generally, homophily is approached as a 4-dimensional construct, consisting of attitude, background, value-morality, and appearance (Ismagilova et al.; Ladhari, Massa & Skandrani, 2020).

In an online environment, inferences about homophily are mainly made through content and information provided on the source's profile (Ismagilova et al., 2020, p. 4). In our experiment, however, respondents were able to see only one review and any additional information was kept to a bare minimum (removal of the name, number of likes, etc.) (cf. 3.2.). Respondents thus had no access to the influencer's social class, residence, cultural or moral values, or any other such indicators of similarity (Ladhari et al., 2020). Examining homophily as a main focus would therefore have required an entirely different approach, but, given the relevance for the influencer context, we still decided to gauge its influence through gender, age and glasses, as respondents would have been able to make inferences about these three factors based on the appearance of the source.

2.3. Consumer attitude and purchase intention

Whereas consumer attitude represents how favourable one feels towards a particular brand, product or advertisement, purchase intention refers to the willingness to buy a certain brand or product. As the rise of social media has changed advertising, so has it changed the relationship between attitude and purchase intention. In traditional advertisements, brand attitude was key in forming purchase intention, but in mobile and SM settings, the advertisement itself seems to play a central role (Friman, 2010, p. 70). On Instagram, these advertisements are increasingly taking the form of sponsored reviews.

In terms of efficacy, it seems requisite that these reviews contain no spelling errors, as they can harm review attitude, product attitude and purchase intention (Maesschalck, 2015, p. 40). The source seems to play an important role as well in this new type of advertising. That is, the

higher the credibility of the influencer posting the review, the higher the levels of purchase intention (Rebelo, 2017, p. 51). The study by Rebelo also suggests that only the source credibility dimensions ‘attractiveness’ and ‘trustworthiness’ influence purchase intention (p. 52). Research by Lim et al., however, produced contradictory results: source attractiveness did not directly affect purchase intention, but only improved brand and product attitude (2017, p. 30). Lim et al. also put forward that source credibility of SM influencers altogether influences neither attitude, nor purchase intention because of their lack of expertise with respect to the endorsed product (2017, p. 29). The study also underlined the importance of product match-up for advertisement success, and revealed a mediating role between consumer attitude and purchase intention (2017, p. 30).

The literature thus is not in complete accord on which factors influence consumer attitude and purchase intention in the context of influencer advertisements. Therefore, it seems useful to re-examine them in the current study. As there is, because of the conflicting results, no certainty that there is a direct effect of source credibility on purchase intention, we propose the following hypotheses for RQ2:

H3: Higher perceived source credibility improves consumer attitude.
H3a: Perceived intelligence enhances the perceived credibility of influencers and therefore positively influences consumer attitude.
H3b: Attractiveness enhances the perceived credibility of influencers and therefore positively influences consumer attitude.
H4: Consumer attitude mediates the relationship between perceived source credibility and purchase intention.
H4a: The effect of perceived intelligence on purchase intention is mediated by consumer attitude.
H4b: The effect of attractiveness on purchase intention is mediated by consumer attitude.

3. METHODOLOGY

3.1. Design

In order to examine the influence of attractiveness and glasses on online credibility, an online experiment was conducted using a two (attractive vs. less attractive) x two (glasses vs. no glasses) between-subjects design. A Dutch survey was distributed online and respondents were randomly assigned to one of the four conditions.

3.2. Stimulus material

Eight profile pictures (four male and four female) and a positive review[2] were created for the experiment. The reviews were posted on fake Instagram accounts to make them as realistic as possible. The reason both male and female pictures were used in the pre-test was to check whether the manipulations produced the same effects for both genders, and, if not, to see in which gender the manipulations had been most successful.

The eight profile pictures were all manipulations of a copyright-free picture found in a database on unsplash.com. FaceApp was first used to turn this photo of a brown-haired male model, smiling towards the camera, into a photo of a female model. This specific picture was chosen as blonde or red hair might induce bias, and the absence of eye contact or a smile could harm trustworthiness and expertise (cf. 2.2.2.). The choice, then, to manipulate this one picture (instead of using a different male and female model, for example) was made to minimize the possible interference of factors such as facial expressions, facial features, hair colour, etc., by keeping them as stable as possible. Also a hat, which was present in the original picture, was removed on Pixlr.com to avoid any possible interference or bias. In a second step, fashionable glasses were added on the site Pixlr.com. This resulted in four 'profile' pictures, which were then each manipulated to look 'less attractive'.

Attractiveness was manipulated by means of weight and blemishes, as smooth skin is generally judged more attractive (Fink, Grammer, & Thornhill, 2001, p. 96) and BMI and perceived physical attractiveness are typically negatively correlated in Western culture (Chithambo &

[2]Originally, two reviews were composed (one with and one without language errors). However, a separate pre-test revealed that only 55.4% of respondents noticed one or more language errors, so we decided not to incorporate them as a variable in our study. Findings from pre-test 2 are included in annex 2.

Huey, 2013, p. 3). Both blemishes and weight were added using FaceApp. It should, however, be mentioned that most of these studies on attractiveness concern only women. Whether the results also apply to men can therefore not be certain.

The glasses which were added in the manipulations had a small gold rim. This type of glasses is currently fashionable, and Influencers are generally expected to follow the latest trends (Djafarova et al., 2021, p. 5). Moreover, Leder et al. found that thicker, full-rim glasses lead to a decrease in attractiveness (2011, p. 220), although results by van der Land et al. did not corroborate these findings (2015, p. 181). Results from the present study might bring further clarification.

The scenarios are referred to with the following abbreviations:

- MANG: Male, Attractive, No Glasses
- MAG: Male, Attractive, Glasses
- MLANG: Male, Less Attractive, No Glasses
- MLAG: Male, Less Attractive, Glasses
- FANG: Female, Attractive, No Glasses
- FAG: Female, Attractive, Glasses
- FLANG: Female, Less Attractive, No Glasses
- FLAG: Female, Less Attractive, Glasses

Figure 1: original picture

Figure 2: thin-rimmed glasses (gold)

Figure 3: MANG

Figure 4: MAG

Figure 5: MLANG

Figure 6: MLAG

Figure 7: FANG

Figure 8: FAG

Figure 9: FLANG

Figure 10: FLAG

Given the fictitious nature of the influencers, and given the removal of additional source cues (see below), respondents are likely to see our stimuli as 'nanoinfluencers' (cf. 2.1.2.), who have the smallest following and are relatively unknown.

The second part of the stimulus material consisted of a fictional review created for the experiment. It was made to look like a sponsored advertisement for wireless earphones from the fictitious brand 'lessmessaudio'. Earphones were chosen as we wanted the product to be gender neutral: not only would this enable a better product match-up with both male and female influencers, a less neutral product (e.g. skin care) would likely also generate different results among male and female respondents. The number of likes and the name of the influencer were removed to avoid possible interference or bias[3]; similar to the number of followers, the number of likes on a post is indicative of popularity and can serve as a credibility cue (cf. 2.2.2.). The correlation between the number of likes and perceived credibility, however, likely varies from person to person. Removing the likes thus precludes possible interference, but might at the same time harm source and message credibility.

[3] Even names can be perceived as (un)attractive. Gebauer, Leary and Neberich, for example, showed how culturally devalued names can lead to prejudice and harm online popularity (2012).

This rationale also applies to the picture accompanying the review; not including a picture might rouse suspicion, but at the same time a visual representation of the product is prone to creating interference. That is, previous experiences with a similar looking product will likely influence the consumer's attitude towards the advertised product. The picture of the product itself was therefore cropped out of the review, and not visible in the survey.

Furthermore, the review itself talked about positive aspects of the product, underlined a link between the reviewer and the product (product match-up) and included an explicit reference to the review being sponsored, which is mandatory (Lee et al., 2020; Naderer et al., 2021). The review contained no spelling or grammatical mistakes.

Finally, the sponsored nature of the review did entail certain limitations; as sponsored content is presumed to be favourable, the message was one-sided and framed positively, even though two-sided reviews are generally considered more credible (Cheung et al., 2018). In terms of argument quality, the information can be seen as relevant and partially verifiable, but objectivity is most certainly not applicable. Lastly, as influencers are supposed to be relatable and accessible (Djafarova et al., 2017), a certain level of informality can be expected, for it can generate a feeling of connection between source and receiver (Lybaert et al., 2012).

Figure 11: review

3.3. Pretest

3.3.1. Sample

154 participants completed the first survey. Two of them were excluded because they did not notice the glasses, and one because he or she was under eighteen years old. This resulted in 151 respondents, who were randomly assigned to eight conditions. 124 respondents were female, and the majority (70.2%) was aged 18-26 years old. Most of the respondents had pursued higher education (92.8%) and were Western (94.7%).

Table 1: demographics pre-test

DEMOGRAPHIC	FREQUENCY	PERCENTAGE
GENDER		
Male	27	17.9
Female	124	82.1
AGE		
Gen Z (1996-2004) (18-26)	106	70.2
Gen Y (1981-1995) (27-41)	19	12.6
Gen X (1965-1980) (42-57)	19	12.6
Gen B (1946-1964) (58-76)	7	4.6
EDUCATION LEVEL		
Primary education	0	0
Secondary education	11	7.3
Bachelor	70	46.4
Master or PhD	70	46.4
ETHNICITY		
Western	143	94.7
Western-Asian	5	3.3
Middle Eastern	2	1.3
Asian	1	.7

3.3.2. Procedure and results

The pre-test consisted of a Dutch survey created with LimeSurvey and was distributed online. Respondents were not renumerated. The first part of the questionnaire was composed of demographic questions, followed by a text instructing participants to look at the picture of the influencer and to complete the questions, which measured physical attractiveness, social attractiveness, perceived intelligence and dynamism with a 7-point Likert scale. The order of the questions was randomized to avoid patterned answers. Respondents were subsequently asked about factors which could possibly influence the results. This included questions on

whether the influencer had worn glasses, whether the participants wore glasses themselves and which device they had used to complete the survey. Enquiring about the device had to ascertain that filling out the questionnaire on a laptop or phone would not lead to different results (because of the visuals). The question on whether the respondents wore glasses also seemed relevant, as they could be more or less critical towards glasses than people who do not wear them, or they could be more susceptible to noticing them. If participants indicated that they had seen glasses, they were also asked whether they thought they were fashionable (on a 7-point Likert scale). This question was included to verify the findings that fashionable glasses do not harm physical attractiveness (cf.2.2.2.2.). Respondents were also able to leave remarks at the end of the survey.

The constructs used to measure physical attractiveness (PA) and social attractiveness (SA) proved to have sufficient levels of internal consistency (PA, α=.747; SA, α=.832). Although sufficiently reliable, the decision was made to remove the fourth item (*I believe the influencer is ordinary looking*) of the PA scale, as it gave rise to some confusion. Removing this item also led to a higher internal consistency for the construct (PA, α=.89). At this point of the experiment, the social attractiveness scale was incorporated for exploratory reasons. The main aim of our pre-test was to check whether the photo manipulations had an effect on perceived attractiveness (and intelligence). We hypothesized that they would influence physical attractiveness (cf. 3.2.), but as research on which physical features affect social attractiveness is scarce, it seemed advantageous to include both attractiveness scales. Perceived intelligence was measured using only one item. In the actual survey, three items of expertise (Ohanian, 1990) were added to compose a construct.

When the pre-test was conducted, there were still multiple venues for the present research. Dynamism (DY), then, was included to keep as many options open as possible. However, the results revealed a low internal consistency (DY, α=.63), and the scale was removed. The low levels of reliability were likely due to the limited number of items, which were only partly taken from the construct used in Seghers, De Clerck, & Lybaert (2021). That is, only the items which applied to the influencer review context were used.

Table 2: scales pre-test

Construct	indicator	Cronbach's Alpha	Items removed	New Cronbach's Alpha
Physical attractiveness	PA1 PA2 PA3 PA4 PA5	.747	PA4 (*ordinary*)	.89
Social attractiveness	SA1 SA2 SA3 SA4 SA5 SA6	.832	-	.832
Perceived intelligence	PI1	/	-	/
Dynamism	DY1 DY2 DY3 DY4	.63	Scale was removed	-

The data from the pre-test was analyzed using an Independent-Samples Mann-Whitney U Test. The findings revealed that the manipulations were successful: glasses significantly enhanced perceived intelligence ($p<.001$) and the influencers without blemishes and added weight were judged significantly more attractive ($p<.001$). The female influencers, however, were judged more physically attractive than the male influencers, regardless of the attractiveness condition ($p<.001$). In fact, even the 'attractive' male influencers were rated fairly low in attractiveness ($M=3.2$ for the 'less attractive' and $M=4.2$ for the 'attractive' male influencers). The score given to the attractive male influencers only differed slightly from the score which the 'less attractive' female influencers received ($M=4.1$). The 'attractive' female stimuli received the highest average score ($M=5.2$).

Interestingly, the manipulations also significantly affected social attractiveness: less attractive influencers were perceived as more socially attractive ($p<.001$). Influencers wearing glasses, too, were attributed higher levels of social attractiveness ($p<.044$). Moreover, social, as opposed to physical attractiveness, was stable across gender. Importantly, glasses did not affect physical attractiveness and the 'physically attractive' influencers were not rated higher in intelligence. Lastly, the glasses were deemed sufficiently fashionable ($M=5.01$) and the type of device used to fill out the questionnaire did not affect the results.

The results from our pre-test verify the findings from Leder et al. that less pronounced, more fashionable glasses (as opposed to the stereotypical thick-rimmed glasses) increase perceptions of intelligence, but do not harm physical attractiveness (2011, p. 220). The finding that glasses enhance social attractiveness is also in line with previous research, stating that glasses positively influence warmth (Fetscherin et al., 2020, p. 341) and trustworthiness (Leder et al., 2011, p. 220).

The pre-test also provides new insights into the concept of 'attractiveness'. That is, most research so far has approached attractiveness as a concept which only qualifies one's physical appearance (Batres, 2020; Ohanian, 1990; Ozanne et al., 2019; Peng et al., 2020; van der Land et al., 2015). Our results, however, suggest that people who are less blessed with good looks may benefit from a different type of attractiveness. Edwards et al. (2015) already employed task and social attractiveness scales to assess the influence of using a LinkedIn profile picture, but to our knowledge, no research has yet been conducted on a social attractiveness halo effect (as opposed to a physical attractiveness halo effect), or on which physical features might affect social attractiveness.

The results from the pre-test, which indicate a discrepancy between physical and social attractiveness, lead us to modify H2, H3b and H4b:

H2a: Reviews by physically attractive influencers are deemed more credible than reviews by less physically attractive influencers.
H2b: Reviews by socially attractive influencers are deemed more credible than reviews by less socially attractive influencers.

H3b: Physical attractiveness enhances the perceived credibility of influencers and therefore positively influences consumer attitude.
H3c: Social attractiveness enhances the perceived credibility of influencers and therefore positively influences consumer attitude.

H4b: The effect of physical attractiveness on purchase intention is mediated by consumer attitude.
H4c: The effect of social attractiveness on purchase intention is mediated by consumer attitude.

As ‘attractiveness’ yielded different results for our male and female stimuli, and given the limited time span of the study, we examined whether the male condition in our study could be dropped. There were two main factors motivating the elimination of the male (instead of the female) condition: the ‘attractive’ male influencers, as opposed to the ‘attractive’ female influencers, were rated fairly low in attractiveness (M=4.2 vs. M=5.2). Moreover, the average score attributed to the ‘less attractive’ male influencers signalled a possible risk for an “ugliness premium” (Peng et al., 2020, p. 81) (cf. 2.2.2.3.).

In order to verify whether the results from the pre-test were still valid for the female condition only, we first assessed the normality of the data. Both the skewness and kurtosis values and the Shapiro-Wilk test confirmed that the PA and SA data for the female stimuli were approximatively normally distributed, so an Independent Samples T-test could be performed. This two-tailed T-test revealed that also within the female condition, physical and social attractiveness differed significantly between ‘attractive’ and ‘less attractive’ influencers (respectively p<.001 and p=.003).

The pretest thus resulted in a 2 (attractive vs. less attractive) x2 (glasses vs. no glasses) experimental design with only female stimuli:

- FANG: Female, Attractive, No Glasses
- FAG: Female, Attractive, Glasses
- FLANG: Female, Less Attractive, No Glasses
- FLAG: Female, Less Attractive, Glasses

Figure 7: FANG

Figure 8: FAG

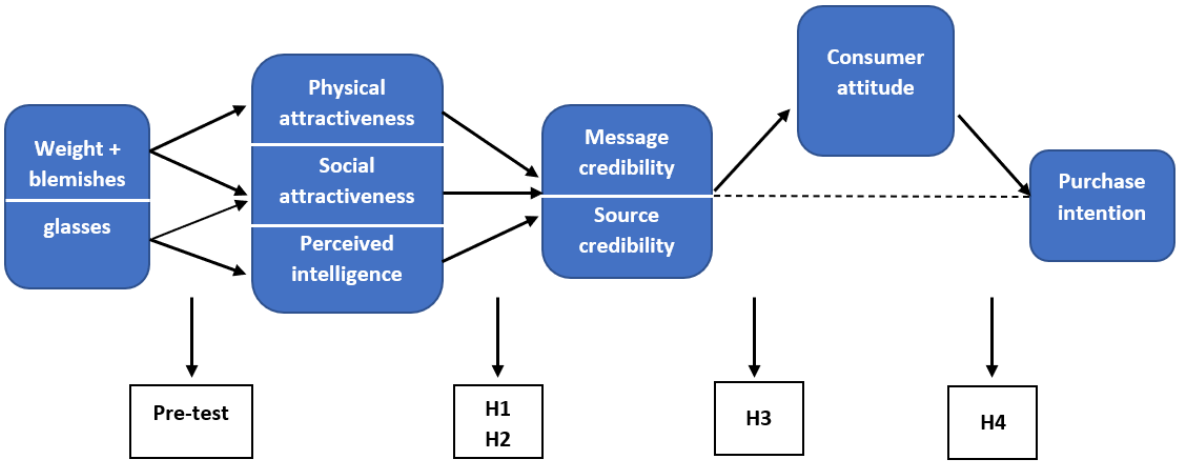
Figure 9: FLANG

Figure 10: FLAG

3.5.Conceptual framework

Based on the literature review and the pre-test, we propose the following conceptual model:

Figure 12: Conceptual model



The arrows illustrate how we believe the variables relate to one another, and are accompanied by the corresponding hypotheses. The first arrows represent the results from our pre-test, which revealed that weight and blemishes affect perceived physical and social attractiveness, and that glasses affect perceived intelligence and social attractiveness.

H1 and H2 state that perceived intelligence, as well as physical and social attractiveness will positively influence message credibility. Although only one review is used in the study, we theorize that, because of the intertwined nature of source and message (cf. 2.2.1.), message credibility will vary according to the source. The message also did not contain any features (such as language errors) which might directly impact its credibility.

Building on H1 and H2, H3 proposes that sources benefitting from higher perceived intelligence, social attractiveness or physical attractiveness will positively influence consumer attitude. Finally, we hypothesize (H4) that consumer attitude acts as a mediator between source credibility and purchase intention.

The final part of this section includes an overview of the hypotheses for each research question:

RQ1: To what extent do physical features (attractiveness and glasses) affect the perceived online credibility of influencers' sponsored reviews on Instagram?

H1: Reviews by influencers wearing glasses are deemed more credible than reviews by influencers without glasses.
H2: Reviews by attractive influencers are deemed more credible than reviews by less attractive influencers.
H2a: Reviews by physically attractive influencers are deemed more credible than reviews by less physically attractive influencers.
H2b: Reviews by socially attractive influencers are deemed more credible than reviews by less socially attractive influencers.

RQ 2: How does the perceived credibility of the influencer affect purchase intention?

H3: Higher perceived source credibility improves consumer attitude.
H3a: Perceived intelligence enhances the perceived credibility of influencers and therefore positively influences consumer attitude.
H3b: Physical attractiveness enhances the perceived credibility of influencers and therefore positively influences consumer attitude.
H3c: Social attractiveness enhances the perceived credibility of influencers and therefore positively influences consumer attitude.
H4: Consumer attitude mediates the relationship between perceived source credibility and purchase intention.
H4a: The effect of perceived intelligence on purchase intention is mediated by consumer attitude.
H4b: The effect of physical attractiveness on purchase intention is mediated by consumer attitude.
H4c: The effect of social attractiveness on purchase intention is mediated by consumer attitude.

3.6.Sample

The final survey yielded a sample size of 336 respondents. However, multiple participants had to be excluded for several reasons: 4 respondents claimed that they were familiar with the brand; 17 found the product advertised not gender neutral; 28 either did not notice the glasses, or indicated having seen glasses when there had been none. This resulted in 287 respondents for the actual experiment, which had 4 conditions participants were randomly assigned to. This average of over 70 participants per scenario was still amply sufficient for the central limit theorem (CLT) to hold, which requires a minimal sample size of 30 participants per condition (Ganti, 2022, para. 4).

Of the 287 respondents, 106 were male, 179 were female and 2 indicated identifying as neither. The majority (40.4%) of participants were 42-57 years old, followed closely by people aged between 18-26 (34.1%). Respondents aged 27-41 (15.3%) and 58-76 (10%) were least represented. Of the 229 who had pursued higher education (79.8%), 126 (43.9%) hold a Bachelor's and 103 (35.9%) a Master's degree (or PhD). 3 participants (1%) indicated not having obtained a diploma of secondary education. Finally, our sample consists almost exclusively of Western participants (97.7%)

Although no central question in this study, these variables will also be discussed in the analysis, as it could be interesting to examine whether the distinct groups display different tendencies, owing to homophily or generation gaps, for example.

Table 3: demographics

DEMOGRAPHIC	FREQUENCY	PERCENTAGE
GENDER		
Male	106	36.9
Female	179	62.4
X	2	.7
AGE		
Gen Z (1996-2004) (18-26)	98	34.1
Gen Y (1981-1995) (27-41)	44	15.3
Gen X (1965-1980) (42-57)	116	40.4
Gen B (1946-1964) (58-76)	29	10.1
EDUCATION LEVEL		
Primary education	3	1
Secondary education	55	19.2
Bachelor	126	43.9
Master or PhD	103	35.9
ETHNICITY		
Western	281	97.9
Eastern	3	1
Asian	3	1

3.7. Procedure

The final survey was created with LimeSurvey and distributed online. Respondents were not renumerated and, again, the first part of the questionnaire was composed of demographic questions. Participants were then shown a text instructing them to read an Instagram review. The text also explained that the reviewer's name had been left out for 'privacy concerns', but that their profile picture was allowed to be used. Participants then saw the review accompanied by one of the four female pictures from the pre-test. Respondents were randomly assigned to one of these four conditions. The pictures of the review and influencer were followed by a number of questions measuring message credibility, brand attitude, product attitude and purchase intention through a 7-point Likert scale. The order of the questions was randomized to avoid patterned answers. Several control questions were integrated, asking whether the influencer in the picture wore glasses and, if yes, whether the respondent thought they were fashionable; whether the respondent was familiar with the brand; whether the respondent thought the product reviewed was gender neutral. Respondents were also asked whether they wore glasses themselves. This was followed by a text asking participants to look at the picture of the influencer once again and to complete the list of questions. These questions enquired about physical attractiveness, social attractiveness, perceived intelligence and expertise using a 7-point Likert scale. The order of the questions was again randomized. Finally, respondents

were asked about their thoughts on the purpose of the study, and were able to leave remarks or clarifications of their answers.

3.8.Scales

The full questionnaire (both in Dutch and English) can be found in annex 1. The right column indicates the source of each of the items. All constructs were measured using a 7-point Likert scale, with 1 being "I completely disagree" and 7 "I completely agree".

Independent variables

Physical attractiveness (α=.903) was measured using 4 items (*attractive, sexy, beautiful, classy*) from Ohanian's (1990) source credibility scale. The social attractiveness scale (α=.843) was composed of statements adapted from McCroskey and McCain (1974) (*this person would fit into my circle of friends; it seems easy to have a chat with this person*) and 4 additional items used in Fiske (2018) (*friendly, trustworthy, warm, sincere*). Perceived intelligence was initially (pre-test) measured with one single item (*intelligence*), as was the case in Wei et al. (2017). For the actual survey, perceived intelligence (α=.859) was combined with three items from Ohanian's (1990) expertise scale (*expertise, knowledgeability, experience*).

Dependent variables

The message credibility scale (α=.893) was composed of items from Friman (2010) (*informative, useful, interesting*) and Appelman and Sundar's (2016) scale of message credibility (*accurate, authentic, believable*). The latter has already been successfully used to evaluate message credibility on Instagram (Mena, Barbe & Chan-Olmsted, 2020). Product attitude (α=.854), then, consisted of 3 items (*perception of sound quality, perception of overall quality, perception of workmanship*) adopted from Rao, Qu & Ruekert (1999). Brand attitude (α=.866) was measured using Kruger, Kühn, Petzer, & Mostert's (1991) brand trust scale (*this brand meets my expectations, I feel confidence in this brand, this brand will not disappoint me*). Message credibility, product attitude and brand attitude together constituted consumer attitude (α=.900). Purchase intention (α=.947), finally, was based on the statements used in Lee, Trail, Lee, & Schoenstedt (2013) (*I am interested in this product, I am interested in this brand, I am inclined to consider purchasing this product, I am inclined to consider purchasing this brand*). Three additional items enquiring about purchase intention (α=.920) (*I want to know*

more about this product, I want to know more about this brand, I want to read more reviews about this product) were included in the questionnaire. They were not part of the purchase intention scale, but were added as a separate construct to see whether there might be a bridge between consumer attitude and purchase intention. That is, in this online advertising scenario, SM users have to make their buying decision based on one single review. In real life, there might be some steps in between, but these are not taken into account in the traditional (offline) purchase intention scales.

4. RESULTS ANALYSIS

This section presents the results and statistical analysis of the data retrieved through the experiment. As the central limit theorem (CLT) states that "the distribution of sample means approximates a normal distribution as the sample size gets larger, regardless of the population's distribution" (Ganti, 2022, para. 4) and our average of 70 respondents per scenario considerably exceeds the requirement of 30 for the CLT to hold (ibid.), we could resort to parametric tests for the statistical analysis of the data. These tests were conducted using SPSS 28 and an online calculation tool for mediation tests (quantpsy.org).

First, we verify whether the manipulations which proved successful in the pre-test (glasses and attractiveness) produced the same effects in the actual experiment. Descriptive statistics, subsequently, are included to provide insights into the characteristics of the data set. Next, we test all hypotheses separately. In the final subsection, other surprising or interesting observations are provided.

4.1. Manipulation check

As in our pre-test, Independent-Samples Mann-Whitney U Tests were conducted to verify whether the manipulations produced the desired effects. Again, the manipulations proved successful: influencers wearing glasses were deemed more intelligent (*M*=4.24, *SD*=1.018 vs. *M*=3.46, *SD*=.886, *p*<.001), physically attractive influencers were rated higher in physical attractiveness (*M*=5.28, *SD*=.840 vs. *M*=3.91, *SD*=1.063, *p*=.000), and physically less attractive influencers were seen as more socially attractive (*M*=4.87, *SD*=.710 vs. *M*=4.54, *SD*=.891, *p*<.001).

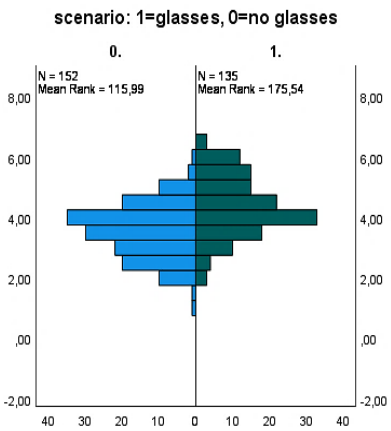


Figure 13: influence glasses on PI

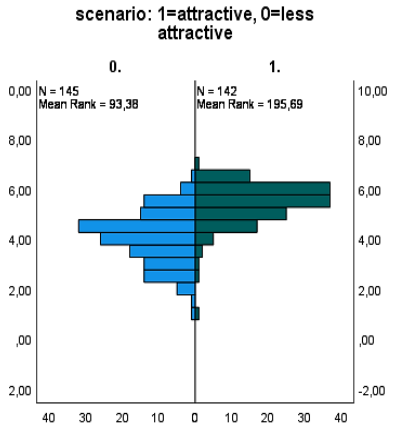


Figure 14: influence PA on perceived PA

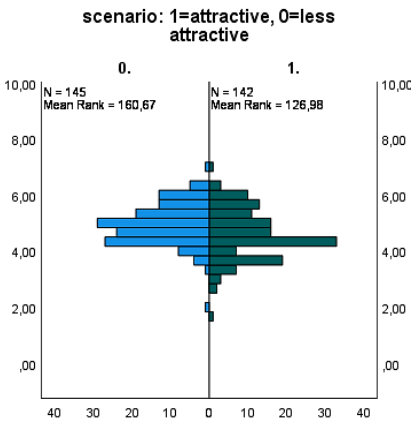


Figure 15: influence PA on perceived SA

4.2. Descriptive statistics

Before resorting to parametric tools to analyse the data, it seems useful to first look at the general trends which can be found in the data. In this way we can already see whether answers by particular groups (gender, education level, etc.) differ from other groups. Of course, any characteristics or tendencies should be met with caution as these results do not differentiate between scenarios. Moreover, certain groups are considerably less represented in the data, so premature generalization needs to be refrained from. Nevertheless, it can be interesting to mention salient results in such cases, as future research could further concentrate on these findings. Finally, descriptive statistics also allow us to generate an overview of mean scores and correlations between variables. These can already give an indication of which results we may expect from the parametric tests.

4.2.1. Mean scores per scenario

- FANG: Female, Attractive, No Glasses
- FAG: Female, Attractive, Glasses
- FLANG: Female, Less Attractive, No Glasses
- FLAG: Female, Less Attractive, Glasses

Figure 16 visualizes the mean scores for the study's key variables across the four scenarios. In general, the scenarios 'FANG' and 'FLANG' on the one hand, and 'FAG' and 'FLAG' on the other hand, display similar tendencies, with the latter obtaining higher scores than the former for all variables except physical attractiveness (PA) and social attractiveness (SA). Interestingly, the attractive female influencer with glasses (FAG) obtains nearly the same average score for social attractiveness as her physically less attractive counterpart without glasses (FLANG) (M=4.71 vs. M=4.74). For perceived intelligence (PI), the FAG scenario even receives the highest mean score of all four conditions (M=4.31).

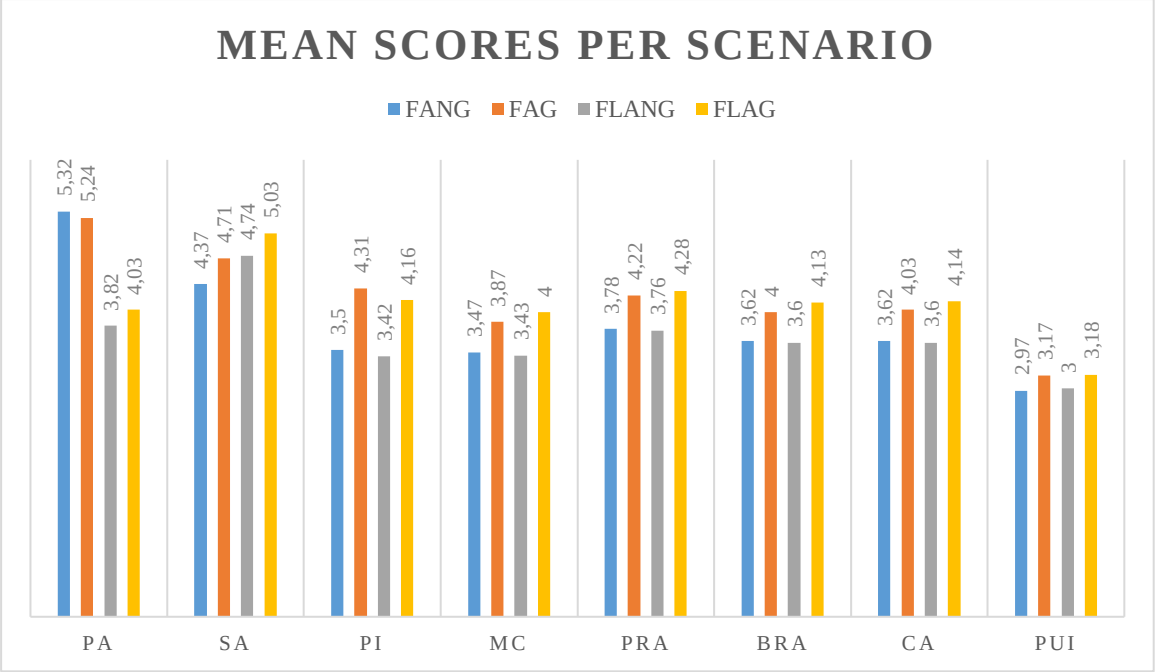


Figure 16: Mean scores per scenario

4.2.2. Group tendencies

As this section mainly serves to highlight striking tendencies or differences between groups, an exhaustive overview of all variables per group seems excessive. Variables for which respondents showed similar tendencies are therefore not included.

In general, the comparisons between groups did not yield many remarkable observations. Most differences remained superficial or did not allow generalizations because of the limited sample sizes (the Eastern and Asian sample each comprise only three respondents). Still, three significant trends emerged from the data which are worth mentioning and further discussed).

First, a comparison of the different age categories revealed that Gen Z (18-26) (M=3.67, SD=1.014) and Gen Y (27-41) (M=4.14, SD=1.172) significantly differed in their score attributions for perceived intelligence (p=.010). This contrast could be meaningful as it might reveal something about the stance of both age groups on influencers.

Male and female respondents, too, varied in their answers. A comparison of these two groups was particularly interesting as the literature suggests that female Instagram users are more likely than male users to be persuaded by influencers (Rebelo, 2017, p. 52), and to impulsively

buy the product advertised (Djafarova et al., 2021, p. 7). While our data suggest that consumer attitude is indeed higher among female respondents (*M*=3.67, *SD*=1.199 vs. *M*=3.93, *SD*=1.011, *p*=.031), purchase intention scores do not differ significantly between male and female participants (*M*=3.02, *SD*=1.408 vs. *M*=3.10, *SD*=1.448, *p*=.326). Lastly, female respondents awarded significantly higher scores for social attractiveness than male respondents (*M*=4.81, *SD*=.770 vs. *M*=4.53, *SD*=.880, *p*=.003).

4.2.3. Correlations between variables

Before testing our hypotheses, we provide a final overview of how all variables in our study are interrelated. Such an overview is especially useful as it discloses not only the direction of the associations (positive or negative), but also their strength, providing valuable insights into the data. Insignificant correlations, on the other hand, signal the absence of a relationship and can be a harbinger of our results in 4.3. However, correlation only shows how the variables move together, whereas regression analysis (cf. 4.3.) will tell how one affects the other.

To assess the linear relationships between variables, a Pearson correlation coefficient was computed for each combination. Although the table below reveals a number of significant correlations, many are fairly weak. They, too, will be mentioned, but the present section mainly serves to highlight the strongest correlations.

Firstly, both perceived intelligence (PI) and social attractiveness (SA) show a positive correlation with message credibility (MC), product attitude (PrA), brand attitude (BrA), consumer attitude (CA), and purchase intention (PuI). The correlations between these variables and social attractiveness are weak, but slightly stronger in relation to perceived intelligence (PI). Social attractiveness and perceived intelligence are positively correlated as well. In fact, this is the strongest correlation either variable has ($r(285)=.494$, $p<.001$), and likely the result of the glasses manipulation (cf. 3.3.2.). As opposed to social attractiveness, perceived intelligence is also positively correlated with physical attractiveness (PA) and the wish for additional information (AddInfo), but again, these correlations are rather weak. Interestingly, there is no negative or positive correlation between physical and social attractiveness. In fact, physical attractiveness appears to be correlated solely with perceived intelligence.

The strongest correlations can be found between consumer attitude and message credibility ($r(285)=.897$, $p<.001$), product attitude ($r(285)=.912$, $p<.001$) and brand attitude ($r(285)=.932$, $p<.001$), which is unsurprising as the latter three constitute the construct consumer attitude. More interesting, however, is the correlation between these variables and purchase intention. Whereas product attitude ($r(285)=.474$, $p<.001$) and brand attitude ($r(285)=.495$, $p<.001$) both display a moderate positive correlation with purchase intention, message credibility shows a fairly strong positive correlation ($r(285)=.605$, $p<.001$). The correlation between consumer attitude and purchase intention is slightly lower ($r(285)=.577$, $p<.001$). Finally, a strong positive correlation can be observed between purchase intention and the wish for additional information ($r(285)=.787$, $p<.001$).

Correlations

		PA	SA	PI	MC	PrA	BrA	CA	PuI	Addinfo
PA	Pearson Correlation	1	,109	,312**	,091	,114	,095	,109	,114	,095
	Sig. (2-tailed)		,064	<,001	,123	,054	,108	,064	,053	,110
	N	287	287	287	287	287	287	287	287	287
SA	Pearson Correlation	,109	1	,494**	,248**	,233**	,198**	,248**	,154**	,111
	Sig. (2-tailed)	,064		<,001	<,001	<,001	<,001	<,001	,009	,060
	N	287	287	287	287	287	287	287	287	287
PI	Pearson Correlation	,312**	,494**	1	,391**	,343**	,340**	,393**	,325**	,229**
	Sig. (2-tailed)	<,001	<,001		<,001	<,001	<,001	<,001	<,001	<,001
	N	287	287	287	287	287	287	287	287	287
MC	Pearson Correlation	,091	,248**	,391**	1	,695**	,747**	,897**	,605**	,458**
	Sig. (2-tailed)	,123	<,001	<,001		<,001	<,001	<,001	<,001	<,001
	N	287	287	287	287	287	287	287	287	287
PrA	Pearson Correlation	,114	,233**	,343**	,695**	1	,815**	,912**	,474**	,267**
	Sig. (2-tailed)	,054	<,001	<,001	<,001		<,001	<,001	<,001	<,001
	N	287	287	287	287	287	287	287	287	287
BrA	Pearson Correlation	,095	,198**	,340**	,747**	,815**	1	,932**	,495**	,274**
	Sig. (2-tailed)	,108	<,001	<,001	<,001	<,001		<,001	<,001	<,001
	N	287	287	287	287	287	287	287	287	287
CA	Pearson Correlation	,109	,248**	,393**	,897**	,912**	,932**	1	,577**	,368**
	Sig. (2-tailed)	,064	<,001	<,001	<,001	<,001	<,001		<,001	<,001
	N	287	287	287	287	287	287	287	287	287
PuI	Pearson Correlation	,114	,154**	,325**	,605**	,474**	,495**	,577**	1	,787**
	Sig. (2-tailed)	,053	,009	<,001	<,001	<,001	<,001	<,001		<,001
	N	287	287	287	287	287	287	287	287	287
Addinfo	Pearson Correlation	,095	,111	,229**	,458**	,267**	,274**	,368**	,787**	1
	Sig. (2-tailed)	,110	,060	<,001	<,001	<,001	<,001	<,001	<,001	
	N	287	287	287	287	287	287	287	287	287

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4: Pearson correlation coefficients

4.3. Hypothesis testing

4.3.1. RQ 1

RQ1: To what extent do physical features (attractiveness and glasses) affect the perceived online credibility of influencers' sponsored reviews on Instagram?

4.3.1.1. Hypothesis 1

H1: *Reviews by influencers wearing glasses are deemed more credible than reviews by influencers without glasses.*

To assess whether glasses positively influenced message credibility, a One-Way ANOVA was performed. This test revealed that there was a statistically significant difference in message credibility between the glasses and no-glasses condition ($F(1, 285) = 11.174$, $p < .001$).

Pairwise comparisons showed that the mean values of the glasses and no-glasses condition differed significantly ($p < .001$, 95% *C.I.* = .200, .772), with reviews by influencers wearing glasses deemed more credible ($M = 3.935$, $SE = .106$) than those by influencers without glasses ($M = 3.448$, $SE = .100$).

Tests of Between-Subjects Effects

Dependent Variable: MC_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	16,895[a]	1	16,895	11,174	<,001
Intercept	3897,311	1	3897,311	2577,663	<,001
SC_Glasses	16,895	1	16,895	11,174	<,001
Error	430,907	285	1,512		
Total	4328,389	287			
Corrected Total	447,802	286			

a. R Squared = ,038 (Adjusted R Squared = ,034)

Estimates

Dependent Variable: MC_scale

scenario: 1=glasses, 0=no glasses	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
0	3,448	,100	3,252	3,645
1	3,935	,106	3,726	4,143

Table 5: results ANOVA glasses-MC

Pairwise Comparisons

Dependent Variable: MC_scale

(I) scenario: 1=glasses, 0=no glasses	(J) scenario: 1=glasses, 0=no glasses	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
0	1	-,486*	,145	<,001	-,772	-,200
1	0	,486*	,145	<,001	,200	,772

Table 6: pairwise comparisons glasses-MC

Tests of Between-Subjects Effects

Dependent Variable: MC_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	54,050[a]	21	2,574	1,732	,026
Intercept	100,174	1	100,174	67,419	<,001
SC_Glasses	4,308	1	4,308	2,899	,090
SC_Glasses * AgeGroup	8,663	6	1,444	,972	,445
SC_Glasses * Gender	6,471	4	1,618	1,089	,362
SC_Glasses * Ethnicity	3,551	4	,888	,598	,665
SC_Glasses * Education	16,110	6	2,685	1,807	,098
Error	393,752	265	1,486		
Total	4328,389	287			
Corrected Total	447,802	286			

a. R Squared = ,121 (Adjusted R Squared = ,051)

Table 7: interaction effects glasses-MC

Although no significant interaction effects were found for education (p=.098), gender (p=.362), ethnicity (p=.665) or age category (p=.445), including the four different factors in the model did nullify the significant effect of the glasses (p=.090). As interaction effects seemed to be the highest for education level, a graph was plotted to depict the differences among groups. The line graph in figure 17 clearly shows a discrepancy between scores attributed by people with primary education and all other groups. However, this is likely due to the very limited number of respondents with primary education (3), which also excludes generalizations.

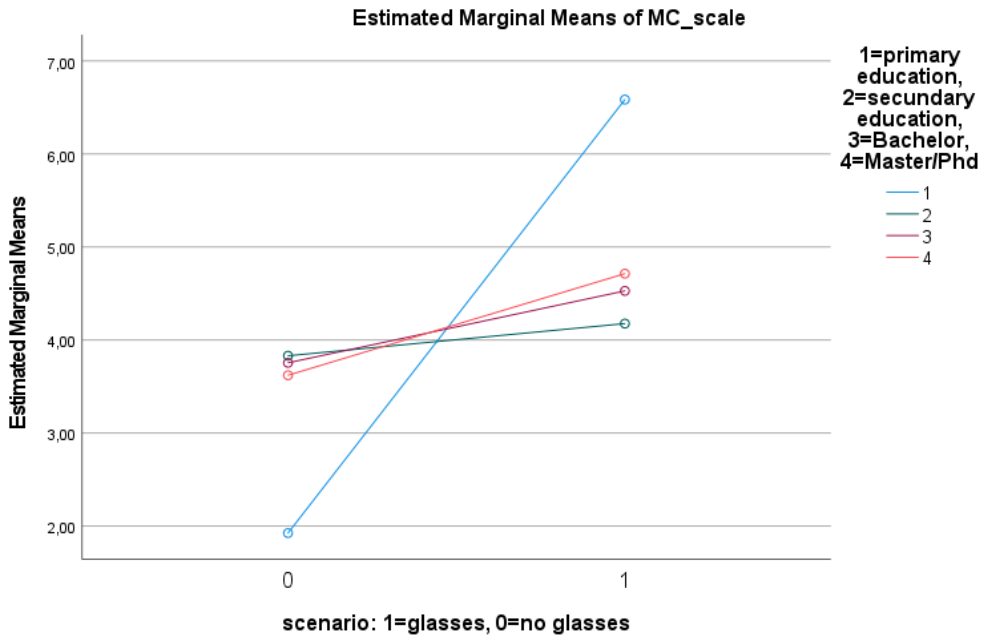


Figure 17: interaction effect education level

For verification purposes, an additional analysis using simple linear regression was conducted to assess whether perceived intelligence (PI), as a scale, significantly enhanced message credibility (MC). The fitted regression model was: MC = 1.854 + .477*PI. The overall regression was statistically significant (R^2 = .153, $F(1, 285) = 51.360$, $p<.001$) and perceived intelligence was shown to significantly predict message credibility (β = .477, $p<.001$).

The analysis thus revealed a significant, directly proportional relationship between perceived intelligence and message credibility. No interaction effects were found for education ($p=.905$), gender ($p=.691$), ethnicity ($p=.425$) or age category ($p=.676$).

Tests of Between-Subjects Effects

Dependent Variable: MC_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	68,377[a]	1	68,377	51,360	<,001
Intercept	65,939	1	65,939	49,529	<,001
PI	68,377	1	68,377	51,360	<,001
Error	379,425	285	1,331		
Total	4328,389	287			
Corrected Total	447,802	286			

a. R Squared = ,153 (Adjusted R Squared = ,150)

Parameter Estimates

Dependent Variable: MC_scale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	1,854	,263	7,038	<,001	1,335	2,372
PI	,477	,067	7,167	<,001	,346	,608

Tests of Between-Subjects Effects

Dependent Variable: MC_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	74,280[a]	11	6,753	4,972	<,001
Intercept	62,348	1	62,348	45,903	<,001
PI	26,747	1	26,747	19,692	<,001
Education * PI	,764	3	,255	,187	,905
Ethnicity * PI	2,332	2	1,166	,858	,425
Gender * PI	1,005	2	,503	,370	,691
AgeGroup * PI	2,074	3	,691	,509	,676
Error	373,522	275	1,358		
Total	4328,389	287			
Corrected Total	447,802	286			

a. R Squared = ,166 (Adjusted R Squared = ,133)

Table 8: regression and interaction effects PI-MC

In brief, H1 can be confirmed only partially: glasses indeed enhance message credibility, but the effect is no longer significant when age, gender, education and ethnicity are taken into account. Although none of the interaction effects is significant by itself, including all four variables in the model obscures the main effect of the glasses. The p- and F-values (p=.098, F= 1.807) do suggest that education level exerts the greatest influence on the relationship between glasses and message credibility, but this is likely due to the very limited number of respondents in the group with primary education (cf. section 7). The influence of perceived intelligence, on the other hand, did remain significant: as perceived intelligence increases, so does message credibility, regardless of the aforementioned variables.

4.3.1.2. Hypothesis 2

H2: *Reviews by attractive influencers are deemed more credible than reviews by less attractive influencers.*

H2a: *Reviews by physically attractive influencers are deemed more credible than reviews by less physically attractive influencers.*

Simple linear regression was used to assess whether physical attractiveness significantly enhanced message credibility. The fitted regression model was: MC = 3.233 + .097*PA. The overall regression was not statistically significant (R^2 =.008, $F(1, 285) = 2.391$, p=.123) and physical attractiveness appeared not to significantly predict message credibility (β =.097, p=.123).

Tests of Between-Subjects Effects

Dependent Variable: MC_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3,725[a]	1	3,725	2,391	,123
Intercept	185,369	1	185,369	118,966	<,001
PA	3,725	1	3,725	2,391	,123
Error	444,077	285	1,558		
Total	4328,389	287			
Corrected Total	447,802	286			

a. R Squared = ,008 (Adjusted R Squared = ,005)

Parameter Estimates

Dependent Variable: MC_scale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	3,233	,296	10,907	<,001	2,650	3,817
PA	,097	,063	1,546	,123	-,026	,220

Table 9: regression PA-MC

H2b: *Reviews by socially attractive influencers are deemed more credible than reviews by less socially attractive influencers.*

Simple linear regression was used to test whether social attractiveness (SA) significantly predicted message credibility (MC). The fitted regression model was: MC = 1.897 + .378*SA.

The overall regression was statistically significant (R^2 = .061, $F(1, 285) = 18.661$, p <.001) and social attractiveness was shown to significantly predict message credibility (β =.378, p <.001).

The analysis thus revealed a significant, directly proportional relationship between social attractiveness and message credibility. No interaction effects were found for education (p=.883), gender (p=.858), ethnicity (p=.239) or age category (p=.648).

Tests of Between-Subjects Effects

Dependent Variable: MC_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	27,450[a]	1	27,450	18,611	<,001
Intercept	30,246	1	30,246	20,507	<,001
SA	27,450	1	27,450	18,611	<,001
Error	420,352	285	1,475		
Total	4328,389	287			
Corrected Total	447,802	286			

a. R Squared = ,061 (Adjusted R Squared = ,058)

Parameter Estimates

Dependent Variable: MC_scale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	1,897	,419	4,528	<,001	1,072	2,721
SA	,378	,088	4,314	<,001	,206	,551

Tests of Between-Subjects Effects

Dependent Variable: MC_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	34,738[a]	11	3,158	2,102	,020
Intercept	29,329	1	29,329	19,526	<,001
SA	20,778	1	20,778	13,833	<,001
Education * SA	,990	3	,330	,220	,883
Ethnicity * SA	4,316	2	2,158	1,437	,239
AgeGroup * SA	2,480	3	,827	,550	,648
Gender * SA	,460	2	,230	,153	,858
Error	413,064	275	1,502		
Total	4328,389	287			
Corrected Total	447,802	286			

a. R Squared = ,078 (Adjusted R Squared = ,041)

Table 10: regression and interaction effects SA-MC

In brief, we can confirm H2b, but have to refute H2a: whereas message credibility increases with social attractiveness, no such relationship can be found for the physical attractiveness of the influencer.

4.3.2. RQ 2

RQ 2: How does the perceived credibility of the influencer affect purchase intention?

4.3.2.1. Hypothesis 3

H3: *Higher perceived source credibility improves consumer attitude*.

H3a: *Perceived intelligence enhances the perceived credibility of influencers and therefore positively influences consumer attitude.*

Simple linear regression was used to test whether perceived intelligence (PI) significantly predicted consumer attitude (CA). The fitted regression model was: CA = 2.239 + .417*PI. The overall regression was statistically significant (R^2 = .154, $F(1, 285) = 52.041$, $p < .001$) and perceived intelligence was shown to significantly predict consumer attitude ($\beta = .417$, $p < .001$).

The analysis thus revealed a significant, directly proportional relationship between perceived intelligence and consumer attitude. No interaction effects were found for education (p=.923), gender (p=.380), ethnicity (p=.703) or age category (p=.846).

Tests of Between-Subjects Effects

Dependent Variable: CA_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	52,157[a]	1	52,157	52,041	<,001
Intercept	96,170	1	96,170	95,957	<,001
PI	52,157	1	52,157	52,041	<,001
Error	285,635	285	1,002		
Total	4550,414	287			
Corrected Total	337,792	286			

a. R Squared = ,154 (Adjusted R Squared = ,151)

Parameter Estimates

Dependent Variable: CA_scale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	2,239	,229	9,796	<,001	1,789	2,688
PI	,417	,058	7,214	<,001	,303	,530

Tests of Between-Subjects Effects

Dependent Variable: CA_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	56,158[a]	11	5,105	4,985	<,001
Intercept	88,835	1	88,835	86,743	<,001
PI	17,109	1	17,109	16,706	<,001
Education * PI	,490	3	,163	,160	,923
Ethnicity * PI	,723	2	,361	,353	,703
Gender * PI	1,987	2	,993	,970	,380
AgeGroup * PI	,832	3	,277	,271	,846
Error	281,634	275	1,024		
Total	4550,414	287			
Corrected Total	337,792	286			

a. R Squared = ,166 (Adjusted R Squared = ,133)

Table 11: regression and interaction effects PI-CA

H3b: *Physical attractiveness enhances the perceived credibility of influencers and therefore positively influences consumer attitude*.

Simple linear regression was used to assess whether physical attractiveness (PA) significantly enhanced consumer attitude (CA). The fitted regression model was: CA = 3.369 + .101*PA. The overall regression was not statistically significant (R^2 = .012, $F(1, 285) = 3.449$, p=.064) and physical attractiveness appeared not to significantly predict consumer attitude (β = .101, p=.064).

Tests of Between-Subjects Effects

Dependent Variable: CA_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4,039[a]	1	4,039	3,449	,064
Intercept	201,258	1	201,258	171,860	<,001
PA	4,039	1	4,039	3,449	,064
Error	333,752	285	1,171		
Total	4550,414	287			
Corrected Total	337,792	286			

a. R Squared = ,012 (Adjusted R Squared = ,008)

Parameter Estimates

Dependent Variable: CA_scale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	3,369	,257	13,110	<,001	2,863	3,875
PA	,101	,054	1,857	,064	-,006	,208

Table 12: regression PA-CA

H3c: *Social attractiveness enhances the perceived credibility of influencers and therefore positively influences consumer attitude.*

Simple linear regression was used to test whether social attractiveness (SA) significantly predicted consumer attitude (CA). The fitted regression model was: CA = 2.282 + .329*SA. The overall regression was statistically significant (R^2 = .062, $F(1, 285) = 18.699$, p <.001) and social attractiveness was shown to significantly predict consumer attitude (β = .320, p <.001).

The analysis thus revealed a significant, directly proportional relationship between social attractiveness and consumer attitude. No interaction effects were found for education (p=.914), gender (p=.517), ethnicity (p=.472) or age category (p=.950).

Tests of Between-Subjects Effects

Dependent Variable: CA_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	20,798[a]	1	20,798	18,699	<,001
Intercept	43,759	1	43,759	39,343	<,001
SA	20,798	1	20,798	18,699	<,001
Error	316,994	285	1,112		
Total	4550,414	287			
Corrected Total	337,792	286			

a. R Squared = ,062 (Adjusted R Squared = ,058)

Parameter Estimates

Dependent Variable: CA_scale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	2,282	,364	6,272	<,001	1,566	2,997
SA	,329	,076	4,324	<,001	,179	,479

Tests of Between-Subjects Effects

Dependent Variable: CA_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	24,875[a]	11	2,261	1,987	,030
Intercept	43,344	1	43,344	38,092	<,001
SA	13,467	1	13,467	11,835	<,001
Education * SA	,595	3	,198	,174	,914
Ethnicity * SA	1,714	2	,857	,753	,472
AgeGroup * SA	,397	3	,132	,116	,950
Gender * SA	1,503	2	,752	,661	,517
Error	312,917	275	1,138		
Total	4550,414	287			
Corrected Total	337,792	286			

a. R Squared = ,074 (Adjusted R Squared = ,037)

Table 13: regression and interaction effects SA-CA

In short, the results support H3a and H3c, but not H3b: consumer attitude increases with the perceived intelligence and social attractiveness of the source, but not with physical attractiveness.

4.3.2.2. Hypothesis 4

H4: *Consumer attitude mediates the relationship between perceived source credibility and purchase intention.*

To assess whether the influence of perceived intelligence, physical attractiveness and social attractiveness on purchase intention was mediated by consumer attitude, linear as well as multiple regression analyses were performed, followed by a Sobel Test.

H4a: *The effect of perceived intelligence on purchase intention is mediated by consumer attitude*.

Simple linear regression showed that perceived intelligence significantly predicted purchase intention (β =.454, p<.001), with purchase intention directly proportional to perceived intelligence. The analysis of the indirect effects revealed that this relationship was significantly mediated by consumer attitude (β =.219, p=.000). Perceived intelligence positively affected consumer attitude (β =.417, p<.001), and consumer attitude in turn positively influenced purchase intention (β =.699, p<.001). However, multiple regression also revealed that, after accounting for the mediating role of consumer attitude, perceived intelligence still significantly influenced purchase intention (β =.163, p=.026).

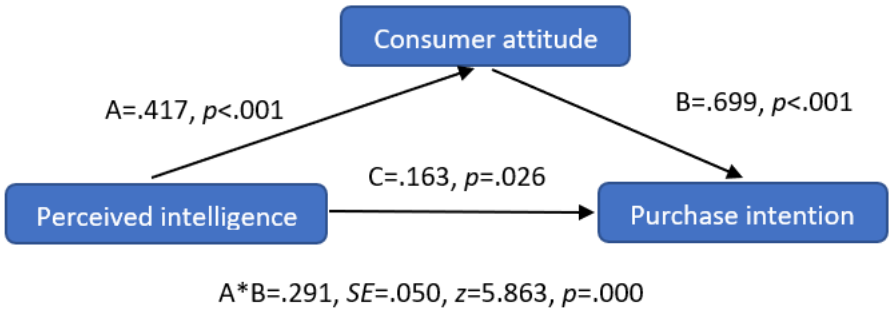


Figure 18: (in)direct effect of PI on PuI through CA

The analysis thus revealed that the effect of perceived intelligence on purchase intention was partially mediated by consumer attitude.

H4b: *The effect of physical attractiveness on purchase intention is mediated by consumer attitude*.

Simple linear regression revealed that physical attractiveness did not significantly predict purchase intention (β =.139, p=.053). As there was no significant total effect, no mediation analysis was performed.

Tests of Between-Subjects Effects

Dependent Variable: Pul_scale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7,665[a]	1	7,665	3,781	,053
Intercept	105,334	1	105,334	51,965	<,001
PA	7,665	1	7,665	3,781	,053
Error	577,699	285	2,027		
Total	3297,438	287			
Corrected Total	585,364	286			

a. R Squared = ,013 (Adjusted R Squared = ,010)

Parameter Estimates

Dependent Variable: Pul_scale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	2,437	,338	7,209	<,001	1,772	3,103
PA	,139	,071	1,945	,053	-,002	,279

Table 14: regression PA-PuI

H4c: *The effect of social attractiveness on purchase intention is mediated by consumer attitude.*

Simple linear regression showed that social attractiveness significantly predicted purchase intention (β =.270, p=.009), with purchase intention directly proportional to social attractiveness. The analysis of the indirect effects revealed that this relationship was significantly mediated by consumer attitude (β =.249, p<.001). Social attractiveness positively affected consumer attitude (β =.329, p<.001), and consumer attitude in turn positively influenced purchase intention (β =.756, p<.001). Multiple regression showed that, after accounting for the mediating role of consumer attitude, the influence of social attractiveness on purchase intention was insignificant (β =.021, p=.810).

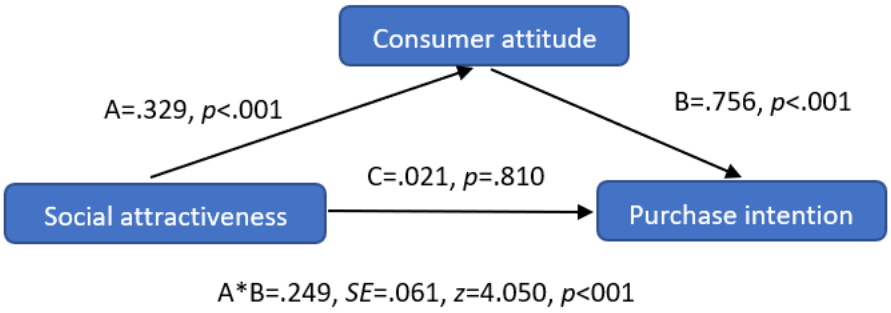


Figure 19: indirect effect of SA on PuI through CA

The analysis thus revealed an effect of social attractiveness on purchase intention which was fully mediated by consumer attitude.

Summary

Before elaborating on the additional findings in this study, we first provide a brief overview of the main hypotheses and to which extent they are supported by the data.

H1: *Reviews by influencers wearing glasses are deemed more credible than reviews by influencers without glasses.*

Hypothesis 1 is only partially supported by our findings. While glasses indeed enhanced message credibility, this positive effect was no longer significant when the interaction effects of gender, age, ethnicity and education were all taken into account. However, the positive influence of perceived intelligence, fostered by wearing glasses (cf. 4.1.), did remain significant.

H2: *Reviews by attractive influencers are deemed more credible than reviews by less attractive influencers.*

Whereas social attractiveness significantly boosted message credibility (H2b), physical attractiveness did not (H2a). H2 is therefore only partially confirmed.

H3: *Higher perceived source credibility improves consumer attitude*.

Both perceived intelligence and social attractiveness of the source significantly enhanced consumer attitude, supporting H3a and H3c. H3b, however, had to be refuted, as we were unable to find a significant impact of physical attractiveness on consumer attitude. H3 is therefore only partially confirmed.

H4: *Consumer attitude mediates the relationship between perceived source credibility and purchase intention.*

Consumer attitude partially mediated the influence of perceived intelligence, and fully mediated the influence of social attractiveness on purchase intention, supporting H4a and H4c.

Still, H4 was not fully confirmed as no significant relationship was found between physical attractiveness and purchase intention (H4b).

4.4. Additional observations

This final analysis subsection serves to highlight other remarkable observations which did not arise from the descriptive statistics or relate directly to the central research questions. More in particular, we examined whether an effect could be found for homophily, whether purchase intention and the wish for additional information differed significantly as scales, and, finally, whether a combined effect could be found of glasses on the one hand and weight and blemishes on the other hand.

4.4.1. Homophily

To assess homophily, a One-Way ANOVA for age and gender was conducted, followed by a Two-Way ANOVA for glasses. These were the three indicators of homophily the respondents would have been able to infer from the influencer's profile picture. A One-Way ANOVA did not suffice to assess the influence of glasses, as the condition that bespectacled participants were assigned to also played a key role. Gender and age, on the other hand, remained constant across all scenarios.

The One-Way ANOVA for age category revealed no significant differences between groups for message credibility ($F(3, 283)=.151$, $p=.929$), social attractiveness ($F(3, 283)=.603$, $p=.614$), physical attractiveness ($F(3, 283)=1.511$, $p=.212$), or perceived intelligence $F(3, 283)=2.312$, $p=.076$).

For gender, the One-Way ANOVA again showed no significant differences in message credibility ($F(2, 284)=.778$, $p=.460$), physical attractiveness ($F(2, 284)=.186$, $p=.830$), or perceived intelligence $F(2, 284)=.654$, $p=.521$). For social attractiveness, however, scores differed significantly between groups ($F(2, 284)=3.934$, $p=.021$): male and female respondents diverged in their answers ($p<.006$, 95% *C.I.* = .082, .474), with female participants attributing higher scores ($M=4.811$, $SE=.061$) than male respondents ($M=4.533$, $SE=.079$).

Finally, the Two-Way ANOVA revealed no significant interaction between the effects of participants wearing glasses and the type of scenario evaluated (glasses vs. no glasses) for social attractiveness ($F(3, 283)=.145$, $p=.704$), perceived intelligence ($F(3, 283)=.568$, $p=.452$), message credibility ($F(3, 283)=.321$, $p=.572$) or physical attractiveness ($F(3, 283)=.112$, $p=.738$).

In summary, close to no effects could be found for homophily. Respondents wearing glasses did not attribute higher scores to influencers wearing glasses, and neither did respondents in the same age category as the influencers. The One-Way ANOVA for gender did reveal a disparity in social attractiveness scores attributed by male and female respondents. The latter rated the female influencers higher in social attractiveness than the male participants, signalling a possible influence of homophily. However, this cannot be said with certainty: our experiment contained no male stimuli, which would have enabled comparing male and female scores for influencers of both genders. As a result, we cannot rule out that the variation in scores is ascribable to women attributing higher social attractiveness scores in general.

4.4.2. Purchase intention vs. additional information

To compare the scores for purchase intention and the wish for additional information, a paired samples t-test was conducted. The test revealed a significant difference in scores between purchase intention ($M=3.07$, $SD=1.431$) and the wish for additional information ($M=4.10$, $SD=1.832$); $t(286) = -15.385$, $p<.001$.

Paired Samples Test

		Paired Differences					t	df	Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1	Pul_scale - AddInfo_scale	-1,02584	1,12957	,06668	-1,15708	-,89460	-15,385	286	<,001	<,001

Table 15: results paired t-test

The significantly higher score for additional information suggests that there might be an important stage between consumer attitude and the intention to purchase a product. Today still, studies typically use the traditional purchase intent scale, which enquires about consumers' interest in a product or brand, and how likely they are to consider buying it. This scale does, however, not take into account the possibility to look up additional information or reviews, a

practice which is widely prevalent today. Although a modification of the purchase intent scale was not this study's primary aim, we hope these exploratory findings can encourage more research into a modernized version of the scale, which is more adjusted to current consumer behaviour (see also section 7).

4.4.3. Combined manipulation effects

Our results have already indicated that glasses on the one hand, and weight gain and blemishes on the other hand significantly increase social attractiveness (cf. 4.1.). However, at this point it is not yet clear whether a combination of both manipulations is statistically more beneficial than only glasses, or only blemishes and weight gain. To this aim, we performed a final Two-Way ANOVA, examining whether combined effects of our manipulations can be found for social attractiveness. For verification purposes, physical attractiveness and perceived intelligence were measured for all scenarios as well.

As expected, physical attractiveness did not differ significantly within each attractiveness condition depending on whether the influencer wore glasses or not: the Two-Way ANOVA revealed a statistically significant difference between at least two scenarios ($F(3,283) = 49.621$, $p<.001$), but pairwise comparisons showed that the value of physical attractiveness differed significantly only within the glasses conditions (($p<.001$, 95% *C.I.* = 1.189, 1.802) for the glasses condition and ($p<.001$, 95% *C.I.* =.910, 1.560) for the no-glasses condition), but not within the attractiveness condition (($p=.232$, 95% *C.I.* = -.123, .507) for the less-attractive condition and ($p=.670$, 95% *C.I.* = -.386, .248) for the attractive condition).

In brief, scores for physical attractiveness were not affected by glasses, and only differed significantly between the 'physically attractive' and 'physically less attractive' influencers.

Pairwise Comparisons

Dependent Variable: PA_scale

scenario: 1=glasses, 0=no glasses	(I) scenario: 1=attractive, 0=less attractive	(J) scenario: 1=attractive, 0=less attractive	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
0	0	1	-1,495*	,156	<,001	-1,802	-1,189
	1	0	1,495*	,156	<,001	1,189	1,802
1	0	1	-1,235*	,165	<,001	-1,560	-,910
	1	0	1,235*	,165	<,001	,910	1,560

Based on estimated marginal means

*. The mean difference is significant at the 0,05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Pairwise Comparisons

Dependent Variable: PA_scale

scenario: 1=attractive, 0=less attractive	(I) scenario: 1=glasses, 0=no glasses	(J) scenario: 1=glasses, 0=no glasses	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
0	0	1	-,192	,160	,232	-,507	,123
	1	0	,192	,160	,232	-,123	,507
1	0	1	,069	,161	,670	-,248	,386
	1	0	-,069	,161	,670	-,386	,248

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 16: pairwise comparisons PA

Perceived intelligence also did not differ depending on the attractiveness condition: the Two-Way ANOVA revealed a statistically significant difference between at least two scenarios ($F(3,283) = 16.459$, $p<.001$), but pairwise comparisons showed that the value of perceived intelligence differed significantly only within the attractiveness conditions (($p<.001$, 95% *C.I.* = .407, 1.031) for the less-attractive condition and ($p<.001$, 95% *C.I.* =.515, 1.144) for the attractive condition), but not within the glasses condition (($p=.257$, 95% *C.I.* = -.136, .508) for the glasses condition and ($p=.624$, 95% *C.I.* = -.228, .380) for the no-glasses condition).

In short, 'physically attractive' influencers were not rated higher in perceived intelligence than 'physically less attractive' influencers, or vice versa. Bespectacled influencers, on the other hand, were attributed significantly higher scores than their counterparts without glasses.

Pairwise Comparisons

Dependent Variable: PI_scale

scenario: 1=attractive, 0=less attractive	(I) scenario: 1=glasses, 0=no glasses	(J) scenario: 1=glasses, 0=no glasses	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
0	0	1	-,719*	,159	<,001	-1,031	-,407
	1	0	,719*	,159	<,001	,407	1,031
1	0	1	-,830*	,160	<,001	-1,144	-,515
	1	0	,830*	,160	<,001	,515	1,144

Based on estimated marginal means

*. The mean difference is significant at the 0,05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Pairwise Comparisons

Dependent Variable: PI_scale

scenario: 1=glasses, 0=no glasses	(I) scenario: 1=attractive, 0=less attractive	(J) scenario: 1=attractive, 0=less attractive	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
0	0	1	-,076	,154	,624	-,380	,228
	1	0	,076	,154	,624	-,228	,380
1	0	1	-,186	,164	,257	-,508	,136
	1	0	,186	,164	,257	-,136	,508

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 17: pairwise comparisons PI

For social attractiveness, however, the Two-Way ANOVA revealed combined effects of the manipulations. First, blemishes and weight (p<.001) and glasses (p<.001) were both shown to significantly impact social attractiveness. Within the attractiveness condition, then, social attractiveness differed significantly as well, depending on whether the influencers wore glasses or not: (p=.036, 95% *C.I.* = .018, .537) for the less attractive influencers, with bespectacled influencers rated higher in social attractiveness (M=5.020, SE=.097) than influencers without glasses, and (p=.0308, 95% *C.I.* = .093, .616) for the attractive influencers, with bespectacled influencers again rated higher in social attractiveness (M=4.727, SE=.095) than influencers without glasses.

An additional independent samples t-test revealed that the attractive influencer wearing glasses (M=4.71, SD=.822) and the less attractive influencer without glasses (M=4.74, SD=.725) were not attributed significantly different scores (p=.818).

In summary, the findings revealed combined effects of glasses, blemishes and weight gain on social attractiveness. Whereas the 'less attractive' influencer with glasses (FLAG) received the highest scores for social attractiveness, the attractive influencer without glasses (FANG) received the lowest social attractiveness scores. The influencers with only glasses (FAG) or

only blemishes and weight gain (FLANG) did not significantly differ in social attractiveness, suggesting that both manipulations carry approximately the same weight.

Estimates

Dependent Variable: SA_scale

scenario: 1=attractive, 0=less attractive	scenario: 1=glasses, 0=no glasses	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
0	0	4,743	,089	4,567	4,918
	1	5,020	,097	4,828	5,212
1	0	4,372	,093	4,190	4,554
	1	4,727	,095	4,540	4,915

Pairwise Comparisons

Dependent Variable: SA_scale

scenario: 1=attractive, 0=less attractive	(I) scenario: 1=glasses, 0=no glasses	(J) scenario: 1=glasses, 0=no glasses	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
0	0	1	-,278*	,132	,036	-,537	-,018
	1	0	,278*	,132	,036	,018	,537
1	0	1	-,355*	,133	,008	-,616	-,093
	1	0	,355*	,133	,008	,093	,616

Based on estimated marginal means

*. The mean difference is significant at the 0,05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 18: pairwise comparisons SA

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
						Significance		Mean Difference	Std. Error Difference
		F	Sig.	t	df	One-Sided p	Two-Sided p		
SA_scale	Equal variances assumed	2,592	,110	-,230	145	,409	,818	-,02938	,12763
	Equal variances not assumed			-,228	134,875	,410	,820	-,02938	,12883

Table 19: t-test SA

5. DISCUSSION

The present section discusses the results of the statistical analyses in section 4. We consider the implications of these results for our hypotheses and compare our findings with previous research. Additionally, we include some of the remarks left by respondents at the end of the survey (cf. 3.7.). These remarks either contradict or substantiate our own and previous findings, and are included mainly to provide more insight into what motivated respondents' answers. However, these comments should be met with considerable caution; answering the final question of the survey (*"Do you have any remarks or do you wish to elaborate on any of your answers*?") was optional, and only 30 out of the 336 participants formulated remarks which served to clarify their answers. Moreover, the question did not enquire about a specific topic, and answers naturally varied. As a result, no conclusions can be drawn from these comments and neither should they be seen as representative of all respondents' views. Nevertheless, allowing respondents to elaborate on their motivations can provide valuable insights, and it might be useful to give a more prominent place to such open questions in follow-up research (see also section 7).

RQ1: To what extent do physical features (attractiveness and glasses) affect the perceived online credibility of influencers' sponsored reviews on Instagram?

Hypothesis 1 stated that glasses would enhance message credibility. The results showed that reviews by influencers wearing glasses were indeed deemed significantly more credible than reviews by those without glasses. However, the significance did not hold after incorporating respondents' gender, age category, education and ethnicity in the model. Hypothesis 1 was thus confirmed only to an extent. Interestingly, none of these interaction effects was significant. Although respondents' education level seemed to affect the relationship between glasses and message credibility the most, a separate graph revealed that the interaction was due to extreme answers from the group with primary education, which only counted three respondents. This not only makes it impossible to make generalizations, it also raises the question whether the main effect of glasses would have held with a larger or more evenly distributed sample size.

Although not a physical feature, we also examined how perceived intelligence influenced message credibility. This seems all the more relevant as our results showed glasses to be significant indicators of perceived intelligence. Even after taking into account the possible

interaction effects, perceived intelligence was shown to significantly predict message credibility, meaning that message credibility increases with perceived intelligence. These results could also point in the direction of a possible indirect effect of glasses on message credibility, mediated by perceived intelligence. This was not examined in the current study, but could be addressed by future studies on the topic.

The finding that wearing glasses can (to an extent) positively influence message credibility is in line with former studies indicating that glasses increase perceptions of competence (Fetscherin et al., 2020, p. 341) and expertise (van der Land et al., 2015, p. 181). Our results do, however, contradict previous research suggesting that higher source credibility in influencers does not generate higher message credibility (Lee et al., 2020, p. 244).

Additionally, our results confirm the finding from Leder et al. that thin-rimmed, fashionable glasses increase perceived intelligence, but do not influence physical attractiveness (2011, p. 220). Glasses in our study thus did not enhance physical attractiveness, as reported in van der Land et al. (2015, p. 181). Our results did reveal, however, that glasses are indicators of social attractiveness. This seems to be in line with previous research in which glasses increased levels of warmth (Fetscherin et al., 2020, p. 341) and trustworthiness (Leder et al., 2011, p. 220).

Hypothesis 2 stated that both physical and social attractiveness of the influencer would positively influence message credibility. Whereas social attractiveness (H2b) significantly predicted message credibility, physical attractiveness (H2a) did not. Hypothesis 2 was thus only partially supported by our data. Our findings suggest that consumers are influenced by a social attractiveness halo effect when attributing message credibility, but not a physical attractiveness halo effect. This is not in line with results from Ozanne et al., who found that physically attractive reviewers generate a halo effect in positive reviews (2019, p. 733). This could, however, be due to the sponsored nature of the Instagram review. The fact that influencers are approached by brands and paid for their reviews might evoke a state of alertness, leading to central processing (cf. 2.2.1.). As attractiveness of the reviewer is a heuristic cue (Ozanne et al., p. 729), such a cognitive approach could undermine its halo effect.

Although previous research put forward that the #sponsored reference does not harm the review's or influencer's credibility (Djafarova et al., 2017; Lee et al., 2020), several comments

formulated by respondents in our study suggest otherwise. That is, certain participants explicitly mentioned that the review being a paid promotion harmed its credibility. Respondent 277, for example, stated: "In vele gevallen heb ik het gevoel dat influencers betaald worden om een product te promoten of het gratis krijgen waardoor de geloofwaardigheid deels wegvalt" [In many cases I feel like influencers are paid to promote a product or get it for free, which makes it lose some of its credibility], and participant 63 affirmed: "Heb een groot wantrouwen voor influencers die gesponsord worden om goederen te promoten. Te weinig transparant" [I greatly distrust influencers who are sponsored to promote goods. They are too little transparant].

Still, this does not necessarily explain why social attractiveness did influence message credibility, as this was also derived from the profile picture. It could be the case that the #sponsored reference does not entirely offset the reliance on heuristic cues, but does nevertheless raise caution. As social attractiveness, as opposed to physical attractiveness, is not immediately associated with persuasive intent and marketing strategies, that could explain why the former is more effective than the latter.

RQ 2: How does the perceived credibility of the influencer affect purchase intention?

Next, perceived intelligence and social attractiveness of the source both appeared to be significant predictors of consumer attitude, so H3a and H3c were confirmed. H3b, on the other hand, was rejected, as physical attractiveness did not significantly predict consumer attitude. This finding contradicts results from previous literature, which showed physical attractiveness, as opposed to expertise and trustworthiness, to be the only source credibility dimension positively influencing consumer attitude (Lim et al., 2017). Results from Ozanne et al., too, suggest that attractive reviewers enhance brand evaluations (2019).

However, the study by Lim et al. also stressed the importance of product match-up for such a positive relation (2017, p. 30). Although a product match-up was included in our review (the influencers mentioned that they were fervent podcast listeners, and that the earphones allowed them to walk around the house while doing so), we also tried to keep the product as gender neutral as possible (cf. 3.2.). This might have resulted in a product match-up which was less effective than an attractive (female) model promoting skin care, for example. However, such

an approach would likely have engendered a disparity in male and female scores for product attitude and purchase intention.

Several remarks left at the end of the survey also seem to corroborate that lack of expertise with regard to the product hinders a favourable consumer attitude (Lim et al., 2017, p. 29). One respondent (63) even expressed that good looks do not equal expertise: “Een Jeroen Meus zou mij een pan kunnen verkopen, een knappe jongedame of -heer zonder bekende achtergrond niet” [Someone like Jeroen Meus [popular Belgian TV chef] could sell me a frying pan, a handsome young man or lady without any prestige with regard to the subject could not]. In parallel, participant 112 wrote: “Het is niet mogelijk om via de foto te beoordelen of deze persoon kennis heeft over het product” [It is not possible to derive from the photo whether this person is knowledgeable about the product]. This perceived lack of experience with regard to the product might have been remedied by providing more information on the influencer (in the form of a biography, for example), but this would have also made the scenarios subject to more bias. Moreover, the fact that social attractiveness and glasses did improve consumer attitude may indicate that they provide sufficient levels of perceived expertise not to raise caution.

Further, mediation analyses revealed that perceived intelligence and social attractiveness influenced purchase intention through consumer attitude. While perceived intelligence produced a direct effect on purchase intention which was partially mediated by consumer attitude, the effect of social attractiveness on purchase intention was fully mediated. These findings supported H4a and H4c. H4b was rejected as no total effect could be found of physical attractiveness on purchase intention.

The literature on how influencers impact purchase intention is not unanimous: whereas Lee et al. found that influencers’ source credibility does not predict purchase intention (2020, p. 244), Rebelo suggests purchase intention increases when perceived source credibility is high (2017, p. 51). Our results seem to support the latter, although we could not corroborate the finding that physical attractiveness positively influences purchase intention, as reported in the same study (p. 52) and in Balaban et al. (2019). Our findings do, however, mirror those from Lim et al., who also were unable to find any influence from physical attractiveness on purchase intention (2017, p. 28).

In addition, Rebelo found that expertise did not affect purchase intention (2017, p. 52), whereas Balaban et al. highlighted its role (2019, p. 41). The aforementioned remarks made by respondents provide additional support for the latter, as the need for expertise was explicitly brought up. Furthermore, our findings were not able to validate the importance of homophily for purchase intention (Balaban et al., p. 41). This might, however, be due to our limited analysis of homophily, or the different research approach used. That is, the interview technique employed by Balaban et al. enabled enquiring about influencers whom the participants already followed.

Further, our results support the finding that female SM users are more likely to be influenced by influencer reviews (Rebelo, 2017, p. 52) but not that they are more easily encouraged to impulse buying (Djafarova et al., 2021, p. 7); consumer attitude is higher among female participants, but purchase intention and the wish for additional information do not significantly differ between genders. There are several factors which might explain this discrepancy. First, Djafarova et al. conducted interviews in which participants were asked specifically about their impulse buying behaviour, whereas our approach was experimental and enquired about purchase intention, which is broader and less extreme. Secondly, the influencer in our experiment promoted (gender-neutral) earphones, while the study by Djafarova et al. focused on fashion. It is likely that the choice of product also considerably influenced the results.

Finally, additional testing revealed that blemishes and weight gain on the one hand, and glasses on the other hand, increase social attractiveness equally. This is especially remarkable as the former are typically associated with reducing physical attractiveness (Chithambo et al., 2013; Fink et al., 2001). The ‘physically less attractive’ influencer who also wore glasses was evaluated as most socially attractive. The ‘physically attractive’ influencer without glasses, on the other hand, received the lowest social attractiveness scores. Effects of the two types of manipulations thus do not offset one another. Instead, our findings suggest a combined effect of blemishes and weight gain on the one hand, and glasses on the other hand.

Additional testing also showed that respondents attributed significantly higher scores to the ‘wish for additional information’ than to actual purchase intention. This suggests that a bridge might exist between consumer attitude and the intention to purchase a product. To our knowledge, however, no scale yet exists to enquire about this intermediate stage. Given the

ever increasing popularity of e-commerce (cf. 2.1.1.), more research into a new or modified purchase intention scale seems desirable (cf. section 7).

6. CONCLUSION

Reviews and their credibility are a widely studied topic, and due to the increasingly visual nature of today's social networking sites, there has been a growing interest in the profile pictures accompanying them. Such research seems especially worthwhile as numerous studies have underlined that reviews do not occur in isolation (Armstrong et al., 2009; Craciun et al., 2009; Lim et al., 2014; Lin et al., 2017; Shan, 2016; Yoo et al., 2007), and a picture of the source has been shown to be one of the factors influencing consumers along with the textual information (Ozanne et al., 2019; Xu, 2014).

Nevertheless, the literature still contains a number of gaps. Certain media, for example, have not yet been covered. Instagram, as such, is a medium whose use is widely prevalent today, and which is increasingly used to share opinions on products and services on. Despite the growing body of literature on Instagram and its 'influencers', however, reviews on the platform have not received ample attention, and there is still considerable ambiguity with regard to how influencers exactly impact purchase intention (Balaban et al., 2019; Djafarova et al., 2017; Lee et al., 2020; Lim et al., 2017; Rebelo, 2017). Extending eWOM research to the platform of Instagram also seems particularly valuable as many of its reviews are sponsored, possibly raising caution and generating results inconsistent with previous research.

In addition, many features altering a source's appearance, such as glasses, have not yet been examined in a review context. Glasses, as an indicator of intelligence, have been shown to increase perceived credibility in job applicants (Fetscherin et al., 2020; van der Land et al., 2014), but it has not been verified whether these results also apply to reviewers.

In an attempt to partly fill these gaps, this study examined how perceived intelligence and attractiveness affect message credibility, consumer attitude and purchase intention on the platform of Instagram. Throughout the study, it became clear that attractiveness as a concept may be problematic, and we decided to distinguish between physical and social attractiveness.

In this way, we were able to provide new insights into which physical features affect social attractiveness: both glasses and the combination of blemishes and weight gain seemed to positively influence the construct. Our results were also able to confirm that blemishes and

weight gain diminish physical attractiveness (Chithambo et al., 2013; Fink et al., 2001), but (thin-rimmed) glasses do not (Leder et al., 2011).

Our actual study, then, centred around two main questions, the first asking **how physical features (blemishes and weight gain, glasses) affect the perceived online credibility of influencers' sponsored reviews on Instagram**. Our findings suggest that although blemishes and weight gain diminish physical attractiveness, they do not lead to a decrease in message credibility. Instead, they enhance social attractiveness, which does positively influence message credibility. To our knowledge, no research has yet been conducted studying the halo effect of social attractiveness. The findings from this study, however, point towards the idea that, in this specific context, social attractiveness produces a stronger halo effect than physical attractiveness. Given these intriguing results, we hope that this study is a springboard for more research into the effects of social attractiveness.

Furthermore, glasses, as indicators of perceived intelligence, improved message credibility as well, but only when interaction effects of participants' gender, age, ethnicity and education were not taken into account. Although none of the interaction effects was significant, respondents' education level seemed to affect the impact of glasses the most. However, the graph depicting the interaction effect of education (figure 17) revealed that its influence is likely due to the very limited number of participants with primary education (3), who happened to attribute extremely high message credibility scores to the scenarios with, and extremely low scores to the scenarios without glasses. However, message credibility does increase with perceived intelligence, which is significantly higher for influencers wearing glasses.

To answer our second research question, we examined **how the perceived credibility of the influencer affects purchase intention.** Firstly, both perceived intelligence and social attractiveness were shown to positively impact consumer attitude and purchase intention. Whereas the effect from social attractiveness was fully mediated by consumer attitude, the effect from perceived intelligence was only partially mediated, revealing a direct effect on purchase intention as well. Physical attractiveness, however, did not affect consumer attitude, nor purchase intention.

Finally, several other findings emerged which did not immediately relate to our research questions. First, virtually no influence of homophily (gender, age, glasses) was found on

perceived intelligence, message credibility, social or physical attractiveness. Female respondents did rate the female influencers higher in social attractiveness than the male respondents, but due to the design of our study it is not clear whether this can be attributed to homophily. That is, there were no male stimuli to compare scores with. Secondly, our additional scale measuring participants' wish for additional information yielded significantly higher scores than the traditional scale measuring purchase intention, possibly signalling the need for a renewed scale adapted to today's society, in which e-commerce occupies a salient position (cf. section 7). Finally, our results revealed a combined effect of glasses and the combination of blemishes and weight on social attractiveness: the influencer with both types of manipulations (FLAG) was rated significantly more socially attractive than the influencers with either just glasses (FAG) or blemishes and weight gain (FLANG). The physically attractive influencer without glasses (FANG) was attributed the lowest scores for social attractiveness. The stimuli with only one manipulation did not significantly differ in social attractiveness, suggesting glasses on the one hand, and blemishes and weight gain on the other hand carry approximatively the same weight.

7. LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The experimental nature of our study entailed several limitations. This section therefore provides an overview of the shortcomings which should be taken into account when interpreting the results, and which may be useful should future studies want to examine this particular topic in greater depth. To this aim, a number of suggestions for future research are included as well.

The first limitation of our study is due to the sample size: of the 287 valid respondents, only 106 (36.9%) were male. Female participants thus were overrepresented. In addition, 97.1% of all respondents were of Western ethnicity, and 79.8% had pursued higher education. Also Gen Y (27-41) (15.3%) and Gen B (58-76) (10.1%) were underrepresented. It would be interesting to conduct a similar experiment with a more balanced dataset, as different cultures may have different perceptions of social and physical attractiveness (see, for example, Swami, Caprario, Tovée & Furnham, 2006), and people in different age categories or with divergent education levels may have a different approach to social media. Respondents with primary education (3), for example, seemed to diverge in their answers from groups with another educational background (cf. 4.3.1.1.), but it is impossible to tell whether this would have also been the case with a larger sample size.

Secondly, attractiveness remains a subjective judgement. Whether certain characteristics of the source are considered attractive could thus very well be ascribed to personal taste or even present-day fashion trends (e.g. hair colour, attire, beardedness and glasses, etc.). However, to avoid raising suspicions about the purpose of the study, and to restrict the number of questions in the survey, we did not enquire about respondents' view on attractiveness, or whether they considered themselves to be up-to-date with the latest trends. For this same reason we did not include detailed questions on homophily, such as how the respondents would rate themselves in attractiveness, whether they enjoyed podcasts, etc. The elimination of the influencers' personal details and exclusion of their profile information (likes, biography, name, followers) to avoid bias also rendered comparisons related to homophily impossible. Due to our study's design, similarity to the influencer was thus only studied peripherally, which might explain why virtually no influence was found. Given the apparent positive influence of homophily on message persuasiveness, source credibility and consumer attitude (Ismagilova et al., 2020), however, it seems conducive to have a clearer picture of the role homophily plays in influencer

marketing, especially as influencers generally communicate to a niche audience (Lim et al., 2017, p. 20). One could, for instance, conduct a within-subject experiment, in which participants have to express their opinion on the credibility of multiple influencers, belonging to different age categories, with different backgrounds, etc., but also provide more information about their own appearance, preferences, etc. afterwards.

Thirdly, respondents were shown one isolated review of an influencer they could not know. This artificial setting may have triggered a more critical analysis than normally would have been the case. In addition, respondents were unfamiliar with the influencer, because they were fictitious. This complicated the influencers' task to use their reputation to add value and positive associations to the product. Often, influencers are relevant in specific domains or within certain groups. Their credibility, then, could stem from their expertise with regard to the product, but also the sense of familiarity or identification their followers might feel. Our study's respondents, however, are not part of that niche and hence unlikely to be influenced to the same extent. Moreover, in this artificial setting, the review/advertisement is intrusive (Bóveda-Lambie et al., 2012), which it would not be for actual followers. Another complicating factor related to the review was the product promoted. Although we tried to opt for a gender-neutral item (cf. 3.2.), we were unable to anticipate the attitude of the respondents towards the product. It is therefore entirely possible that previous experiences with, or the recent purchase of a similar item influenced the results. For instance, respondent 261 stated: "is geen artikel dat *[sic]* ik naar op zoek ben, dus geen interesse in haar verhaal" [this is not an item which I am looking for, so I am not interested in what she has to say], whereas respondent 64 commented: "ik ben toevallig net aan het overwegen om oortjes aan te kopen. Dat heeft m'n antwoorden natuurlijk gestuurd" [as it happens, I am currently considering buying earphones. Of course, this influenced my answers].

Fourthly, because of the limited time span in which the study was conducted, and because of the minimal differences in physical attractiveness scores for the male influencers (cf. 3.3.2), our experiment was conducted using only female stimuli. It would, however, be interesting to replicate the study with (different) male stimuli, especially since actual Instagram influencers are predominantly female (84% in 2019) (Statista, 2021c). Male influencers might therefore generate dissimilar impressions. It might also be useful to consider blemishes and weight gain as two separate manipulations. In the present study, we only examined the combination of both features, but distinguishing between them could reveal whether both affect social attractiveness

to the same extent. A final limitation related to the actual design of the study concerns the order of the survey questions. Enquiring about the influencer's appearance and perceived intelligence before asking about review credibility and purchase intention would have likely raised suspicions about the purpose of the study. We therefore reversed this order, and first showed the review (accompanied by the influencer), followed by a picture of only the influencer (cf. 3.7.). However, and although we asked to judge the influencer solely based on their appearance, the sponsored review may have triggered bias among some respondents, possibly influencing scores for social attractiveness, for example. Future research conducting a similar experiment could distribute additional surveys enquiring solely about the influencer based on his/her appearance. Results could then be compared to examine the impact of the review.

Fifthly, we enquired about purchase intention not only through the traditional scale (cf. 3.8.). We feel that, in today's digitised society, where online shopping has become ingrained, the established questions on users' intention to buy no longer suffice. The significantly higher scores for our 'additional information' scale, as well as comments from participants, seem to corroborate this. Participant 121, for instance, stated: "Ik ga niet snel een product kopen van een merk/influencer dat ik niet ken. Ik zou eerst meer informatie opzoeken over het product en reviews van andere mensen lezen" [I do not easily buy products from a brand/influencer which/whom I do not know. I would first look up additional information about the product and read reviews from other people]. We believe an approach to purchase intention which includes questions about additional information would lend itself better to present-day advertising and marketing. We therefore hope our study can serve as a stimulus to consider fine-tuning and subsequently implementing a modernized version of the purchase intention scale. Statements such as "*I want to visit the brand's website", "I want to look up additional information online", "I want to read more reviews about this product"*, etc. might resonate more with respondents than the rather straightforward *"I am interested in this product"*, or *"I am inclined to consider buying this product"*. A scale combining both types of statements might therefore offer a more complete picture of (online) consumers' purchase intentions.

Further, simply enabling participants to leave a remark at the end of the survey led to many insights and a better understanding of what motivated some of their answers. It therefore seems interesting to give a more prominent place to open questions in follow-up research. Respondents could, for example, be asked to elaborate on why they chose particular answers. Open questions already play a key role in interview-based research, but such studies are often

time consuming and consequently limited in sample size. Combining the valuable insights of open questions with the potential reach of (online) questionnaires therefore seems to offer the best of both worlds. For instance, this approach could have revealed more about why Gen Y (Millennials) (27-41) attributed higher scores for perceived intelligence than Gen Z (18-26) (cf. 4.2.2.).

A final suggestion for future research concerns language errors in reviews and how they interact with the appearance of the source. Originally, we wanted to implement language errors as a variable in our study. However, as relatively few respondents noticed the errors in the pre-test, they were left out of the study. Still, the combination of language errors and source appearance seems to provide fertile soil for further research: not only can language errors harm the credibility of the source (Lybaert et al., 2020), the source can also create certain expectations language-wise. When these expectations are violated, this can negatively affect the source's credibility (see, for instance, Jensen, Averbeck, Zhang & Wright, 2013 for the language expectancy theory (LET) in product reviews). As glasses have been shown to be an indicator of perceived intelligence, this might generate the expectation of flawless writing. Future studies could examine whether bespectacled sources committing language errors are 'penalized' more severely (i.e. rated lower in credibility) than sources without glasses. This need, however, not be in an influencer context, as other settings could be more suitable. That is, language errors are probably more likely to occur in 'regular' reviews (TripAdvisor, Airbnb), reviews posted on business' Facebook pages, blog posts, etc. than in sponsored content.

BIBLIOGRAPHY

Adegbola, O., Gearhart, S., & Skarda-Mitchell, J. (2018). Using Instagram to engage with (potential) consumers: A study of Forbes Most Valuable Brands' use of Instagram. *The journal of social media in society, 7*(2), 232-251. Retrieved from https://www.thejsms.org

Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly, 93*(1), 59-79. doi: 10.1177/1077699015606057

Armstrong, C.L. & McAdams, M.J. (2009). Blogs of Information: How Gender Cues and Individual Motivations Influence Perceptions of Credibility. *Journal of Computer-Mediated Communication, 14*(3), 435-456. doi:10.1111/j.1083-6101.2009.01448.x

Balaban, D., & Mustățea, M. (2019). Users' perspective on the credibility of social media influencers in Romania and Germany. *Romanian Journal of Communication and Public Relations, 21*(1), 31-46. doi: 10.21018/rjcpr.2019.1.269

Batres, C. (2020). *PSA001 secondary analysis: Examining the "attractiveness halo effect"*. [Preprint]. PsyArXiv. doi:10.31234/osf.io/c7hf3

Bóveda-Lambie, A. M., & Hair, N. (2012). Advertising versus invertising: The influence of social media B2C efforts on consumer attitudes and brand relationships. In *Online Consumer Behavior* (pp. 243-270). Retrieved from https://www.researchgate.net

Broeder, P. & Remers, E. (2018). Eye contact and trust online: the effect of profile pictures on Airbnb booking. *The IEEE 12th International Conference on Application of Information and Communication Technologies,* 336-339. doi:10.1109/ICAICT.2018.8747121

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

Chadha, R. (2018, Jan 30). For Influencers, Instagram Is the Clear-Cut Favorite. [Article]. Retrieved December 19, 2021, from https://www.emarketer.com/

Cheung, C. M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. *International journal of electronic commerce, 13*(4), 9-38. doi:10.2753/JEC1086-4415130402

Cheung, C. M. Y., Sia, C. L., & Kuan, K. K. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *Journal of the Association for Information Systems, 13*(8), 2. doi:10.17705/1JAIS.00305

Chithambo, T. P., & Huey, S. J. (2013). Black/white differences in perceived weight and attractiveness among overweight women. *Journal of Obesity, 2013*. doi: 10.1155/2013/320326

Craciun, G. & Moore, K. (2019). Credibility of Negative Online Product Reviews: Reviewer Gender, Reputation and Emotion Effects. *Computers in Human Behavior, 97,* 104-115. doi:10.1016/j.chb.2019.03.010

Djafarova, E., & Bowes, T. (2021). 'Instagram made Me buy it': Generation Z impulse purchases in fashion industry. *Journal of Retailing and Consumer Services, 59,* 102345. doi: 10.1016/j.jretconser.2020.102345

Djafarova, E. & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior, 68,* 1-7. doi:10.1016/j.chb.2016.11.009

Dwivedi, Y. K., Kapoor, K. K., & Chen, H. (2015). Social media marketing and advertising. *The Marketing Review, 15*(3), 289-309. doi: 10.1362/146934715X14441363377999

Edwards, C., Stoll, B., Faculak, N., & Karman, S. (2015). Social presence on LinkedIn: Perceived credibility and interpersonal attractiveness based on user profile picture. *Online Journal of Communication and Media Technologies, 5*(4), 102-115. doi:10.29333/OJCMT/2528

Eisend, M. (2006). Two-sided advertising: A meta-analysis. *International Journal of Research in Marketing, 23*(2), 187-198. Retrieved from https://www.empgens.com

Expertise. (s.d.). In *Lexico.* Retrieved 09/02/2022 from https://www.lexico.com/definition/expertise

Fetscherin, M., Tantleff-Dunn, S., & Klumb, A. (2020). Effects of facial features and styling elements on perceptions of competence, warmth, and hireability of male professionals. *The Journal of social psychology, 160*(3), 332-345. doi: 10.1080/00224545.2019.1671300

Fink, B., Grammer, K., & Thornhill, R. (2001). Human (Homo sapiens) facial attractiveness in relation to skin texture and color. *Journal of Comparative Psychology, 115*(1), 92. doi: 10.1037//0735-7036.115.1.92

Fiske, S. T. (2018). Stereotype Content: Warmth and Competence Endure. *Current Directions in Psychological Science, 27*(2), 67–73. doi: 10.1177/0963721417738825

Flanagin, A. J., Winter, S., & Metzger, M. J. (2018). Making sense of credibility in complex information environments: The role of message sidedness, information source, and thinking styles in credibility evaluation online. *Information, Communication & Society, 23*(7), 1038-1056. doi:10.1080/1369118X.2018.1547411

Friman, J. (2010). *Consumer attitudes toward mobile advertising*. (Master's thesis). Retrieved from http://epub.lib.aalto.fi/

Ganti, A. (2022, March 06). Central Limit Theorem (CLT). [Article]. Retrieved March 28, 2022, from http://www.investopedia.com/

Gebauer, J. E., Leary, M. R., & Neberich, W. (2012). Unfortunate first names: Effects of name-based relational devaluation and interpersonal neglect. *Social Psychological and Personality Science, 3*(5), 590-596. doi: 10.1177/1948550611431644

Horns and halo effect. (2009). In *Oxford Reference online*. Retrieved 03/04/2021 from https://www.oxfordreference.com/view/10.1093/oi/authority.20110810105106599

Instagram Business Blog (2016, May 31). Coming soon: New Instagram Business Tools. [Blog]. Retrieved November 16, 2021 from https://business.instagram.com/

Instagram Business Blog (2017, November 30). Celebrating a community of 25 million businesses. [Blog]. Retrieved November 16, 2021 from https://business.instagram.com/

Instagram Business Blog (2019, March 19). New to Instagram shopping: Checkout. [Blog]. Retrieved November 16, 2021 from https://business.instagram.com/

Ismagilova, E., Slade, E., Rana, N. P., & Dwivedi, Y. K. (2020). The effect of characteristics of source credibility on consumer behaviour: A meta-analysis. *Journal of Retailing and Consumer Services, 53*, 101736. doi: 10.1016/j.jretconser.2019.01.005

Jacobson, J., Gruzd, A., & Hernández-García, Á. (2020). Social media marketing: Who is watching the watchers? *Journal of Retailing and Consumer Services, 53*, 101774. doi: 10.1016/j.jretconser.2019.03.001

Jensen, M. L., Averbeck, J. M., Zhang, Z., & Wright, K. B. (2013). Credibility of anonymous online product reviews: A language expectancy perspective. *Journal of Management Information Systems, 30*(1), 293-324. doi:10.2753/MIS0742-1222300109

Johnston, E., Jackson, S., & Kelley, M. (2009). Dress For the Job You Want: An Examination of the Relationship Between Work Attire and Perceived Level of Education. *Intuition: BYU Undergraduate Journal of Psychology, 5*(1), 15-20. Retrieved from https://scholarsarchive.byu.edu/

Kibo. (2020, Dec 14). Younger Generations Show Increased Shift to Long-Term Online Shopping. [Article]. Retrieved November 2, 2021, from https://www.businesswire.com

Krämer, N. C., & Winter, S. (2008). Impression management 2.0: the relationship of self-esteem, extraversion, self-efficacy, and selfpresentation within social networking sites. *Journal of Media Psychology, 20*(3), 106–116. doi:10.1027/1864-1105.20.3.106

Kruger, L-M., Kühn, S.W., Petzer, D.J. & Mostert, P.G. (2013). Investigating brand romance, brand attitude and brand loyalty in the cellphone industry. *Acta Commercii 13*(1), 178. doi: 10.4102/ac.v13i1.178

Leder, H., Forster, M., & Gerger, G. (2011). The Glasses Stereotype Revisited. *Swiss Journal of Psychology, 70*(4), 211-222. doi: 10.1024/1421-0185/a000059

Lee, D., Trail, G. T., Lee, C., & Schoenstedt, L. J. (2013). Exploring Factors that Affect Purchase Intention of Athletic Team Merchandise. *ICHPER-SD Journal of Research, 8*(1), 40-48. Retrieved from https://eric.ed.gov/

Lee, K. T. & Koo, D. M. (2012). Effects of attribute and valence of e-WOM on message adoption: Moderating roles of subjective knowledge and regulatory focus. Computers in Human Behavior, 28(5), 1974-1984. doi:10.1016/j.chb.2012.05.018

Lee, S., & Kim, E. (2020) Influencer marketing on Instagram: How sponsorship disclosure, influencer credibility, and brand credibility impact the effectiveness of Instagram promotional post. *Journal of Global Fashion Marketing, 11*(3), 232-249. doi: 10.1080/20932685.2020.1752766

Li, M. Q., Lee, Y. C., & Wu, L. L. (2020). Sidedness Effect of Online Review. *Proceedings of the 24th Pacific Asia Conference on Information Systems: Information Systems (IS) for the Future, PACIS 2020*, 125. Retrieved from https://aisel.aisnet.org/

Lim, X. J., Radzol, A. M., Cheah, J., & Wong, M. W. (2017). The impact of social media influencers on purchase intention and the mediation effect of customer attitude. *Asian Journal of Business Research, 7*(2), 19-36. doi: 10.14707/ajbr.170035

Lim, Y. S., & Van Der Heide, B. (2014). Evaluating the wisdom of strangers: The perceived credibility of online consumer reviews on Yelp. *Journal of Computer-Mediated Communication, 20*(1), 67-82. doi:10.1111/jcc4.12093

Lin, C. A. & Xu, X. (2017). Effectiveness of online consumer reviews. The influence of valence, reviewer ethnicity, social distance and source trustworthiness. *Internet Research, 27*(2), 362-380. doi:10.1108/IntR-01-2016-0017

Long-Crowell, E. (10.05.2013). The Halo Effect: Definition, advantages & disadvantages. [Ch.7, lesson 4]. Retrieved March 1, 2021 from https://study.com/academy/

Lybaert, C., De Clerck, B., Plevoets, K., Loete, T., & Depovere, S. (2020). De impact van taal en gender in negatieve reviews: Een experimentele studie naar gepercipieerde geloofwaardigheid en consumentengedrag. *Tijdschrift voor Taalbeheersing, 42*(1), 55-94. doi:10.5117/tvt2020.1.003.lyba

Maesschalck, V. (2015). *'Wat een schitterent hotel!' Een experimenteel onderzoek naar de impact van spelfouten in online reviews.* (Master's thesis). Retrieved from https://libstore.ugent.be

McCroskey, J.C. & McCain, T.A. (1974) The measurement of interpersonal attraction. *Speech Monographs, 41*(3), 261-266. doi: 10.1080/03637757409375845

Melton, J. (2019, March 26). Millennials now do 60% of their shopping online. [Article]. Retrieved November 2, 2021, from https://www.digitalcommerce360.com

Mena, P., Barbe, D., & Chan-Olmsted, S. (2020). Misinformation on Instagram: The impact of trusted endorsements on message credibility. *Social Media + Society, 6*(2), 1-9. doi: 10.1177/2056305120935102

Mohsin, M. (2021, Feb 16). 10 Instagram stats every marketer should know in 2021. [Infographic]. Retrieved October 30, 2021, from https://www.oberlo.com

Naderer, B., Matthes, J., & Schäfer, S. (2021). Effects of disclosing ads on Instagram: the moderating impact of similarity to the influencer. *International Journal of Advertising, 40*(5), 686-707. doi: 10.1080/02650487.2021.1930939

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: evidence for unconscious alteration of judgments. *Journal of personality and social psychology, 35*(4), 250-256. doi:10.1037/0022-3514.35.4.250

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39–52. doi:10.1080/00913367.1990.10673191

Ozanne, M., Liu, S.Q., & Mattila, A.S. (2019). Are attractive reviewers more persuasive? Examining the role of physical attractiveness in online reviews. *Journal of Consumer Marketing, 36*(6), 728-739. doi:10.1108/JCM-02-2017-2096

Peng, L., Cui, G., Chung, Y., & Zheng, W. (2020). The Faces of Success: Beauty and Ugliness Premiums in e-Commerce Platforms. *Journal of Marketing, 84*, 67 - 85. doi:10.1177/0022242920914861

Petty, R. E., & Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123-205. doi:10.1016/S0065-2601(08)60214-2

Rao, A. R., Qu, L., & Ruekert, R. W. (1999). "Signaling Unobservable Product Quality Through a Brand Ally," Journal of Marketing Research, *36*(May), 258–268. doi: 10.1177/002224379903600209

Rebelo, M. F. (2017). *How Influencers' Credibility on Instagram is perceived by consumers and its impact on purchase intention*. (Doctoral dissertation). Retrieved from https://repositorio.ucp.pt/

Remue, L. (2021). *Visual cues and their impact on online credibility: a contextualised approach*. Unpublished Bachelor paper. Faculty of Arts and Philosophy, Ghent University.

Schindler, R. M. & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behaviour, 11*(3), 234-243. doi:10.1002/cb.1372

Shan, Y. (2016). How credible are online product reviews? The effects of self-generated and system-generated cues on source credibility evaluation. *Computers in Human Behavior, 55*, 633-641. doi:10.1016/j.chb.2015.10.013

Sharma, G. (s.d.). Why do Brands Prefer Instagram over Other Social Media Platforms? [Article]. Retrieved November 1, 2021, from https://www.socialpilot.co

Statista. (2021a, Sep 7). Instagram: distribution of global audiences 2021, by age group. [Statistics]. Retrieved October 30, 2021, from https://www.statista.com

Statista. (2021b, Sep 10). Global social networks ranked by number of users 2021. [Statistics]. Retrieved October 30, 2021, from https://www.statista.com

Statista. (2021c, Aug 13). Distribution of influencers creating sponsored posts on Instagram worldwide in 2019, by gender. [Statistics]. Retrieved May 07, 2022, from https://www.statista.com

Swami, V., Caprario, C., Tovée, M. J., & Furnham, A. (2006). Female physical attractiveness in Britain and Japan: a cross-cultural study. *European Journal of Personality:*

Published for the European Association of Personality Psychology, 20(1), 69-81. doi: 10.1002/per.568

Tafesse, W., & Wood, B. P. (2021). Followers' engagement with instagram influencers: The role of influencers' content and engagement strategy. *Journal of Retailing and Consumer Services, 58,* 102303. doi: 10.1016/j.jretconser.2020.102303

Trustworthiness. (s.d.). In *Lexico.* Retrieved 08/02/2022 from https://www.lexico.com/definition/trustworthiness

Takeda, M. B., Helms, M. M., & Romanova, N. (2006). Hair color stereotyping and CEO selection in the United Kingdom. *Journal of human behavior in the social environment, 13*(3), 85-99. doi:10.1300/J137v13n03_06

van der Land, S. F. & Muntinga, D. G. (2014). To shave or not to shave? How beardedness in a Linkedin profile picture influences perceived expertise and job interview prospects. *HCI in Business: Lecture Notes in Computer Science, 8527*, 257–265. doi:10.1007/978-3-319-07293-7_25.

van der Land, S. F., Willemsen, L. M., & Unkel, S. A. J. (2015). Are Spectacles the Female Equivalent of Beards for Men? How Wearing Spectacles in a LinkedIn Profile Picture Influences Impressions of Perceived Credibility and Job Interview Likelihood. *HCI in Business: Lecture Notes in Computer Science, 9191*, 175–184. doi:10.1007/978-3-319-20895-4_17

Walther, J. B. (1993). Impression development in computer-mediated interaction. *Western Journal of Communication, 57*(4), 381-398. doi:10.1080/10570319309374463

Wang, X. (2016). *The Impact of Profile Picture Facial Expressions on Review Credibility*. (Master's thesis). Retrieved from https://etd.ohiolink.edu/

Wei, X. & Stillwell, D. (2017). How smart does your profile image look? Estimating intelligence from social network profile images. *In Proceedings of the tenth ACM international conference on web search and data mining*, 33-40. doi:10.1145/3018661.3018663

Wilton, B. (2015). *Professional personal branding: An Investigation of the Process How Recruiters Judge the Perceived Credibility of Job Applicants Based on Their LinkedIn Profile Picture.* (Master's thesis). Retrieved from https://thesis.eur.nl/

Xu, Q. (2014). Should I trust him? The effects of reviewer profile characteristics on eWOM credibility. *Computers in Human Behavior, 33*, 136-144. doi:10.1016/j.chb.2014.01.027

Yoo, K. H., Lee, K. S., & Gretzel, U. (2007). The role of source characteristics in eWOM: What makes online travel reviewers credible and likeable? In M. Sigala, L. Mich, J. Murphy, and A. Frew (Eds.), *Proceedings of the 14th International ENTER Conference in Ljubljana, Slovenia, 2007* (pp. 23-34). UK, Axon Imprint.

APPENDIX

Appendix 1: questionnaire

DUTCH	ENGLISH	Adopted or adapted from
Message Credibility		
De review is nauwkeurig.	This review is accurate.	Appelman & Sundar (2016)
De review is authentiek.	This review is authentic.	
De review is geloofwaardig.	This review is believable.	
De review is informatief.	This review is informative.	Friman (2010)
De review is nuttig.	This review is useful.	
De review is interessant.	This review is interesting.	
Product Attitude		
Het product dat gepromoot wordt, lijkt me een goede geluidskwaliteit te hebben.	The product that is promoted seems to have a good sound quality.	Rao et al. (1999)
Het product dat gepromoot wordt, lijkt me degelijk.	The product that is promoted seems decent.	
Het product dat gepromoot wordt, lijkt me kwaliteitsvol.	The product that is promoted seems qualitative.	
Brand Attitude		
Dit lijkt me een merk dat aan mijn verwachtingen voldoet.	This seems like a brand that will meet my expectations.	Kruger, Kühn, Petzer, & Mostert (2013)
Dit lijkt me een merk dat me niet zal teleurstellen.	This seems like a brand that will not disappoint me.	
Dit merk lijkt me betrouwbaar.	This seems like a reliable brand.	
Purchase Intention		
Ik heb interesse in dit product.	I am interested in this product.	Lee, Trail, Lee, & Schoenstedt (2013)
Ik heb interesse in dit merk.	I am interested in this brand.	
Ik ben geneigd om de aankoop van dit product te overwegen.	I am inclined to consider purchasing this product.	
Ik ben geneigd om de aankoop van dit merk te overwegen.	I am inclined to consider purchasing from this brand.	
Physical Attractiveness		
Deze persoon heeft een aantrekkelijk uiterlijk.	This person has an attractive appearance.	Ohanian (1990)
Deze persoon heeft een sexy voorkomen.	This person has a sexy appearance.	
Deze persoon heeft een mooi uiterlijk.	This person is beautiful.	
Deze persoon straalt klasse uit.	This person is classy.	
Social Attractiveness		
Deze persoon zou tot mijn vriendengroep kunnen behoren.	This person could fit into my circle of friends.	McCroskey & McCain (1974)
Het lijkt me makkelijk een gesprek met deze persoon te voeren.	It seems easy to have a chat with this person.	
Deze persoon ziet er vriendelijk uit.	This person looks friendly.	Fiske (2018)
Deze persoon heeft een warme persoonlijkheid.	This person has a warm personality.	
Deze persoon is betrouwbaar.	This person is trustworthy.	
Deze persoon is oprecht.	This person is sincere.	
Perceived intelligence & expertise		
Deze persoon ziet er intelligent uit.	This person looks intelligent.	Wei (2020)
Deze persoon ziet er deskundig uit.	This person looks knowledgeable.	Ohanian (1990)
Deze persoon ziet er ervaren uit.	This person looks experienced.	
Deze persoon heeft expertise.	This person looks like an expert.	

Appendix 2: pre-test 2

Originally, we had the intention to integrate language errors as a variable in our study: language errors in reviews have been shown to reduce perceived intelligence and credibility of respectively source and message (Lybaert et al., 2020). It therefore seemed particularly interesting to examine the interplay between attractiveness and glasses on the one hand, and mistakes on the other hand. If glasses increase perceived intelligence, for example, the question remains whether bespectacled reviewers would be 'penalized' more severely than reviewers without glasses because 'they are presumed to know such things', or, if they are attractive, whether the halo effect would exonerate them.

Pre-test 2, then, mainly served to check whether participants noticed the mistakes in the review: two rather obvious mistakes (verb conjugation) and two less noticeable mistakes. The number and type of errors were based on Maesschalk (2015). The review was only available in Dutch, but the mistakes used can be found below with an English equivalent for the two non-conjugation errors.

oplaadt → oplaad	Conjugation error
Gezegd → gezegt	Conjugation error
Interessant → intressant	Similar to interesting → intresting
Jouw → jou	Similar to your → you're

Ik ben een enorme fan van podcasts. True crime, debatten of lifestyle, ze interesseren me allemaal!

Daarom was ik ook superblij toen ik onlangs deze draadloze oortjes ontving van @lessmessaudio. Niet alleen zien ze er heel erg cool uit, ze zijn ook nog eens supercomfortabel en blijven heel goed zitten. Nu kan ik bijvoorbeeld het hele huis doorlopen terwijl mijn gsm oplaadt, en toch nog steeds genieten van mijn favoriete podcast 👌.

Had ik trouwens al gezegd dat de oortjes waterproof zijn, een batterijduur van meer dan 20 uur hebben én beschikbaar zijn in allerlei verschillende kleuren? Bovendien krijg je met de code maar liefst 15% korting bij je aankoop!

Wat zijn jullie favoriete podcasts? 🎧

#lessmessaudio #sponsored #podcasts

Ik ben een enorme fan van podcasts. True crime, debatten of lifestyle, ze intresseren me allemaal!

Daarom was ik ook superblij toen ik onlangs deze draadloze oortjes ontving van @lessmessaudio. Niet alleen zien ze er heel erg cool uit, ze zijn ook nog eens supercomfortabel en blijven heel goed zitten. Nu kan ik bijvoorbeeld het hele huis doorlopen terwijl mijn gsm oplaad, en toch nog steeds genieten van mijn favoriete podcast 👌.

Had ik trouwens al gezegt dat de oortjes waterproof zijn, een batterijduur van meer dan 20 uur hebben én beschikbaar zijn in allerlei verschillende kleuren? Bovendien krijg je met de code maar liefst 15% korting bij jou aankoop!

Wat zijn jullie favoriete podcasts? 🎧

#lessmessaudio #sponsored #podcasts

Sample

In total, 56 respondents completed the survey of pre-test 2, which only had one condition. Again, the participants were predominantly Western (96.4%), and most had pursued higher education (85.7%), but gender and age were more spread than in pre-test 1. Of the 56 respondents, 39 were female and 17 were male. The largest group was still aged between 18 and 26 (35.7%), but also 27 to 41 (23.2%) and 42 to 57-year-olds (28.6%) constituted a large share of the participants. 58 to 76-year-olds are least represented with 7 respondents (12.5%).

DEMOGRAPHIC	FREQUENCY	PERCENTAGE
GENDER		
Male	17	30.4
Female	39	69.6
AGE		
Gen Z (1996-2004) (18-26)	20	35.7
Gen Y (1981-1995) (27-41)	13	23.2
Gen X (1965-1980) (42-57)	16	28.6
Gen B (1946-1964) (58-76)	7	12.5
EDUCATION LEVEL		
Primary education	0	0
Secondary education	8	14.3
Bachelor	22	39.3
Master or PhD	26	46.4
ETHNICITY		
Western	54	96.4
Asian	1	1.8
Middle Eastern	1	1.8

Procedure and results

Similarly to pre-test 1, pre-test 2 consisted of a Dutch survey created with LimeSurvey which was distributed online. There was no renumeration for respondents. The first part of the questionnaire was composed of demographic questions, followed by a text instructing participants to read the Instagram review. The text also explained the absence of the influencer's name (privacy concerns). Participants were then asked to answer a list of questions measuring message credibility, brand attitude, product attitude and purchase intention through a 7-point Likert scale. The order of the questions was randomized to avoid patterned answers. Again, several control questions were included on factors which might influence the results: respondents were asked whether they were familiar with the brand in the review and whether they found the product gender neutral to make sure the product and brand name were well chosen. The next question asked whether they had noticed any language errors in the review. When respondents answered yes, they were requested to provide as many examples as possible.

In the final question, participants had to indicate the device they had used to complete the survey to ascertain that filling out the questionnaire on a laptop or phone did not lead to different results (because of the visuals). They were also able to leave any remarks they might have.

Calculating Cronbach's Alpha showed that all constructs measured were sufficiently reliable. High levels of internal consistency were found for message credibility (MC, α=.854), product attitude (PrA, α=.819), brand attitude (BrA, α=.821) and purchase intention (PI, α=.903). Also consumer attitude, composed of MC, PrA and BrA, was internally consistent (CA, α=.843).

Construct	Indicator	Cronbach's Alpha
Message credibility	MC1 MC2 MC3 MC4 MC5 MC6	.854
Product attitude	PrA1 PrA2 PrA3	.819
Brand attitude	BrA1 BrA2 BrA3	.821
Purchase intention	PuI1 PuI2 PuI3 PuI4	.903
Consumer attitude	MC PrA BrA	.843

The results showed that the product and brand name were both appropriate for the experiment: 94.6% of the respondents reported that they thought the product was gender neutral, and 94.6% indicated they were not familiar with the brand. From the 56 respondents, only 31 (55.4%) noticed one or more language errors. Although all types of errors were spotted, the conjugation errors were reported notably more. From the 31 respondents who found the errors, 28 found at least one conjugation error, while the 'jou' error was reported 13 times, and the 'intressant' error only 6 times. It is possible that participants did notice the other errors and were simply not able to remember them while answering the final questions, but still, this suggests that the different types of mistakes do not carry the same weight.