Aalto-yliopisto
Aalto-universitetet
Aalto University

Master's Programme in Business Analytics

TikTok Influencer Ad Disclosure Types on Consumer Trust

How Different Sponsorship and Commercial Disclosure Styles Impact Engagement

Victoria Schumacher

Master's thesis
2025

Aalto-yliopisto
Aalto-universitetet
Aalto University

Author Victoria Schumacher

Title of thesis TikTok Influencer Ad Disclosure Types on Consumer Trust

Programme Master of Science in Business Analytics

Major Business Analytics

Thesis supervisor Prof. Ilkka Leppänen

Date 24.11.2025 **Number of pages** 62 **Language** English

Abstract

This thesis examines how TikTok influencer sponsorship disclosures affect consumer trust, using engagement as a proxy. Drawing from existing works like the Persuasion Knowledge Model, Parasocial Interaction Theory, and the Authenticity Management Framework, this study analyses 2,250 TikTok posts across 45 beauty influencers with both WEKA machine learning models and JASP regressions. The results affirm that visible sponsorship does have a trust penalty, indicated by a drop in engagement, on TikTok posts. In terms of commercial cues beyond traditional sponsorship, it could not be concluded that explicit disclosure of free PR products impacted engagement, ambiguous disclosure of free PR products reduced engagement, and affiliate codes increased engagement. This suggests that economic incentives can boost perceived value to viewers and drive interest rather than diminish authenticity and trust. Regarding different clarity tiers for different disclosure types (low, medium, high), there is a directional trend toward greater engagement penalties, but statistical evidence remains inconclusive due to multicollinearity within the disclosure tiers. Additionally, although it is shown that a positive general tone increases engagement overall for beauty influencers on TikTok, it does not moderate the engagement penalty of sponsorship. Thus, the results indicate that a generally positive product review tone across an influencer's content and sponsorship effects work additively rather than interactively. Overall, these findings expand previous social media influencer ad research into TikTok and highlight the importance of transparent and consistent disclosure practices, clarify how different commercial cues impact how audiences respond to influencer marketing, and provide guidance for influencers, brands, and platforms navigating the ethical and regulatory expectations associated with social media advertising.

Keywords TikTok, Influencer Marketing, Sponsorship Disclosure, Authenticity, Parasocial Interaction Theory, Persuasion Knowledge Model, PR Products, Affiliate Marketing, Social Media Marketing, Beauty Influencers, Paid Partnerships, Influencers

Table of contents

Preface and acknowledgements

I want to thank my thesis advisor, Professor Ilkka Leppänen, for his guidance and advice as I explored this topic.

I'd also like to dedicate the completion of this thesis to my beloved family member who passed during the writing process, Ginger Schumacher. Additionally, a special dedication to Mixy Schumacher.

Otaniemi, 24 November 2025
Victoria Schumacher

Symbols and abbreviations

Symbols

B	Unstandardized coefficient
β	Standardized coefficient
R	Correlation coefficient
t	t-statistic
p	p-value

Abbreviations

PKM	Persuasion Knowledge Model
PIT	Parasocial Interaction Theory
AMF	Authenticity Management Framework
SE	Standard error
RMSE	Root mean square error
MAE	Mean absolute error

1 Introduction

1.1 Background and Motivation

Social media dominates a lot of modern life. It's how we exchange information with not only our inner circles but also the world as a whole, shaping how we view the world and ourselves; in fact, 40% of teenagers report that social media has increased their worries about their own body image (Krzymowski, 2024).

However, the social media landscape isn't as simple as it seems: it's not all thoughtless sharing of whatever content is entertaining, as brands have latched onto social media as an opportunity to sell and market products. With influencer marketing, brands can seamlessly slide exposure and recommendations for their products into consumers' feeds in a native way. Depending on the strategy employed — free PR products, affiliate marketing, paid ads disclosed one or more of a variety of ways, or other avenues — many individuals may not even realize that the content they are consuming contains an ad.

Particularly with the rise of TikTok and the ability to buy products directly from social media apps, incentivizing impulse purchasing, there's a need for guidelines to protect consumers and our broader society from misleading business practices related to influencer marketing. In addition to helping consumers, understanding how different ways of disclosing ads impacts viewer trust and engagement is helpful for platforms like TikTok to manage their content and for companies that want to employ influencers to effectively use influencer marketing as a tool. Overall, it's a win-win for everyone for there to be a clearer understanding of how influencer ads should be disclosed be it visually, verbally, or otherwise to create positive engagement and avoid suboptimal outcomes.

1.2 Objectives and Scope

Overall, the goal of this thesis is to create guidelines companies and influencers should follow for ad disclosure when participating in influencer marketing. Additionally, there's a desire to simply define and understand the influencer marketing landscape better. While the data used in this thesis will focus specifically on the beauty industry, which will be expanded upon shortly, it can serve as a base to extrapolate into other categories and future research. Not only will this goal allow companies to maximize the value of their marketing dollars and help influencers maximize their value to both their audiences and sponsors, it will also further discourse about marketing transparency. As research and regulation tend to lag technological evolution, understanding the impact of different communication strategies in

advertising — particularly in spaces where it might be confused for native content — can help inform regulation even as technology continues to evolve in new ways.

To achieve this goal this thesis will dissect the impact of different forms of ad disclosures on engagement, which will be used as a proxy for consumer trust. This leads to the main research question (RQ1). Additionally, a sub-research question (RQ2) will be investigated to explore whether the typical tone (positive vs. negative) influencers use to discuss products affects the impact of ads on engagement. These research questions are formalized below:

RQ1: How do influencer ad disclosure styles on TikTok affect follower trust?

RQ2: Does the typical tone of an influencer's reviews moderate an engagement drop caused by sponsorships?

Regarding the scope of this investigation, there are a few considerations. To provide a more narrow, comparable group of influencers, a single niche was selected: beauty influencers. With beauty being one of the most profitable influencer niches (Molenaar, 2025) and the main physical product category sold by US content creators (Yltava, 2023), it makes a good choice to streamline this thesis. Additionally, sponsored posts, affiliate links, free PR boxes, and other advertising content are extremely common to see in this niche, making this research relevant. Lastly, as beauty standards are directly related to self-image, focusing on the beauty industry directly tackles a key societal issue; per a 2019 study by the Mental Health Foundation in the UK, 22% of adults and 40% of teenagers say that social media content has led them to worry about their body image, feeding the roughly one third of people that feel anxious or depressed about their body image.

Another key piece of this thesis is its exclusive focus on TikTok as the target social media platform. While influencer marketing is a well-known field of research, Chinese-born TikTok is a newcomer in the social media landscape with relatively little research focused on it to date despite its rapid rise to fame; despite only emerging globally in 2017, TikTok reached 5.139 billion cumulative downloads by the end of 2024 (Tafradzhiyski, 2025). Coupled with a notable lack of research related to TikTok ads in comparison to other social media platforms' ads, TikTok's position as a recently emerged and growing social media platform makes it ideal for focus.

1.3 Structure of Thesis

To answer the research question, this thesis will take the following form. With the context now established in Section 1, Section 2 will begin by broadly

considering existing research that has been done in related areas of study. This includes investigating marketing persuasion models and existing ad studies on social media. Section 3 will describe this thesis's data collection strategies and methodologies, while Section 4 will analyze that data, Section 5 will discuss the findings in relation to the research questions and Section 6 will wrap up any last conclusions and a summary. In the end, disclosures and references will also be included.

2 Literature review

This literature review will focus on existing research in the field of influencer marketing and related topics.

To begin, it will start most broadly by understanding topics that are relevant to the psychology of influencer marketing. This will be important for setting a baseline for the whole field of influencer marketing and for understanding the psychology behind social media viewers' behavior. These frameworks will be directly applicable to understanding the results of this thesis's data analysis.

After, the literature review will drill down and discuss two relevant segments of influencer marketing: existing research related to influencer marketing in general and then more specifically research focused on TikTok. By beginning with the broader category of social media marketing before diving into TikTok specifically, this thesis will be able to tap into a wider body of research and focus on works that are relevant to the topic of disclosure. This is important with TikTok being a relatively new social media platform, as there is significantly less research related to it than, for example, Instagram. Consequently, while the research that does exist about TikTok is more directly applicable to this thesis and will be examined, there are fewer close comparisons to the research question.

2.1 Broader Applicable Topics to Influencer Marketing

Friestad and Wright's Persuasion Knowledge Model (PKM)

Diving into psychology relevant to influencer marketing and cited in several other papers referenced throughout this thesis, Friestad and Wright (1994) offer an important baseline with their PKM. Though Friestad and Wright's paper is from a few years before the era of social media and even further from the era of social media influencers, their model is still extremely relevant to this research; the PKM framework suggests that people are not passive participants in marketing persuasion. Rather, Friestad and Wright suggest that "consumers' persuasion-coping knowledge enables them to recognize, analyze, interpret, evaluate, and remember persuasion attempts".

Thus, with the PKM there's a suggestion that once an attempt to persuade is detected people may "refine their attitudes toward products and marketers' messaging" rather than take things at face value. While these actions are not strictly in opposition to the goals of the persuader (agent), the persuadee (target) will attempt to choose response tactics that allow them to "maintain control over the outcome(s)". This is done by drawing on three facets of knowledge:

1. Agent Knowledge: what the person knows about the persuader.
2. Topic Knowledge: what the person knows about the subject matter.
3. Persuasion Knowledge: what the person knows about persuasion tactics, motives, and strategies.

The figure below, pulled from Friestad and Wright's work, shows how these facets come together in the system of a persuasion attempt. As viewing the agent's side is useful for understanding the greater picture of persuasion but is not as relevant to this thesis, going forward the target's side of the system will be the focus.

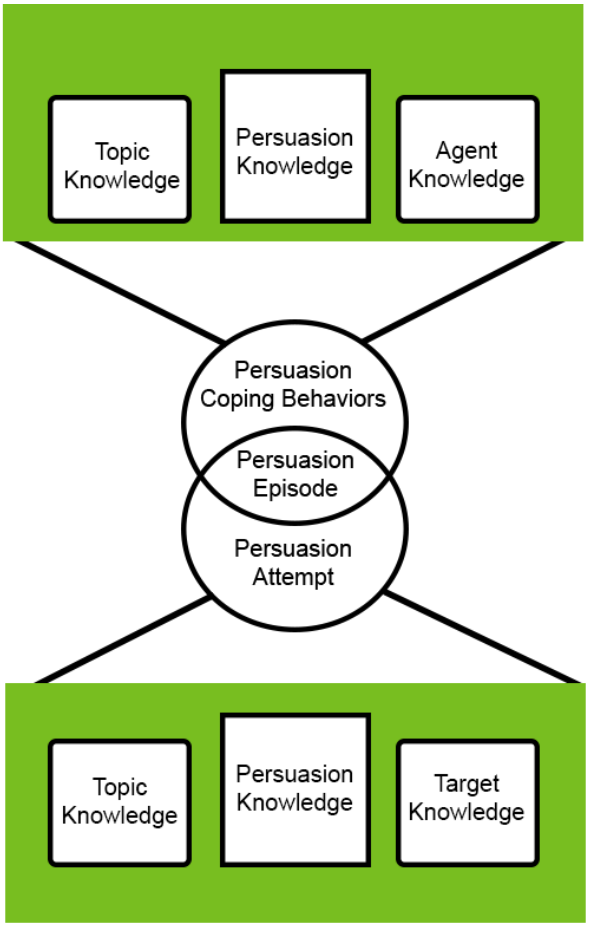

Figure 1. Friestad and Wright's PKM (1994), Graphic Recreation by Author

Agent knowledge as it relates to this thesis primarily comes from influencers. In the case where a viewer has never seen an influencer's content before, agent knowledge largely comes from how the influencer presents themself through their physical space, appearance, one-time content choices, editing, dialogue, and how they handle any relevant disclosures (if applicable). Without any deeper knowledge, most of what a viewer would get

from a video would be surface level. However, influencers inherently seek to build large followings, meaning the people seeing their content are likely to have some form of parasocial relationship with them rather than to have a fully removed opinion. In this case, they can form stronger opinions on agent knowledge based on personal experience with former product recommendations, a wider breadth of content choices, and a more well-rounded perception of how genuine the influencer seems in their content. Additionally, any scandals or positive press relating the influencer's personal or professional life are more likely to be known in this case.

A secondary source of agent knowledge comes from the companies sponsoring ads. As companies are the ones making beauty products and the claims about them, viewers of influencer advertising are likely to consider previous experiences with that company's products when judging them as an agent. Just like with influencers, companies are also not free from the potential impacts of a scandal or positive press. As a final note, companies' motivations are very inherent; regardless of whether beauty companies may want to help their customers with specific issues, ultimately their goal boils down to selling products. While influencer content can put a bit of distance between a product and that goal by adding distance between consumers and a brand's direct messaging, the producing companies have little ability to change the perception that they are primarily trying to push sales with their own messaging.

Topic knowledge with these research questions relates to how much someone knows about the beauty industry and possibly even more narrowly what someone knows about a particular category of product. For example, a person with dry skin may know very little about products meant for oily skin and someone who doesn't use eyeliner on a regular basis won't feel like an expert in that product category despite being a regular makeup user. Depending on their knowledge, TikTok viewers will have different abilities to tell whether an influencer's claims seem reasonable and may find a persuasion attempt more or less convincing as a result.

Looking lastly at persuasion knowledge, this relates to how well a viewer understands how influencer marketing and beauty marketing in general work. Relating to influencer marketing specifically, this knowledge could start as simply as answering the question, "does this feel like an ad?" Diving deeper, while most people probably understand that a labelled ad is a paid-for exchange between an influencer and a company there are many sneakier forms of brand involvement. To provide one example, it is commonplace in the beauty industry for companies to send out free PR packages to influencers to promote new products. Despite influencers having an obvious incentive in those cases to say good things about the brands sending them things they often seemingly perceive that they are not required to disclose anything when discussing free PR products due to a lack of formal contract or exchange of money. In another example, posting affiliate codes that provide product

discounts or sending Amazon product links directly to commenters are often disguised as generosity when in reality influencers are gaining a commission off sales through their code or link. Looking more broadly, beauty marketing has plenty of access to photoshop or filters to manipulate product effects and in some cases their products directly feed off insecurities – companies will sell products to solve problems people did not even know they had!

Per Figure 1, these facets of knowledge come together to inform a target's, which in this case would be a social media viewer, persuasion coping behaviors. However, noting one other remarkable piece from this model, Friestad and Wright discuss how persuasion tactics are only recognized as such when someone sees how it might affect them psychologically. In the case of beauty influencers, this might mean that a viewer acknowledges that a brand is trying to borrow something from the influencer by partnering with them. In many cases, this is likely viewer trust and credibility to give credence to a product(s) as it is a given that brands will only hype their products rather than be critical in the vast majority of cases.

Eisend's Source Credibility Analysis

Friestad and Wright's work leads directly to a question addressed by a 2006 work done by Eisend: why would a brand bother with influencers in the first place? After all, they could simply focus on other advertising techniques that keep their messaging more controlled and require less outsourcing. In his model, he defines three dimensions used by people to determine source credibility:

1. Inclination toward the truth (Will the source tell the truth?)
2. The potential of truth (Does the source know the truth?)
3. Presentation (Does the source appear to tell the truth?)

Going further, later on in his work Eisend diagrams a version of this source credibility framework that is specific to salespeople. While influencers are not traditional salespeople they are acting as such in the context of a particular sponsored post, or serving as product promoters in many non-sponsored posts. Consequently, Figure 2 shows a slightly more nuanced version of the above.

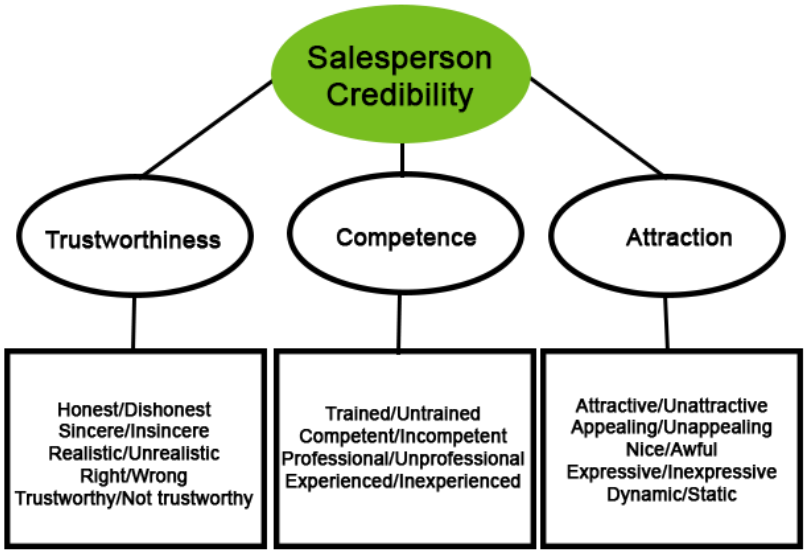

Figure 2. Eisend's Factor Structure of Salesperson Credibility (2006), Graphic Recreation by Author

As beauty influencers are likely to have very clearly built some amount of credibility on the second point it is less interesting to analyze than the other two bullets. This is due to clear experience in whatever space they are an actively posting influencer in and the fact that they are most often discussing products they are in the process of using. Meanwhile, the other two are a more nuanced discussion to understand.

Touching on presentation first, appearing to tell the truth can have multiple dimensions to it. Per the author, presentation contains but is not limited to "easily visible characteristics… like dynamism, attraction or attractiveness". Additionally, Eisend mentions that an exciting presentation as opposed to a boring display also drives credibly via the presentation dimension.

In the world of influencers, in conjunction with their unchangeable physical attributes like skin color, this means that presentation encompasses everything from the way an influencer dresses, speaks, edits their content, decorates their room, uses their makeup products, presents via their username, and much more. Regardless of whether their "vibe" is attempting to be a relatable, laid-back beauty friend or a clinical expert, Eisend's work suggests that businesses have good reason to focus on relationships with influencers whose brands and content they believe align with their own; the presentation of the influencer potentially rubs off on the product or brand being promoted and they reinforce each other. On the other hand, Eisend's model suggests that a strong deviation from an influencer's typical, trusted content could set off red flags in consumers regarding credibility from presentation, reinforcing the value of alignment.

Jumping back to Eisend's first point, inclination toward the truth is something hard to judge with influencers. While it would be fair to assume that most people are inclined toward the truth without other motivations, influencers try to build a personal brand, make content go viral, get positive attention and gifts from beauty brands, become famous, monetize content, and more. Per Freistad and Wright (1994) this might make an influencer's viewers feel that there are persuasion attempts occurring in their content (regardless of whether they are consuming specific paid ad by a company) and lead them to some level of distrust in the influencer.

This plays heavily into both the main research question and the secondary research question for this thesis. For the main question, when an influencer sees that something is an ad they are immediately going to be alerted that there is a persuasion attempt occurring and question whether the influencer they're watching will tell the truth with even more scrutiny. To them, the influencer getting paid is a clear motivator to say whatever a third party wants the influencer to say and that is likely to be a positive message. Consequently, overall this work suggests it is in the influencer's best interest to be careful about presenting ads in a native way that is consistent with their brand to avoid damaging credibility. Additionally, businesses are incentivized to select influencers that have high perceived inclination for the truth.

With that in mind as a segway, Eisend's third point also relates to secondary question for this thesis about whether the way an influencer typically discusses products affects the impact of ads on engagement. If an influencer's perceived inclination for the truth impacts their credibility one would assume that how they discuss products, regardless of payment, would impact that credibility. Coupling this with the higher scrutiny and distrust from the inclination toward the truth dimension when a piece of content is labelled as an ad and it is suggested that some impact should be seen during the analysis.

2.2 Generalized Influencer Advertising

The authenticity of content as a whole is important to consider in part because it helps to build the parasocial relationships that feed influencers' followings and engagement rates. As early as 1956 the core of Parasocial Interaction Theory (PIT) emerged, which describes the illusion of a personal and reciprocal relationship that people feel they have with a celebrity or other public figure (Horton & Wohl). In short bursts this can simply create one-sided parasocial "interactions" while longer term these "interactions" can develop into parasocial "relationships". In these cases, the person will feel like they have a strong, personal connection to the public figure while notably that public figure does address that specific person with their words or

actions. In short, they don't know a specific fan even exists while that same fan feels a strong bond.

With the emergence of new forms of media, such as social media, PIT has been further expanded upon and used in new applications. Research by Dandona et al. (2025), for example, explicitly references PIT and investigates how influencer credibility and perceived authenticity affect consumer trust. In this study, they show that consumer trust is "greatly influenced by perceived authenticity and influencer credibility, both of which are important predictors of brand engagement". Diving into this quote, influencers feeling authentic and credible, which again pulls from Eisend's 2006 framework, creates conditions for a parasocial relationship. This in turn leads to more engagement on social media. A notable piece from this paper along with other works (Barari et al., 2025; Pan et al., 2024) is that it provides a credible basis for using engagement as a proxy for consumer trust, which will be utilized as a core piece of this thesis during the analysis phase.

As a result of PIT and the other frameworks touched upon already, it can be said that for social media influencers authenticity management is a key to creating, maintaining, and growing their personal brand. Consequently, while advertising arrangements can be great for influencers monetarily they can also cause tensions that require influencers to manage with various strategies. This has led to research like Audrezet et al.'s 2018 work which interviewed social media influencers to come up with the Authenticity Management Framework (AMF) in Figure 3. The AMF is a matrix that looks at two authenticity strategies deliberately used by social media influencers to guide their brand partnerships.

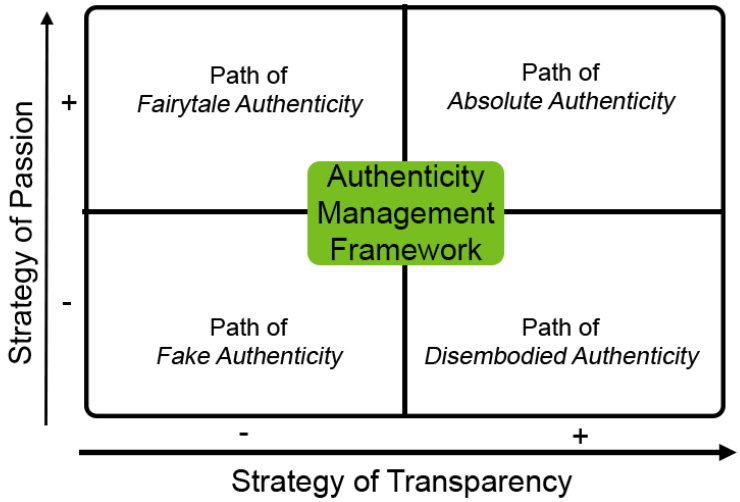

Figure 3. Audrezet et al.'s Four Quadrants of the AMF (2018), Graphic Recreation by Author

The vertical dimension of the AMF is strategy of passion, which relates to the degree to which the influencer's behavior is perceived as being driven by

intrinsic motivation, genuine enthusiasm, or personal values as opposed to extrinsic rewards like fame or monetary compensation. To fall high on this axis, ad content would need to be emotionally activating and clearly rooted in personal experience or interest. On the other hand, low passion ad content is going to appear commercially motivated or transactional. Overall, this dimension judges the authenticity of an influencer's intent with their post(s).

The horizontal dimension of the AMF is strategy of transparency, which relates to the degree to which an influencer reveals their sponsorships, advertising partnerships, or other commercial activities. To fall high on this axis advertising material would have to be almost impossible to miss and proactively disclosed. To be far on the other side of this axis, sponsored content would either mislead viewers or present a lot of ambiguity.

With this thesis the horizontal axis is most relevant to the main research question regarding the impact of different ad disclosure types on engagement as a proxy for user trust. Meanwhile, how that plays into the vertical axis begins to tie in the secondary question relating to how an influencer's usual way of discussing products changes the way an ad affects engagement; it's mentioned that "constant, consistent enthusiasm for every single brand mentioned could raise questions about whether [influencers] can be always very excited about each partnership and whether they are exaggerating their passion" (Audrezet et al, 2018).

In the upper left quadrant the authors discuss so-called "fairytale authenticity", described as content from influencers that genuinely enjoy producing content about a product(s) or brand(s) but don't disclose sponsorship ties. For example, one influencer discussed in the paper had a partnership with a shoe company called GOLA and would make posts discussing how their shoes are what she typically wears on a hot day to stay comfortable without ever mentioning their relationship. While the passion for the products is authentic, believing that passion and truth compensates for the lack of transparency leads to "fairytale authenticity".

In the lower left quadrant is "fake authenticity", is the strategy with the least authenticity overall. For influencers who participate in this strategy, they may find that their ad content sticks out unnaturally from their normal content due to requirements to comply with their partners' brand guidelines and/or a lack of interest on their part. As an example, the authors here mention an influencer they interviewed who promoted a Valentine's Day gift idea without disclosing their partnership when they themselves also question the whole institution of the holiday.

In the bottom right quadrant is "disembodied authenticity", which is when influencers are transparent about the fact that they're doing an ad but don't seem particularly enthused by whatever they're promoting. In short, the ad is not hidden from their viewers and the product they're promoting does not align well with their usual brand. An influencer from the study that fell into this category described a fashion week challenge from a company where she

had to come up with an outfit that fit certain requirements set by said company. While she accomplished the challenge and was clear about why she was doing it with her followers, she even later admitted that the sponsorship wasn't a particularly good fit for her passions, just her monetary interests.

In the upper right quadrant is "absolute authenticity", which is going to be the optimal point or holy grail of presenting authentically to an audience. In this case an influencer is going to be searching for sponsorship opportunities that are the ultimate win-win. With these engagements, the influencer gets to make money while feeling like they are connecting their audience with products and brands they genuinely believe in and think they should know.

Overall, this framework is crucial for understanding how social media influencers manage their personal brands and build a feeling of authenticity. As previously mentioned, this framework is particularly relevant when considering whether the way influencers discuss products in general affects the impact that ads have on engagement, as it would feed into their perceived authenticity. Even if they are not getting paid, this framework would still be somewhat relevant whenever a beauty influencer is giving a product review; they could be making content for the sake of it or because they're legitimately excited to explore a product or brand. They also could in some way have been influenced by the brand to try it through something like a PR package or alternatively have come about whatever they are reviewing organically.

More than just the framework, a key conclusion from Audrezet et al.'s study that is relevant to this thesis is that marketers should let influencers have a lot of flexibility in how they create an ad, what products they promote, and what opinions they express to maximize authenticity. However, despite this conclusion and Dandona et al. (2025) showing that high disclosure transparency improved the effect of authenticity on trust, there are still not proper, researched guidelines to how influencers should disclose their partnerships to reap those gains, which this thesis will expand upon.

2.3 TikTok-Specific Influencer Advertising

Much of the existing research related to TikTok does align with previous findings from other social media platforms.

Notably, it has been shown that perceived expertise and authenticity can build trust with an influencer, and that overt sponsorship can weaken authenticity (Tian et al., 2025). Tian et al. (2025) also offer the conclusion that disclosures can increase viewers' recognition of an advertising attempt, which aligns with the research done by Friestad and Wright (1994), but that their impact on deeper consumer outcomes, such as brand perception and purchasing behavior, have mixed conclusions and is still underexplored. While not directly the topic of this thesis, this is pointing to a similar space for further exploration.

Meanwhile, Dupuis et al. (2024) discusses relevant adjacent research focused on ad disclosures in the food and beverage space of TikTok. In their study, they found that some common ways of helping disclose partnerships, such as tagging a brand, did not meaningfully differentiate to adolescents (the focus of their study) between sponsored content and organic content. They conclude that in order to protect the public from undue influence by brands, clearer sponsorship labelling is needed. This ties closely into this thesis as it will investigate different forms of ad disclosure and make recommendations about guidelines for disclosures. Despite focusing on a different category of TikTok influencer, both food and makeup heavily feature branded products even when content isn't sponsored, with Tian et al. (2025) finding 47.6% of the videos in their dataset featured at least one.

2.4 Summary and Hypotheses

Overall, existing research suggests that authenticity and transparency are important for influencers who wish to maintain their brand and consumer trust, which can use engagement as a proxy. Additionally, when it comes specifically to sponsored content, social media viewers are likely to adjust how they process the content they view once they determine a persuasion attempt is happening. Consequently, despite there being minimal research around the relationship of sponsorship disclosures and consumer trust on the relatively new social media platform of TikTok, there is enough knowledge about human behavior and sponsorships on other social media platforms for hypotheses to be made.

For the main research question (RQ1), the hypothesis (H1) for this thesis centers around the learning that sponsorship reduces consumer trust and that persuasion attempts can impact how consumers react to messaging. Additionally, sponsored content can be less authentic and reduce credibility. The hypothesis is formally denoted and summarized below.

H1: More explicit ad disclosures will lead to higher engagement, suggesting higher viewer trust.

Regarding the second research question (RQ2), the hypothesis (H2) is centered around the learnings that influencer credibility and perceived passion impact how influencers are seen by their audience. Thus, it is presumed that influencers who are generally quite positive will seem more authentic when promoting an ad because ads would blend in better with their normal content. The hypothesis is formally denoted below.

H2: A more positive typical review tone reduces the engagement drop caused by sponsorships.

3 Research material and methods

The dataset used in this thesis was hand-collected by the author over the period July 17th, 2025 to October 4th, 2025. For each of 45 different influencers, 23 different attributes were collected initially. After adding calculated variables and preparing the data for analysis with WEKA and JASP, each TikTok had roughly 40 associated attributes.

The number of influencers and attributes collected were partially premeditated and partially determined during the data collection phase. Regarding the number of influencers, the goal was to balance statistical significance with the time needed to collect the data by hand. Tests were run at 35 collected influencers to see if any statistical significance could be found with the data and then more influencers were added to create a larger buffer. Regarding the initial attributes selected and the descriptions defining them, they were created after conducting the analysis found in section 3.2 Disclosure Analysis. That process considered how to keep track of the TikToks used so they could be referenced and determining what information would be needed to test the research questions.

Not all collected or calculated attributes were ultimately used in this thesis to draw conclusions due to both the focus the thesis and a lack of significance from some of the attributes. However, all attributes were left in the dataset for possible future use and analysis.

For each of the 45 influencers, data was collected for the most recent 50 posts on their profile excluding pinned posts. This was to provide a fair point of comparison across influencers and prevent pinned posts, which are essentially promoted posts, from skewing results. Additionally, by recording data for 50 TikToks per influencer there was enough data per influencer that the information was meaningful within its own bubble.

3.1 Influencer Selection

While numbers regarding demographics for beauty influencers are murky at best, 27% more white influencers have their first paid engagement within a year than black, indigenous, or people of color influencers (MSL Staff, 2023) and, while not strictly focusing on beauty influencers, 77% of monetized influencers are female (Thorpe, 2023). This lines up with the predominant suggestion that the “current leading face of beauty creators is white, female, and 24 to 38 years old” (Gant, 2025).

As a result, the majority of the influencers selected were young, white women. However, there was an intentional effort to also include a number of popular male influencers, older influencers, and influencers of color. The majority of these influencers were found by searching famous influencers in the beauty space on TikTok (such as @nikkietutorials and @jeffreestar),

keeping track of influencers tagged by other influencers as inspiration for certain posts, or through algorithm suggestions. For this thesis, ⅕ of the selected influencers identify as male and roughly ¼ are people of color.

In terms of shared characteristics for this group, there are two primary key points. First, the influencers all largely focus on beauty content that relates to makeup and facial skincare, as skincare directly affects makeup. This means beauty influencers who focus on fashion, nails, fragrances, or hair products were not considered to create a more homogenous group. As a second similarity, only influencers with 100,000 or more followers were considered. This was done to maximize the value of each influencer to the dataset, as it was assumed that influencers with larger numbers of followers would be more likely to have ad content. This assumption was made based on the idea that a brand would have more perceived value in collaborating with influencers who have larger audiences due to a higher presumed reach. Additionally, there is a higher chance that brands know about an influencer in the first place if they have more followers.

As a final requirement for influencers, their content had to be primarily done in English. This was so data collection could be accurately done by the English speaking data collector, but also has the added benefit of focusing this study on posts made in a dominate global language.

3.2 Disclosure Analysis

Prior to selecting the attributes that would be collected for each TikTok watched for the dataset, first there had to be an understanding of the different options for disclosing ads on TikTok. Disclosure options were ultimately broken down into four distinct categories: verbal, on-screen, description, and TikTok disclosure tags.

However, before diving deeper into those disclosure types it is worth discussing regulatory and platform requirements for disclosure to understand the broader context. As influencers can come from anywhere in the world the only fully unifying rules for how ads need to be disclosed can come from the shared platform, TikTok. For better or for worse, these rules are incredibly short and simple on TikTok, with their support page showing that the only real rule for posting videos that contain advertising is, “when posting content that promotes a brand, product, or service on TikTok, you must turn on the content disclosure setting”. This would mean that the TikTok disclosure tags category that will be discussed should always be turned on, though this thesis will discuss that further in the discussion section.

In terms of requirements from countries, many do have some guidelines for disclosing ads. In many places, including the USA, UK, and EU, these rules include some version of being required to disclose any meaningful relationship with a brand with simple, understandable verbiage in a

prominent way (European Commission, 2023; Federal Trade Commission, 2019, Advertising Standards Association, n.d.). Not only does this include any content made in exchange for monetary content, gifted items, and affiliates, it also includes other personal relationships with brands or the people involved. In some countries, such as the USA, there is even a requirement for how to disclose content: the Federal Trade Commission (2019) asks that influencers disclose ads both verbally and in text. While TikTok has the right to remove content that does not comply with their regulations, influencers who don't follow their local guidelines are at risk of being fined.

Verbal Disclosures

These disclosures stem from the auditory piece of a given video, meaning someone could be looking away from the video and still hear a disclosure. While influencers don't usually directly say "this is an ad", they may mention that the video is sponsored by or was a challenge from a certain brand. Alternatively, the influencer may simply mention that they are partnering with a brand for the video.

A key drawback to this disclosure style is that audio is not a required piece of watching TikTok videos. In fact, it is estimated that between 69-75% of videos are watched on mute, especially when people are in public places (McCue, 2019; Sharethrough, 2021). While some influencers include their own captions in the post-production of their videos and uploaded videos receive automatically generated captions, this still doesn't fully resolve the drawbacks of videos being watched on mute. Notably, this changes the disclosure cue from an auditory one back into a visual one that must be read, changing the disclosure type from a verbal disclosure to an visual disclosure.

Overall, this disclosure type is fairly central to the content itself. Assuming the video viewer does have their sound on and is engaging with the TikTok, this would be hard to miss unless the influencer were to use misleading phrasing. Additionally, when paired with other types of disclosure it can reinforce the point that a video is sponsored. However, since on its own it can be missed due to viewers not paying full attention constantly and many users watching videos silently, influencer- or auto-generated captions are a key partner. This will be expanded upon further in the next section while discussing captions.

On-Screen Disclosures

These disclosures are those that can be understood without audio and without looking at the description of the TikTok, meaning it appears visually in the video itself. This will be broken up into two categories, as there are two

common ways on-screen disclosures show themselves: partner disclosure and captions.

Most often, partner disclosure comes in the form of a very short hashtag or phrase such as #MaybellinePartner superimposed somewhere on the side of the video, though it could include something like a pop-up graphic with disclosure information.

While none of these are inherent to this style of disclosure, in practice there are a few notable drawbacks to this disclosure style. One common drawback is that these partner disclosures tend to be relatively subtle; they are often small text placed in a corner or some other place that is out of the way in terms of where a viewer would be inclined to look. Anecdotally, during data collection the author often would have missed this type of disclosure if it had not been standard practice to deliberately look for it. Another issue with them is a lack of uniformity, as there are no standards for where these disclosures should be or how they should look. For example, a bold, dark font of a more notable size could be in the bottom left corner on one TikTok while a tiny white disclosure could be placed underneath the TikTok navigation bar in another. Additionally, in a portion of the TikToks that use partner disclosure on-screen disclosures it's notable that the disclosure is only visible for the first few seconds of the TikTok. This was another challenging issue that required extra, deliberate attention during the data collection phase.

Overall, this type of disclosure would not generally be considered very upfront or visible in most cases. While it could theoretically be very transparent, the way it is used in actuality is not.

Captions are another, somewhat more interesting form of on-screen disclosure. In theory, they should perfectly transcribe what is being said by an influencer (or make a correction). As mentioned previously, this makes for a great reinforcement of any verbal disclosures and serves as a fail-safe if audio is off. However, when users use the fast forward (2X speed) feature on TikTok by holding down the edge of their screen, automatically generated captions are not visible. This means that for the sake of disclosure, influencer-added captions either done manually or through a paid service like TikTok's CapCut editing platform are superior to the automatically generated captions from TikTok uploads. Additionally, much like a verbal disclosure a caption requires a viewer to be paying attention at a particular moment in time to notice that the video is sponsored if used in isolation.

Overall, captions are a great way to help reinforce ad disclosures. However, because captions are so closely linked to verbal disclosures captions were not counted as a part of the On_Screen_Ad_Disclosure attribute in Section 3.3 Attribute Selection; they are not considered a distinct disclosure from the verbal disclosure because the captions are simply a supplement to the core verbal disclosure. This is especially salient as all TikToks for the sake of this thesis were watched with sound on to absorb the

full intended experience of the TikTok from the side of the influencer posting the video.

Description Disclosures

The description is the place where the bulk of additional information can be added and in many different ways, as they can be up to 4,000 characters long and include regular text, hashtags, and account tags. Despite the fairly large maximum description length, it is important to note that to see the full description on a TikTok video a viewer needs to tap on the description. Otherwise, only approximately two lines (50-100 characters) are visible. This means that the front pieces of a description are more likely to be viewed regardless of whatever is contained later in the text, as it requires no additional activation to view it.

Likely in an effort to comply with regulations that want a clear disclosure of partnerships, a common way influencers disclose their partnerships is by writing AD somewhere in their text, particularly in the beginning of their caption. Alternatively, when influencers state that a post is sponsored in their description they may tag the sponsoring brand and state that they are a partner in the post.

Hashtags are an interesting part of disclosures. Hashtags group content across TikTok posters into a shared place based on post content, meaning that users would associate hashtags with descriptor information. Consequently, a hashtag like #ad could be a good way for influencers to share that their post is an ad. However, while a sponsorship hashtag at the start of a description could be quite clear, hashtags are often placed at the end. In fact, it's quite common for influencers to place a large block of hashtags relevant to the content of their video at the end of their descriptions to help the TikTok algorithm spread their content. Unfortunately, this makes hashtag blocks less important to the people actually viewing the video as it describes content they have already seen and does not generally add to the experience of the TikTok. This is especially true as more hashtags continue to get added, making a singular #ad or similar get drowned out. Another note is that influencers disclosing an ad in their hashtags often don't use something generic like #ad – instead, they will use more specified hashtags like #MaybellinePartner or #SephoraSquad, which can be less clear than the generic terms.

Descriptions are also the most common place for affiliate codes to be found. Generally, they are placed within a list of products like an aside, stating something like "(CODE for 10% off)" or "(CODE to save $$)". Alternatively, influencers using affiliate links will state somewhere in their description that they will send links to a product(s) discussed to users who comment a certain phrase on their video or say that they have products linked in their bio.

While uncommon and not directly linked to the description, it is worth mentioning that some influencers do add additional context or information about a TikTok in the comments on occasion, which could include information about partners. As this is a rare occurrence in the first place and scrolling the comments is not an inherent experience when viewing a TikTok, any additional comments of this nature were not considered.

Overall, descriptions are a great place for influencers to include information. However, the degree to which that information can be considered featured depends both on how it is included and where it is included. If it is included as a hashtag for instance, it runs the risk of being drowned out by other hashtags or skipped over by viewers. Regarding placement, whatever information is placed at the top of the description is likely to have a bigger impact both due to it being more likely to be read even if a viewer stops reading part way through a description and because it is shown even without the description getting expanded.

TikTok Commercial Content Disclosure Tags

Commercial content disclosure tags are the only advertising tools exclusively used for advertising. For this research, there were two forms seen and one that is described in TikTok's support pages that was never seen during the data collection phase.

1. Paid Partnership: used for promoting or posting branded content on behalf of another business.
2. Promotional Content: used for influencers posting to promote their own brand or business. This is the commercial content disclosure label that was not seen during data collection.
3. Creator is Eligible for a Commission: for influencers who have both registered for and been approved for TikTok Shop's own affiliate program, allowing them to promote products and earn commissions directly through the TikTok platform.

Notably, per TikTok's support page, "turning on the content disclosure setting won't affect the distribution of your post in feeds on TikTok." This should mean that there is no incentive for creators to avoid turning on commercial content disclosure tags aside from the general effects of their content being an ad. This tag cannot be removed from a post later by the influencer and is a visible banner without additional actions on the part of a viewer.

Overall, the commercial content disclosure tags are an obvious way for influencers to disclose their content as an ad due to their unambiguous and exclusive nature as an advertising tool.

Disclosure Type Tiers

Based on analysis of the different forms of disclosure and the data collection experience, disclosure attributes have been broken into three tiers. While the attributes will simply be listed below and broken into categories, the details of these attributes will be expanded upon in Section 3.3 Attribute Selection.

1. High clarity disclosure methods. These are disclosure methods that are immediately apparent for the duration of a video from the very beginning. They require no additional actions on the part of the TikTok viewer to see.
    a. Sponsor_Hashtag_Start
    b. Paid_Partnership_Disclosure
2. Medium clarity disclosure methods. These are methods that are directly stated but are only present for part of a TikTok or require additional steps by the viewer to determine, such as clicking on the description.
    a. Stated_Ad
    b. Written_Stated_Ad
3. Low clarity disclosure methods. These are a disclosure methods that are the most likely to be missed if they are the sole way of disclosing an ad due to being buried within the content in a way that's likely to be skipped over, such as in the background of a video or in the description.
    a. Sponsor_Hashtag_End
    b. On_Screen_Ad_Disclosure

When discussing the sponsorship tiers considered in this analysis it is worth noting the absence of both explicit PR (PR_Product) and implicit PR (PR_Product_Implied) as well as affiliate codes or links (Affiliate_Codes). While these attributes count as commercial content, they do not fall under traditional sponsorship criteria, which was the primary focus of this thesis prior to the data collection period.

3.3 Attribute Selection

One of the most crucial decisions for this thesis was the selection of which attributes would be recorded for each post during the data collection process. Overall, attributes for the spreadsheet were narrowed down into four categories: core post information, advertising information, content type information, and calculated information. The different information types are defined below with relevant attributes defined and discussed in the subsequent tabled:

1. Core post information: general data related to the post and poster.

Table 1. Core Post Information Attribute Definitions and Considerations

Attribute Name	**Value Type**	**Description and Considerations**
TikTok_Handle	String	Influencer's TikTok username
Post_Code		Self-created label to differentiate between posts and track the number of posts per influencer
Link	String	Direct link to the logged TikTok
Followers	Integer	Follower count as displayed on the app
Likes	Integer	Post's likes as shown on the app. Value sometimes shows slight shift on desktop view
Comments	Integer	Post's comments as shown on the app. Value sometimes shows slight shift on desktop view
Shares	Integer	Post's shares as shown on the app. Value sometimes shows slight shift on desktop view
Brands_Tagged	Integer	Number of brands tagged in the post caption. Does not include people, only brands. This attribute in some cases straddles the line of being advertising information.
Products_Review_Type	Enumerated	A subjective measure of how the influencer talks about products in the video. The default state is Not Applicable. (Not Applicable/Positive/Negative/Mixed/Both)
Pos_Review	Boolean	Indicates whether an influencer had exclusively positive comments about a product(s)
Neg_Review	Boolean	Indicates whether an influencer had exclusively positive comments about a product(s)
Mixed_Review	Boolean	Indicates whether an influencer had both positive and negative comments about a product(s)
Makeup Content	Boolean	Indicates whether the post features makeup and/or facial skincare. Subjective

Within this series of attributes several could benefit from further elaborations. In these cases, an understanding of the judgements being made and use cases are key.

Beginning with Products_Review_Type, this was the original way data was collected for Pos_Review, Neg_Review, and Mixed_Review before it was formatted for use in WEKA. As stated in the table the default state was Not

Applicable, which meant that the influencer did not appear to give an opinion on a clearly identifiable product. This means expressing an opinion about a makeup technique, generic product type (i.e. wooden lip liners, tubing mascara), or something like a place did not change this attribute from the default state. This was decided to ensure this category would remain relevant to understanding if and how endorsements fit into influencers' typical content.

Regarding the other potential values for this attribute, Positive and Negative were assigned if at least one product featured had an opinion expressed about it. In the case of mixed, this means that at least one product got both positive and negative comments. Both meant that at least one product had exclusively positive things said about it, while at least one more had exclusively negative things said about it. In a few cases it's relevant to note that mixed overrode Positive and that Negative/Both were considered to override Mixed, as it goes more against the grain to speak negatively about products than to say positive things. This is because speaking positively may make an influencer appear more likable, help garner favor from brands, induce purchase from an affiliate link, and more, while speaking negatively could do the opposite. Essentially, speaking poorly about products has more inherent risk.

When formatting the data for WEKA, Products_Review_Type was split into the other three review categories and turned into Boolean values. In the case of Not Applicable, all three columns would show 0. In the case of Both, Pros_Review and Neg_Review would both be 1. For only positive, negative, or mixed comments, just the respective column would have a 1.

With the makeup content attribute most posts are very straightforward to identify. Particularly when looking at influencers with over 100,000 followers in the beauty niche, a large portion of content is inherently going to be indisputably makeup content. However, there are two places classifications can get murky. First, because for the purposes of this thesis only makeup and related topics like skincare count there were times whether a judgement has to be made about whether something relates enough to makeup. For example, if an influencer was figuring out what accessories match their bumblebee makeup look that would still count as makeup content. Second, in some cases "transition videos" or reels may not have focus heavily on makeup but use it as a piece of a narrative. Overall, the guiding question in times of doubt for classifications was, "would the tone, point, message, understanding, or enjoyment of the piece change if the makeup were removed or different?"

In the cases of more subjective attribute designations, there is likely to be some arguably misclassified data. However, it can be argued that it is only a small fraction of the data. With the first attributes discussed relating to review types, in most cases influencers are incentivized to be more extreme rather than less as extreme opinions are more likely to evoke a reaction in

their audience (Konovalova, 2024). This means influencers are likely to remove doubt in most cases by explicitly stating “I hate this product” or “I love this product but I don’t like this thing about it”. However, it is notable that the most common cases that require thought to classify were cases where it was the tone rather than the words that betrayed whether something was an opinion or fact: for example “it’s $40” can be said as a simple fact or with tonal cues that suggest that product is a good or bad deal. In the second case involving makeup content classification, it is also quite likely that a beauty influencer clearly posts beauty content or clearly does not if, for example, they are sharing something personal. If they posted beauty content that did not have to do with their makeup like a hair tutorial, that was also usually a clear focus. As a result, while these attributes are subjective and there are likely multiple posts that could have their classifications argued, over the whole dataset these debatable cases should be a relatively small portion of the data.

2. Advertising information: attributes applicable in the presence of advertising

Table 2. Advertising Information Attribute Definitions and Considerations

Attribute Name	**Value Type**	**Description and Considerations**
Affiliate_Codes	Boolean	Indicates the presence of an affiliate code or text/verbal cues directing people to an affiliate link or TikTok Shop
Affiliate_ Profit_Disclosure	Boolean	Does the influencer disclose that they make money when an affiliate code or link is used either verbally, in text, or with a "creator earns a commission" disclosure label? (Yes/No)
Promoting_ Own_Brand _Or_Event	Boolean	Indicates the post is promotional for their own event or a product(s) from their own brand. Sometimes subjective
No_Ad	Boolean	Indicates the influencer indicates a product discussion is not a paid ad or they say or do something that implies this, such as saying they bought it or going shopping. Subjective
Stated_Ad	Boolean	Indicates the influencer verbally states that their video contains an ad or partnership
Written_Stated _Ad	Boolean	Indicates the influencer writes their video contains an ad or partnership in the description. Does not include hashtag disclosures
On_Screen_Ad_ Disclosure	Boolean	Indicates that on the video itself, the influencer has included a text overlay disclosing a paid ad or partnership

Sponsor_Hashtag _Start	Boolean	Indicates a generic (#ad) or specific partner (#TirTirPartner) hashtags are at the start of the description.
Sponsor_Hashtag _End	Boolean	Indicates a generic (#ad) or specific partner (#TirTirPartner) hashtags are in the description somewhere beyond the beginning words, most likely in hashtags at the post's end
Paid_Partnership _Disclosure	Boolean	Indicates the influencer has turned on TikTok's in-app labeling indicating an ad
PR_Product	Boolean	Indicates that a free product or experience gifted from a brand was featured
PR_Product_Implied	Boolean	Indicates that it was not explicitly stated, but it was implied that a product or experience gifted from by a brand was featured

In this category there are a few potentially subjective attributes that warrant further discussion and one attribute that could use further elaboration to fully understand.

For attribute Promoting_Own_Brand_Or_Event, while promotional content for an influencer's own brand or event may not be labeled with partner hashtags or use TikTok's labeling tools, it does tend to be quite clear. In general this is because in cases where this attribute is activated influencers are, as University of Michigan Ross School of Business professor Marcus Collins discusses in his book *For The Culture*, trying to motivate their viewers to move. This could be convincing them to sign up for a livestream event, purchase something, check out an event listing, or more. Worth noting is that this attribute only dealt with events and products related to their personal brand or a brand they own, not collaborations with outside brands.

No_Ad as an attribute could be broken out more similarly to how ads have been done, such as a stated or written version, but it was decided to keep it as one attribute. Similarly, Affiliate_Profit_Disclosure could be broken down further but was not. This is because the primary focus of this research is on disclosing ads, not disclosing things that are not ads. This could be unpacked in future research. However, for the sake of this thesis something triggering the No_Ad column could be extremely explicit or more implicit in nature. As more explicit examples, an influencer could state that an endorsement is not an ad or that they purchased a product with their own money. As a more subtle or implied trigger, an influencer may talk about former repurchases or appear to be going shopping at the store. In some cases, this may mean that a misleading ad could still activate the No_Ad attribute.

PR_Product focuses on a strategic marketing tool commonly employed by beauty companies; companies will send influencers that fit their brand and have enough reach product packages, often featuring new products or collections to create buzz and awareness (Kuras, 2024). Often these packages

may contain over-the-top or unusual packaging like jumbo sizing, likely to incentivize influencers to show the PR package on their channel. Regardless of packaging, influencers are already incentivized to quickly share about PR packages as it is an easy, free opportunity to be one of the first people to discuss a new launch and build goodwill with brands.

While this thesis will discuss some considerations and concerns about PR product disclosure more in depth in the discussion section, for now the focus will stay on the different values this attribute can have in the data. Yes is quite clear and indicates that the influencer has directly stated something within a video was PR or that a brand sent them something they're talking about. No can technically be misleading as it does not mean that no PR products are present but rather that the video has not given a specific indication that one is included. For example, multiple cases videos had to be labeled No for this attribute despite knowledge that products from PR packages were included because of previous video parts from the same influencer or other influencers' videos about the same product. The last possible value for this attribute is Implied, which was assigned when an influencer never explicitly stated that they are using products from PR packages but it is either implied by the packaging – which is usually wildly different from in-store packaging – or by saying something like the product will be coming out, suggesting the product was not acquired through publicly accessible means.

While technically many influencers should be disclosing free products in a similar manner to paid advertisements per their local regulations (Federal Trade Commission, 2019; European Commission, 2023), this does not seem to be a standard practice.

The final attribute that needs more discussion is Affiliate_Codes. Per Salesforce (2024), "affiliate marketing involves promoting another business's products or services online in return for a small affiliate commission after each sale". While this is not traditional advertising in the sense that an influencer is getting paid to produce a specific piece of content, they are incentivized to share on their own in order to reap a financial reward.

In this context, this often takes one of two forms. The first is affiliate codes, which see an influencer sharing a unique code that often has their name or TikTok handle included in it. At checkout on partnered sites their viewers can then input this code for a discount on their purchase. Meanwhile, the influencer then makes a small commission on the purchase. Similarly, the second form is affiliate links. In this case an influencer will either directly send sales pages for specific products to interested users (who may express their interest by commenting on a video) or keep the link posted in their bio and direct viewers there. In this case, there's usually not a discount for the purchaser and the sales page does not look different than if the purchaser had searched an item up on their own. However, the influencer still receives a commission when a purchase is made as the URL is specific to them. An

example of this can be Amazon storefronts, which many influencers use to link the products they use in their videos.

An issue with affiliate marketing is that it is generally not very transparent. While an influencer may tell their audience that their code will provide a checkout at purchase, viewers without marketing subject knowledge will likely not understand that the influencer is also getting paid. Similarly, linking products used in videos in an Amazon storefront, TikTok Shop, or similar can come across like the influencer is trying to help their audience find the products they are talking about while in reality they are trying to turn them into a revenue source.

While this is not to say that affiliate marketing can not be win-win, it does leave a huge question about transparency. That is why Affiliate_Profit_Disclosure was created to track whether influencers actually disclose in any way that they earn a commission off purchases. This can be disclosed in a number of ways at the behest of the influencer: verbally, on-screen, or in caption for example. Additionally, there is one standardized, automatic measure that can come into play when TikTok Shop items specifically are linked. When this happens, there is also an automatic "creator earns a commission" content label that is turned on.

3. Calculated information: attributes that do not contain unique information, but are rather calculations based on core post or advertising information.

Table 3. Calculated information attribute definitions and considerations

Attribute Name	**Value Type**	**Description and Considerations**
Norm_Likes	Numeric	Likes/Followers
Norm_ Comments	Numeric	Comments/Followers
Norm_Shares	Numeric	Shares/Followers
Norm_Saves	Numeric	Saves/Followers
Engagement_ Rate	Numeric	(Likes+Comments+Shares+Saves) /Followers
Log_Followers	Numeric	Log(1+Followers)
Windsorisation_ER	Numeric	Winsorized Engagement Rate, 95%
Low_Clarity	Boolean	Indicates post has low clarity ad disclosure
Med_Clarity	Boolean	Indicates post has medium clarity ad disclosure
High_Clarity	Boolean	Indicates post has high clarity ad disclosure
Sponsored	Boolean	Indicates post contains any form of ad disclosure (low, med, and/or high clarity)
Commercial_Content	Boolean	Indicates post contains any form of ad disclosure, affiliate link(s), and/or implied or confirmed PR materials

These attributes are created from other collected data for data analysis purposes, meaning they all contain formulas and are not based on new information. Those formulas and each attribute's purpose will be discussed below.

The attributes Norm_Likes, Norm_Comments, Norm_Shares, and Norm_Saves exist to be normalized versions of the traditional Likes, Comments, Shares, and Saves attributes. These are more useful for calculations than their raw counterparts because they attempt to account for variation from follower size. This puts a post from an influencer with 100,000 followers that receives 10,000 likes on the same level as a post that receives 100,000 from an influencer that has 1,000,000 followers. Without normalizing the data, larger influencers would have the opportunity to skew results dramatically.

This also ties into Engagement_Rate, which was intended to be the primary dependent variable as opposed to just, for example, likes. The engagement rate is a standardized metric often used in influencer marketing research and includes a sum of metrics that have to do with audience engagement divided by the follower count (Library of Congress, n.d.), which allows for better comparisons. This includes likes, comments, shares, and saves in this case, and will be expanded upon in Section 3.4 Analysis Tools and Data Analysis Plan.

While not originally in the plan, as discussed in Section 4 Analysis and Results both a Log_Followers attribute and a winsorized version to the 95th percentile were eventually added. These were added to improve regressions' ability to properly work with the data by reducing the extreme impact of viral posts and differences between influencers' sizes. These do not perfectly eliminate the issues associated with the data set but they do improve the outcomes.

Low_Clarity, Med_Clarity, and High_Clarity indicate whether posts had disclosure types that meet the categories previously designated in Section 3.2 Disclosure Analysis. These were created to help determine whether the different classes of ad disclosure lead to different outcomes on a more categorical rather than granular level.

Sponsored and Commercial_Content serve similar purposes for analysis but have one key difference. Posts with a 1 in the Sponsored column are those that have any type of traditional ad disclosure, regardless of the clarity of those disclosure methods. Commercial_Content is more comprehensive, as it includes any posts that have confirmed or implied PR products being used as well as those with affiliate links too. While these are not traditional advertising, they certainly count as commercial content and are a manipulation by brands of what content is consumed by the public.

Sponsored and Commercial_Content serve as post-type labels along with Promoting_Own_Brand_Or_Event, which shows promotional content, and

Affiliate_Codes, which shows affiliate marketing content, from the advertising attributes group.

While it was not certain whether all of these calculations would ultimately be useful during modeling, these were created with the intention that they could be experimented with.

3.4 Analysis Tools and Data Analysis Plan

This thesis employed machine learning for its quantitative analysis. Specifically, it first employed a kiwi software suite known as Waikato Environment for Knowledge Analysis (WEKA) followed by University of Amsterdam-based Jeffrey's Amazing Statistics Program (JASP).

WEKA was chosen because of its strong compatibility with structured datasets and visual exploration capabilities while also being an open-source machine learning resource. Additionally, it has the benefit of accessibility for testing multiple model types in one place without the need for manual coding on the part of the researcher. Extending off that, WEKA's ability to combine multiple model types (including classification, clustering, and regression models) in a single environment made it a good choice for this thesis.

Before importing the dataset into WEKA, the data was cleaned in Excel and exported as CSV file to make it compatible. Cleaning the data meant converting categorical variables like Featuring_Free_PR_Product and Products_Review_Type from single columns into series of binary variables. For example, Products_Review_Type was split into Positive_Review, Negative_Review, and Mixed_Review to accurately demonstrate that in the case of "Both" an influencer was expressing both positive and negative content. Additionally, cleaning the data also meant also meant verifying that all continuous variables such as Likes, Followers, and Engagement_Rate were assigned as numeric variables, as incorrect labeling could lead to incorrect analytics.

As mentioned at the start of Section 3, the final dataset collected contains 2,250 observations from across 45 influencers, each representing one individual TikTok post. While each post ultimately included 39 attributes relating to influencer characteristics, disclosure styles, discussion tone, and engagement performance, some attributes are more critical to analysis than others. Given the overall goal of this analysis is to evaluate how different types of advertising disclosures and product review tones influence engagement as a proxy for consumer trust, the most important attribute is the engagement rate. This can be calculated with the following formula:

$$Engagement\ Rate = \frac{Likes + Comments + Shares + Saves}{Followers}$$

After preparing the data, which involved further transforming the engagement rate to reduce skewness, multiple analytical methods were tested in WEKA's Explorer and a linear regression was selected to be the main model for this thesis. This makes sense because of its ease of interpretation and clear alignment with the research questions; linear regression allows for the estimation of how each disclosure type individually contributes to engagement rate variation while also controlling for other factors like affiliate marketing. This model also made sense because it allows for links between this thesis's empirical findings and the theoretical frameworks discussed earlier, such as the Persuasion Knowledge Model (Friestad & Wright, 1994), which predicts measurable shifts in perception when persuasion attempts are recognized.

A second analysis run on WEKA created piecewise linear regressions were done to look at the impact of different clarity tiers on engagement. The goal of this was to see how different disclosure types changed the winsorized engagement rate and to see if there was a notable difference in their impact.

In addition to the regression models above, there was also a desire to explore whether the tone of reviews moderated the effects of branded content. Consequently, interaction terms (Sponsored x Neg_Review, Sponsored x Mixed_Review) were created to test whether endless positivity or honest "deinfluencing" comments impacted the potential engagement decrease associated with commercialized content. These interaction effects dive into the secondary research question of this thesis regarding authenticity and trust.

Regarding the performance of the models, WEKA's included analysis features were employed. First, WEKA's 10-fold cross-validation feature was used to evaluate consistency across the sample and reduce the risk of overfitting. Complementary diagnostic metrics, including the correlation coefficient (R), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), were also used to gauge predictive performance. Finally, coefficient significance from the regression outputs were analyzed in Section 4.2; these provided not just standardized coefficients, but also additional values for each independent variable. Per standard practice, it was important to consider significance as well, which will be touched upon shortly; this allowed for attributes' effects on engagement to be judged as meaningful or not, helping conclusions pulled from the data to be considered statistically important and not just driven by random variation.

Overall, WEKA's modeling environment was a good choice for this thesis because it was an easily accessible software with a range of analysis options built in. Linear regressions can also be considered a fitting tool for the majority of this thesis's analysis. However, while WEKA is great for preprocessing and modeling large variable sets efficiently, allowing this thesis to test multiple model specifications (clarity levels, tone interactions) and generate cross-validated performance measures (MAE, RMSE), as

mentioned previously this data needs to be supplemented to be robust. Consequently, the data was also run on JASP to assess which effects were significantly significant with p-values and standardized coefficients.

Although both platforms implement linear regression, they have different interfaces, assumptions, and diagnostic outputs. Based on the hypotheses, it was expected that WEKA models would expose consistent directional patterns related to engagement, sponsorship penalties, and the impact of tone, while JASP would confirm which of these effects are significant. Consequently, they provide complementary insights into the research questions that allow for a robust understanding of the data. Some similarities and differences between the two programs that are relevant to this thesis can be found in Table 4.

Table 4. Core Comparisons Between WEKA and JASP

Aspect	WEKA	JASP	Relevance to Thesis
Primary Focus	Machine learning and predictive modeling	Statistical inference & hypothesis testing	WEKA tests multiple variable combinations efficiently; JASP assesses statistical significance (p-values, CIs)
Regression Type	Ordinary Least Squares (OLS) via data-mining framework	Classical OLS via frequentist or Bayesian methods	Both yield comparable coefficients; JASP adds inferential tests
Output Metrics	Correlation coefficient, MAE, RMSE, relative errors	β coefficients, SE, t, p, R^2, adjusted R^2	WEKA emphasizes model fit; JASP quantifies statistical reliability
Variable Handling	Handles binary/nominal and continuous predictors automatically, supports filters for transformations	Requires preprocessed numeric/categorical variables defined by user	WEKA's filters (Standardize, AddExpression) simplify transformations like interactions; JASP requires manual creation
Interactions & Transformations	Must be manually created using AddExpression	Automatically generated via interaction options	Interactions (i.e. Sponsored × Positivity) are created manually in WEKA but are automatically specified in JASP
Model Validation	Cross-validation built-	No built-in cross-validation (focus on	WEKA provides out-of-sample performance metrics;

	in for predictive performance	fit to current sample)	JASP verifies in-sample statistical significance
Interpretation Style	Predictive (machine learning lens)	Inferential (statistical lens)	WEKA results show pattern stability; JASP provides confirmatory hypothesis testing

4 Analysis and Results

In total, 2250 data points were used and of those 298 (13.24%) fell under traditional sponsored post advertising. On a raw basis mean engagement was calculated as being 3.11% for sponsored posts averaging 2.00% and non-sponsored posts being 3.28%, giving a raw difference of -1.28 percentage points.

If affiliate advertising is included, the number of posts jumps from 298 to 451 posts (20.04%), of which 26 contain both traditional ad content and affiliate marketing. Even without considering the implications of PR packages, this immediately reemphasizes the importance and value of the purpose of this study; one in five posts contains some form of paid advertising. If including PR gifts, 690 posts (30.67%) contain commercial content.

4.1 Data Transformations

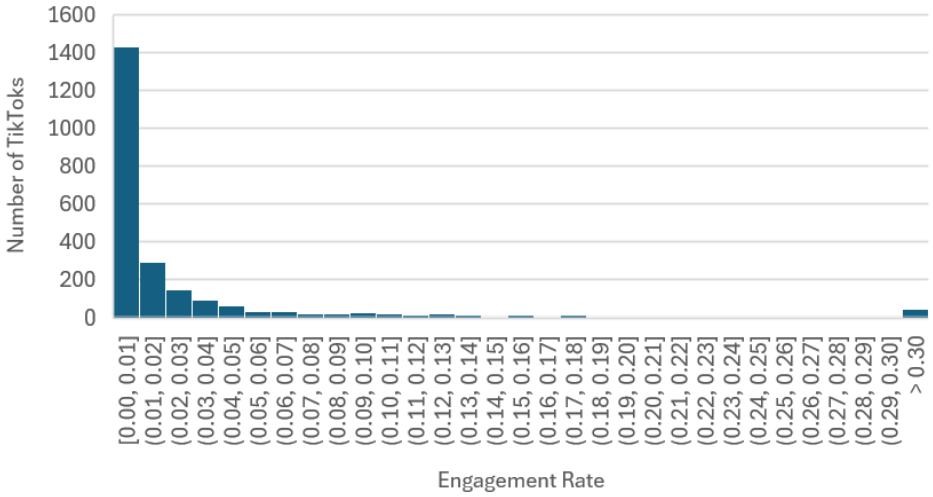

Figure 4. TikTok Engagement Rate Frequency for Beauty Influencers with 100,000+ followers (n=2250)

Before uploading the collected data to WEKA in CSV format, it was determined that one more data transformation originally not planned for was needed. To start, the Engagement_Rate variable was turned from a standard numeric variable into Log_ER with the following formula to see if it would fix an obvious issue with the data as seen in *Figure 4*.

$$Log\_ER = log(1 + Engagement\ Rate)$$

This was done to deal with engagement data being very skewed. While most TikTok posts have low engagement rates, with 1426 (63.38%) boasting

an engagement rate below 1%, the highest engagement rate recorded was from the TikTok itsbabykelz_31, which had a 315.45% engagement rate. The high engagement posts could be considered outlier viral posts and hint that a transformation could be needed. Using Excel to perform skewedness and kurtosis evaluations on the unmodified engagement rate this was confirmed; the data showed a 13.20 skewedness score, making it extremely right-skewed, and a 248.17 kurtosis value, which makes the data extremely leptokurtic with lots of extreme outliers.

Unfortunately, while transforming the data into log_ER as in *Figure* 5 reduced the extreme skewing in the data it did not eliminate it enough to proceed. While it's generally desired to have a skewness around +/-2 and kurtosis of around +/-7 (Hair et al., 2010; Bryne, 2010), the transformation still left a skewness of 8.14 and kurtosis of 92.34. This reflects the presence of a few viral posts with essentially unprecedentedly high engagement. Fortunately, this is not unexpected in the environment of social media.

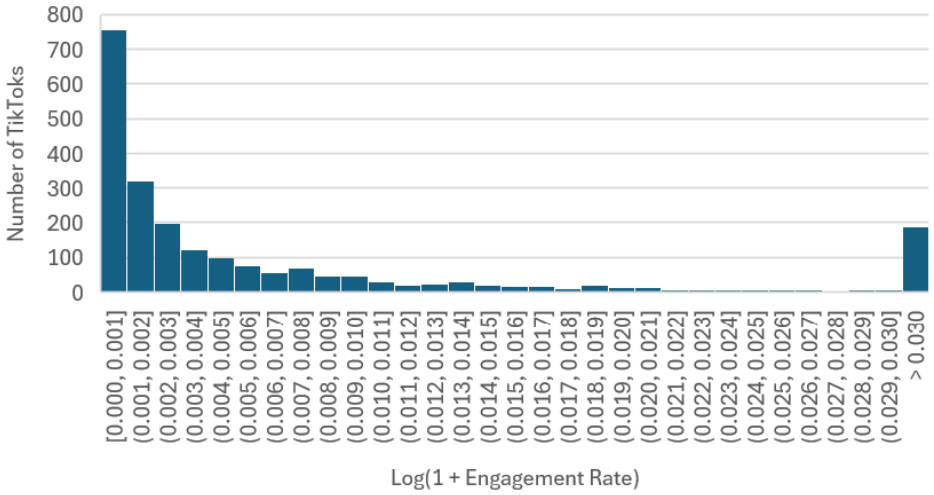

Figure 5. Logarithmic TikTok engagement rate frequency for beauty influencers with 100,000+ followers (n=2250)

This led to a final adjustment to the data before WEKA analysis: winsorization, which would help limit the influence of outliers on the dataset by flattening peaks. With the intension to using linear regressions during the WEKA analysis and the nature of viral posts being seemingly unpredictable, this step is appropriate in this context. After calculating the 95th percentile as 0.12, all values greater than that value were reduced to it for a winsorized engagement rate column. 95th percentile was selected as it reduced skewness to 2.04 and kurtosis to 3.945. This indicates a substantially reduced right-tail. While this data still isn't perfectly normal, which is to be expected with organic data like social media data, it falls within a reasonable range to be able to perform analysis. Ultimately, this means that 113 values were

considered outliers and flattened for this analysis. The final histogram can be seen in Figure 6.

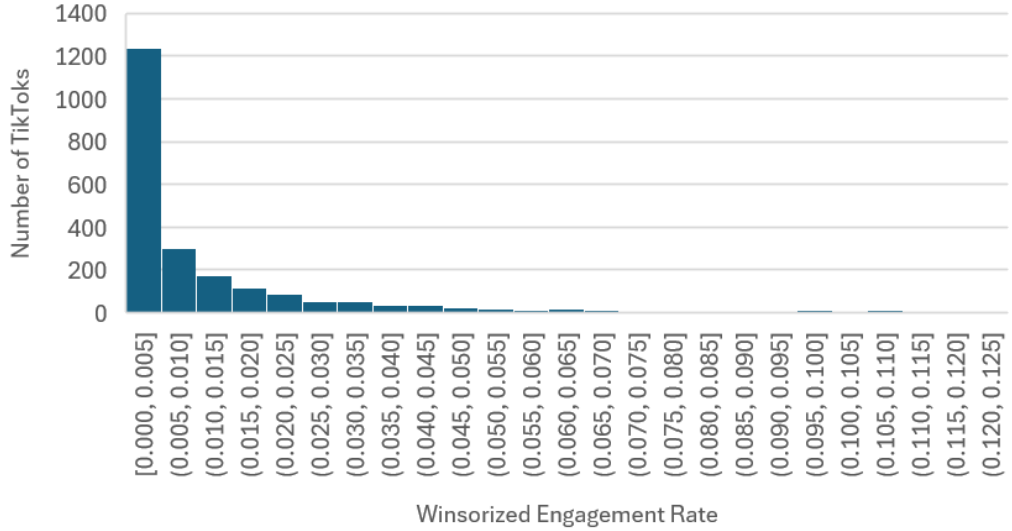

Figure 6. TikTok 95th Percentile Winsorized Engagement Rate Frequency for Beauty Influencers with 100,000+ Followers (n=2250)

4.2 WEKA and JASP Analysis

Before diving in more granularly about disclosure variables, it was decided to first run a higher-level analysis about what influences engagement rate. Based on previous research, if the data is sound it should have unsurprising results.

With the winsorized engagement rate as the class (dependent variable) a linear regression (OLS) with 10-fold cross validation using greedy attribute selection to keep the output tidy was run. The classifier ClassifierAttributeEval with a Ranker search method was used. Repetitive variables such as normal Likes and Commercial_Intent were removed from analysis to prevent multicollinearity issues. The following regression was formed:

$$\begin{aligned} Winsorized_{ER} &= 0(Brands\_Tagged) + 0.0071(Affiliate\_Codes) \\ &+ 0.0041(Stated\_No\_Ad) + 0.0044(Neg\_Review) \\ &+ 0.0042(Mixed\_Review) + 0.0163(Pos\_Only\_Percent) \\ &- 0.0041(Implied\_PR\_Product) - 0.0029(Sponsored) \\ &+ 0.0041 \end{aligned}$$

With this formula there are five attributes included that improve the engagement of the post. The first is Affiliate_Codes, which show slightly higher engagement. This could be for one of two reasons: it could be that the audience responds positively to the value of discount codes, particularly since

the commercial link is rarely exposed, or that the type of content that tends to have affiliate codes is popular; affiliate codes are often included on videos that show off products in some sort of tutorial. An influencer stating or implying that their post isn't an ad is an unsurprising addition to this list, as it would inherently seem more authentic.

The last three positive indicators are interesting when taken together. Starting with Neg_Review and Mixed_Review, they provide a positive boost to a TikTok and likely for similar reasons; they suggest an influencer is engaging in authentic "deinfluencing" behavior and possibly even allowing for some controversy that would be perceived as transparent. However, Pos_Only_Percent has a positive coefficient. With this attribute being positive it suggests that accounts that post mostly positive content earn a higher engagement rate. This could be for a number of reasons including positivity bias or the way people's algorithms are set up (it would be safe to assume that people generally watch content they enjoy). Regardless, this leads to an interesting point that more positive content overall on an influencer's feed is a smart strategic choice for engagement but that can also lead to negative or mixed reviews being more impactful. If an influencer's reviews are trusted and typically positive, this could explain why it is a bigger deal when they have something negative to say.

On the other hand, there are two attributes that decrease the likelihood of engagement. The first is Implied_PR_Product, which could be aligned with the suggestion that a lack of transparency reduces the assumption of authenticity and triggers viewers to act as though they're exposed to an undisclosed persuasion attempt. The second is sponsored content, which reflects both Friestad and Wright's PKM (1994) and Eisend's suggestions that individuals will question whether a knowledge source's inclination for the truth when judging credibility.

Brands_Tagged had some amount of significance such that the model did not remove it, but its value was small enough that it was pruned and rounded to zero.

Table 5. Winsorized Engagement Rate Linear Regression on WEKA Model Fit Statistics

Metric	**Value**
Correlation Coefficient (R)	0.104
Mean Absolute Error (MAE)	0.0137
Root Mean Square Error (RMSE)	0.0217
R^2	0.011

While at first blush the correlation coefficient and R^2 do not seem particularly impressive, as the model only finds a weak linear association and accounts for just 1.1% of variance, this is does not on its own mean that the model should be ignored; prior influencer engagement research suggests that

R^2 values will not be huge in this research area to the amount of noise found in social media data (Ki et al., 2020). In fact, other research related specifically to ads in social media supports the idea that models are often modest in effect but significant (Jin & Ryu, 2021).

Before diving into determining statistical significance however, it's worth noting that this model did not account for standardization. That means that while the direction and real-units impact of attributes on the model can be assessed with coefficients, the coefficients are not able to be compared to each other for relative strength. After applying WEKA's attribute standardization, which converts all numeric variables to mean 0 and standard deviation 1 while leaving binary indicator variables unchanged, the formula changes to:

$$\begin{aligned} Winsorized_{ER\_Standardized} &= 0.0336(Brands\_Tagged) + 0.0916(Affiliate\_Codes) \\ &+ 0.0488(Stated\_No\_Ad) + 0.0504(Neg\_Review) \\ &+ 0.0568(Mixed\_Review) + 0.1408(Pos\_Only\_Percent) \\ &- 0.0368(Implied\_PR\_Product) - 0.047(Sponsored) + 0 \end{aligned}$$

Now that the model has been standardized there's a few interesting changes and things to note. First, Brands_Tagged was reintroduced after being pruned in the unstandardized regression due to a near-zero coefficient. However, its effect stayed comparatively negligible, suggesting that brand tagging has relatively little independent influence on engagement. Second, it can be seen that the most influential predictor on engagement here is Pos_Only_Percent and the least influential compared to the others are Implied_PR_Product and, unsurprisingly, Brands_Tagged. As per the unstandardized version of the model, the R^2 in Table 6 still suggests a quite modest effect as would be expected with social media data.

Table 6. Winsorized Engagement Rate Linear Regression on WEKA Model Fit Statistics with Standardization

Metric	**Value**
Correlation Coefficient (R)	0.107
Mean Absolute Error (MAE)	0.648
Root Mean Square Error (RMSE)	1.016
R^2	0.011

From this point it is clear from the WEKA analysis that there are directional impacts from the variables as it relates to engagement based on the coefficients, but statistical significance is still in question. Thus, replicating this linear regression in JASP and running it to obtain inferential statistics like p-values and an F-statistic is the next step; re-estimating the WEKA model in JASP strengthens the credibility of this analysis by diving

deeper into both the model as a whole and individual predictors. The results of this analysis, once again using Winsorized_ER as the dependent variable, can be found in Table 7 and Table 8.

Table 7. Winsorized Engagement Rate Linear Regression on JASP Regression Table

Predictor	**B**	**SE**	**β**	**t**	**p**
Affiliate_Codes	0.012	0.002	0.100	4.881	< 0.001
Brands_Tagged	0.000	0.000	0.094	4.605	< 0.001
Stated_No_Ad	0.000	0.003	0.002	0.098	0.922
Neg_Review	0.006	0.003	0.049	2.347	0.019
Mixed_Review	0.008	0.002	0.072	3.378	< 0.001
Pos_Only_Percent	0.033	0.004	0.188	9.011	< 0.001
Implied_PR_Product	-0.003	0.003	-0.047	-2.310	0.021
Sponsored	-0.006	0.002	-0.065	3.140	0.002

Table 8. Winsorized Engagement Rate Linear Regression on JASP Model Fit Statistics

Metric	**Value**
Correlation Coefficient (R)	0.264
R^2 (Approx)	0.070
Adjusted R^2	0.066
RMSE	0.031
F(8, 2241)	21.02
p (model)	< 0.001

Based on the JASP analysis, it can be said that the overall model for predicting engagement explains roughly 7% of the variance in engagement (Adjusted $R^2 = 0.066$). Coupled with an F-statistic of $F(8, 2241) = 21.02$, $p < 0.001$, the model is confirmed significant; it does a better job of predicting engagement than having no predictors and relying on chance. Several predictors were statistically significant.

Zooming in on the individual predictors, the strongest predictor was Pos_Only_Percent, which suggested that a higher percentage of positive product-related statements leads to substantially higher engagement ($\beta = 0.188$, $p < 0.001$). Posts that include affiliate codes ($\beta = 0.100$, $p < 0.001$) or tag more brands ($\beta = 0.094$, $p < 0.001$) also showed an increase in engagement. While more subtle in their effect, the tone attributes Mixed_Review ($\beta = 0.072$, $p < 0.001$) and Neg_Review ($\beta = 0.049$, $p = 0.019$) rounded out the group of attributes that predicted higher engagement. Meanwhile, sponsored posts showed a notable engagement penalty ($\beta = -0.065$, $p = 0.002$) as well as posts where PR was implied without an explicit disclosure ($\beta = -0.047$, $p = 0.021$).

The Stated_No_Ad variable was not a significant predictor of engagement ($\beta = 0.002$, $p = 0.922$). This indicates that explicitly stating a post is not an

ad does not affect engagement rates. The very high p-value suggests that the true effect is likely zero and the observed coefficient is indistinct from random noise in the data, which is essentially bound to be high with real social media data. In these cases, TikTok viewers are not activated to deal with a persuasion attempt as they would not recognize one occurring (Friestad & Wright, 1994).

After establishing this baseline model, it made sense to push more specifically into the research question and look more closely at different forms of ad disclosure. This will dive into another linear regression focused on the winsorized engagement rate. Regarding which attributes will be relevant for this analysis, it will be the three clarity metrics as well as log_followers and makeup_content as controls. During this part of the analysis, no standardization will occur because the clarity variables are already 0/1 indicators.

This process will have three models to create baselines for each clarity metric that are relative to the others. First, Low_Clarity was dropped and the regression was run with the winsorized engagement rate still set as the class.

$$Winsorized_{ER\_LC\_Baseline} = -0.0037(Med\_Clarity) - 0.0021(Log\_Followers) + 0.0246$$

With this outcome, it can be said that medium-clarity posts have, on average, 0.37 percentage points lower engagements than low-clarity posts, controlling for follower count; on average, when posts disclose ads at the medium-clarity level, they get less engagement than low-clarity posts. Additionally, as expected larger accounts have lower engagement rates so there is a scale penalty. The intercept can be interpreted as the baseline for low-clarity posts' engagement rate when log-scale normalized. However, it's worth noting that once again the R^2 value is quite low, at just 0.0014.

Repeating this process after dropping Med_Clarity and adding Low_Clarity back to the model, the following results are revealed:

$$Winsorized_{ER\_MC\_Baseline} = -0.0032(High\_Clarity) - 0.0021(Log\_Followers) + 0.0248$$

With this outcome, it can be said that high-clarity posts have, on average, 0.32 percentage points lower engagements than medium-clarity posts, controlling for follower count; on average, when posts disclose ads at the high-clarity level, they get less engagement than medium-clarity posts. Additionally, as expected larger accounts still have lower engagement rates so there is still a scale penalty. The intercept can be interpreted as the baseline for medium-clarity posts' engagement rate when log-scale normalized. Once again the R^2 value is also quite low, at just 0.0014.

Repeating this process after dropping High_Clarity and adding Med_Clarity back to the model, the following results are revealed:

$$Winsorized_{ER\_HC\_Baseline} = -0.0037(Med\_Clarity) - 0.0021(Log\_Followers) + 0.0246$$

Now that these models have been run, differences between disclosure-clarity levels can be assessed. As a distinct note, there is a sign reversal that occurs in the final formula that has created a negative coefficient. However, it does not strictly reverse the pattern that higher levels of disclosure are yielding higher engagement penalties. This is because there is overlap in the dummies of the clarity tiers due to some posts having both high and low clarity disclosures, creating shared variance and issues with multicollinearity. This makes logical sense because a post using high clarity disclosure practices is likely to use some lesser levels of disclosure as well. This could be interpreted as meaning that, when there is overlap, medium-coded content does not significantly impact engagement once high-clarity disclosure methods are used. In more technical terms, controlling for high-clarity disclosures medium-clarity disclosures don't have a significant effect. While it did not impact the sign in this case, the partial effects impact was likely also relevant for low-clarity overlap.

Drawing on Friestad and Wright's PKM (1994), this thesis would be confident in asserting that the unexpected sign is likely due to the above and not high-clarity disclosures having a lesser impact compared to medium-clarity disclosures. However, further research with a dataset that gives the ability to focus only on datapoints that do not have overlap would be salient to fully confirm this theory. But regardless of the conclusiveness of the differences between low, medium, and high clarity sponsorship disclosure, the WEKA does not rule out decreasing consumer engagement with more obvious sponsorship disclosure.

While a one-way ANOVA would ordinarily be used to further understand the differences between the the clarity tiers, it was not appropriate due to the possibility for posts to fall into more than one group. Seeking out sponsored posts that are mutually exclusive in how they disclose sponsorships in terms of clarity could be an opportunity for useful future research.

Moving on to a final analysis focused more on RQ2, which is functionally a moderation question about whether positivity strengthens or weakens the sponsorship penalty found in the first piece of analysis or whether mixed/negative messaging has an impact. For this, interactions terms between Sponsored and Pos_Only_Percent as well as tone variables Neg_Review and Mixed_Review are needed alongside the standard attributes of Winsorized_ER and Log_Followers. This is outlined below in Table 9.

Table 9. Setup for Interaction Analysis Between Sponsored Posts and Review Type Frequency

Metric	**Role**	**Purpose**
Sponsored	Main Effect	Baseline impact of an ad on engagement, usually negative
Pos_Only_Percent	Main Effect	Does more positive tone increase engagement overall?
Sponsored X Pos_Only_Percent	Interaction	Does sponsorship penalty depend on tone positivity?
Neg_Review	Tone Category	Authenticity cue
Mixed_Review	Tone Category	Authenticity cue
Log_Followers	Control	Scale penalty for larger audiences

The first step in this analysis is to use the AddExpression filter under unsupervised attributes to add 3 interaction terms: Sponsored_Pos_Int (Aponsored x Pos_Only_Percent), Sponsored_Neg_Int (Sponsored x Neg_Review), and Sponsored_Mix_Int (Sponsored x Mixed_Review).

Next, Windsorized_ER was once again selected as the class attribute for a linear regression with 10-fold cross validation. The following formula is the result:

$$Winsorized_{ER\_Interactions} = 0.0042(Neg\_Review) + 0.0041(Mixed\_Review) + 0.0151(Pos\_Only\_Percent) - 0.0092(Sponsored) + 0.0144(Sponsored\_Pos\_Int) + 0.0054$$

Now to use this to unpack the general query of, does the frequency and tone of product discussion (positive, negative, or mixed) affect the impact of ads on engagement, the pieces can be pulled apart and analyzed.

By taking the absolute value of the Sponsored coefficient (0.0092) and dividing it by the interaction coefficient above (0.0144), a value of 64% is found. This is particularly interesting because it suggets that based on coefficient ratios, the penalty for a post being sponsored disappears when approximately 64% or more of product-related statements are positive. More tangibly, roughly two-thirds of an influencer's content would want to be positive to see a boost. This suggests that highly upbeat messaging can compensate for the usual decline in engagement associated with ads, despite negative and mixed reviews showing a slight boost on a more individualized basis.

Aligning this with both the PKM and the idea of the path of absolute authenticity, positive posts may stand out less when the content around it as already largely positive as well. In this way, sponsorships don't stand out like a sore thumb and seem much more authentic. However, a multiple regression with an interaction term between positivity and sponsorship on JASP can provide further insight on whether positive messaging truly reduces the sponsorship penalty.

However, before moving into that analysis, the output of which can be found in Table 10 and Table 11, it's still worth noting that negative and mixed reviews have positive coefficients despite overall positive messaging showing higher engagement. While this is discussing individual datapoints not trends, if significance holds this suggests that positive and negative data can both be useful in the right post. There is a consideration to this though: sponsored ads are very unlikely to have negative commentary in them, as the influencer is being paid to help the company sell. That would make the interaction with the percentage of posts with positive commentary a more useful takeaway within the context of this thesis.

Table 10. Winsorized Engagement Rate Linear Regression on JASP with Sponsored x Pos_Only_Percent Interaction Term Regression Table

Predictor	**B**	**SE**	**β**	**t**	**p**
Pos_Only_Percent	0.032	0.004	0.183	7.990	< 0.001
Mixed_Review	0.007	0.002	0.062	2.963	0.003
Neg_Review	0.005	0.003	0.042	1.996	0.046
Sponsored	-0.010	0.005	-0.109	-2.222	0.026
Pos_Only_Percent x Sponsored	0.010	0.009	0.051	1.016	0.310

Table 11. Winsorized Engagement Rate Linear Regression on JASP with Sponsored x Pos_Only_Percent Interaction Term Model Fit Statistics

Metric	**Value**
Correlation Coefficient (R)	0.221
R^2	0.049
Adjusted R^2	0.047
RMSE	0.031
F(5, 2244)	23.09
p (model)	< 0.001

The results of the JASP analysis show that the interaction term was not statistically significant ($\beta = 0.051$, $p = 0.310$), which indicates that the percentage of positive product-related statements does not meaningfully impact the effect of sponsorship on. Consequently, the results from WEKA appear to be a result of using standardized coefficients without inferential statistics. Despite positivity strongly increasing engagement as a whole ($\beta = 0.183$, $p < 0.001$) and sponsorship independently decreasing engagement as

was previously established with the initial model (β = -0.109, p = 0.026), these attributes do not buffer each other in a meaningful way and can be said to operate additively rather than interactively.

5 Discussion

5.1 Summary of Key Findings

This thesis investigated two primary research questions, which are listed below alongside their associated hypotheses and relevant findings from the analysis:

RQ1: How do influencer ad disclosure styles on TikTok affect follower trust?
H1: More explicit ad disclosures will lead to higher engagement, suggesting higher viewer trust.

With regard to RQ1, results ultimately showed a directional trend in line with H1's expectation of differences with greater penalties at higher clarity levels, but the results were ultimately inconclusive due to multicollinearity. However, the analysis did show that visible sponsorship reduces engagement, suggesting a trust penalty, which is in line with results from other social media platforms. Notably, explicit PR disclosures were not shown to significantly affect engagement, while ambiguous PR disclosures reduced engagement and affiliate codes increased engagement.
Overall, H1 was not supported because clarity tier effects were not statistically significant.

RQ2: Does the typical tone of an influencer's reviews moderate an engagement drop caused by sponsorships?
H2: A more positive typical review tone reduces the engagement drop caused by sponsorships.

With regard to RQ2, results suggested a strong relationship between a typically positive tone when reviewing products and engagement, but it did not moderate the sponsorship penalty. This suggests that the effects are additive rather than interactive.
Overall, H2 was not supported.

5.2 Theoretical Interpretation

As a whole, this thesis is very in line with previous research and can be understood by drawing on the theoretical frameworks previously considered. For example, looking at Friestad and Wright's (1994) PKM helps explain why visible sponsorship reduces engagement; once commercial intent is recognized, they will start evaluating the persuasion attempt and react to maintain some control of the outcome of the engagement. This makes

viewers more likely to be skeptical of the message they are hearing and less likely to respond positively to an ad. This aligns with the clear sponsorship penalty observed in the analysis.

Moving on to PIT there is even more insight. Influencers build ongoing (albeit one-sided) relationship-like connections with their followers, which in turn builds their trust and credibility to their audience. Sponsorship disrupts this dynamic by giving the influencer a clear motive that interrupts their trustworthiness per Eisend's model of salesperson credibility (2006) and changing the interaction from feeling relational to commercial. The reduction in engagement from ads reflects this interruption of the parasocial connection.

The AMF further clarifies the different outcomes that come from various commercial cues. Explicit PR disclosures did not significantly change engagement, which likely reflects the "explicit authenticity" dimension as the influencer is excited to share a free product experience they are having with their viewers. On the other hand, ambiguous PR disclosures reduced engagement, implying that unclear disclosures may create an air of inauthenticity in content. This could align well with either "fake authenticity" or "fairytale authenticity" depending on the specifics of the situation. On the other hand, as affiliate codes increased engagement, this may suggest an alignment with the "absolute authenticity" as often viewers are watching products being promoted get used in an authentic, native way. Additionally, there are potentially perceived benefits for this type of disclosure such as discounts or lowered search costs.

All together, these theories show that transparency alone is not the sole driver of trust. Instead, follower reactions to situations like viewing sponsored content depends on how commercial cues react with perceptions of honesty, relational closeness, audience expectations and benefits, and the influencer's motives.

5.3 Practical Implications

The findings in this study allow for some practical insights that can be applied by influencers, brands, and even platforms like TikTok themselves when navigating how to deal with influencer marketing and commercial content.

For influencers, the key unexpected point is that visible sponsorship will reliably reduce engagement on posts. This means that creators should expect that when they take commercial deals they are accepting a more lukewarm audience response regardless of their post tone or the quality of their content. Additionally, the results show an importance for influencers to find a way to disclose PR items; not only do some countries, such as the United States, have regulations requiring it (Federal Trade Commission, 2019), but explicit disclosures do not see an engagement penalty while ambiguous implied

disclosures do. Additionally, the positive engagement impact of affiliate codes suggests that viewers enjoy tangible value regardless of a commercial element. Affiliate codes would be a way influencers could make money while also not sacrificing engagement.

For brands, this thesis reinforces the importance of choosing influencers whose audiences perceive them as authentic, positive reviewers. Since sponsorship penalties are a given, they should choose influencers with naturally high existing engagement with a posting style that fits their brand to achieve "absolute authenticity". Brands should also provide clear, structured guidance for disclosures to their influencers that avoid ambiguity, particularly as it relates to PR products, and consider growing their affiliate programs to capitalize on consumer interest.

For TikTok and other platforms like it, the results of this thesis point to a need for clearer and more consistent disclosure guidelines and tools. Ambiguous PR content not only harms engagement for influencers, but notably is indisputably a commercial attempt by brands that can easily deceive a viewer. While a few influencers have come up with strategies to improve the transparency of PR products being used, the majority only mention it seemingly in passing or when they first receive the products. In both cases, it doesn't seem like the disclosures are for the purpose of emphasizing transparency, rather to make conversation with the camera.

A small case study from the data illustrates the lack of disclosure standardization around PR: @f.bermann is one of the rare influencers who consistently identifies gifted products. In her case, she has decided to mark gifted products in her TikToks' descriptions with an *. Out of the 50 TikToks that were logged for her, 24 (48.00%) had at least one product marked as gifted. With only 255,300 followers, which makes @f.bermann a relatively small influencer for this list, it wouldn't be unreasonable to assume that some influencers receive more PR packages than she does and use the products even more often. Even extrapolating an equivalent rate across influencers, this means nearly half the content beauty fans see includes gifted products.

While that doesn't turn influencers' TikToks into ads in the traditional sense of a company paying for someone to post nice things, this does show that beauty companies have a significantly more massive pull and influence on what becomes popular than first meets the eye. Whether a full PR package gets its own unboxing video on an influencer's feed or the products within are only used as parts of future makeup looks, PR packages lead to unpaid ads for products with functionally zero de facto expectations of disclosure.

Additionally, social media platforms need to be better at enforcing their rules to protect their users. As previously mentioned, it should be standard practice given TikTok's rules that influencers turn on the content disclosure setting for their sponsored videos. This means that every ad that was recorded in the dataset should have triggered the Paid_Partnership attribute, but only 142 out of 298 posts identified as sponsored content (47.65%)

followed this rule. While the support page claims that failure to comply with the rule may lead to removed or restricted posts, it does not seem like TikTok does a good job of policing this based on the statistics. This means that there is not much consistency with how advertisements are disclosed and that this research is particularly salient.

5.4 Societal/Ethical Implications

With the impact of both social media and beauty on mental health and self-perception, as well as the widespread reach of TikTok, highlighted in the introduction, there are important ethical implications to consider with this thesis. The consistent sponsorship penalty and negative impact of ambiguous PR disclosures show the importance of ensuring transparency in digital environments where commercial messaging is getting increasingly blended in with personal content. In line with Dupuis et al.'s (2024) research in the food space on TikTok, it can be difficult for individuals, particularly younger audiences, to distinguish between authentic recommendations and paid sponsorships, particularly given inconsistently applied and unstandardized disclosures. This leaves consumers vulnerable and raises questions about consumer protection, informed decision-making, and fairness in both social media spaces and digital marketplaces.

Ambiguous PR disclosures are also ethically problematic because they take advantage of the trust developed through parasocial relationships. When influencers do not make it clear which products they have been gifted, viewers may not register a persuasion attempt and thus have their purchasing decisions shaped without adequate awareness of commercial intent. This undermines transparency and authenticity and could ultimately erode trust in influencers as well as brands who partake in influencer marketing.

However, this is not to accuse influencers of deliberately misleading their viewers. While they may have some understanding that brands are using a persuasion technique on them by sending free products, they likely do not realize they are adding opacity to the flow of influence between those brands and the general consumer. Additionally, keeping track of everything that has been purchased versus gifted would likely be a hassle to some extent, but PR products should be treated as commercial initiatives because they are.

Lastly, TikTok only having one clear rule regarding disclosing promotional content and still having under half of commercial posts in compliance with that policy suggests a need for stronger platform-level governance. Rules breed trust in users regarding the protections they should expect, and inconsistent enforcement allows for deceptive and unclear advertising practice to continue. This leads to uninformed decision making and understanding of the content viewers are consuming. Zooming out to a broader societal standpoint, improving disclosure systems and the

enforcement systems associated with them supports more ethical digital marketing practices and safeguards vulnerable audiences. Particularly with sponsorships creating an engagement penalty, unethical businesses and influencers are incentivized to hide the commercial status of sponsored content; there must be enforced rules in place to prevent that.

Overall, these results suggest that clear, standardized disclosures that are well enforced are vital to protect everyone from influencers to brands to viewers; creating a transparent ecosystem helps build trust and protect consumers.

5.5 Limitations

While this thesis conducted a sound investigation, there are some limitations that should be noted. Related to the data collection process, the fact that the data was collected by hand introduces human error and subjectivity. This strategy also resulted in a cross-sectional design that shows a snapshot of a short period of time, which is a key consideration in the rapidly evolving landscape of social media. Additionally, because the data collected focused only on one niche by targeting beauty influencers there is uncertainty regarding the extent to which the results can be generalized to other influencer types.

Lastly, the analysis was impacted by multicollinearity due to overlap in the clarity tiers. This ultimately contributed to an inconclusive result for RQ1 due to limited ability to draw statistically significant results related to the impact of different strengths of disclosure on engagement.

5.6 Future Research

There are a number of ways that this thesis could be built upon. Most directly related, the overlapping clarity tier categories made it difficult to isolate the effects of different disclosure strengths. By gathering more data or using experimental designs and focusing on data with mutually exclusive disclosure types, it could be more meaningfully clarified whether higher disclosure clarity influences consumer engagement. A longitudinal study would be another valuable option for expanding this research, as it could track how viewer trust and engagement evolve over time as influencers post more commercial content. A longitudinal study would also be interesting to compare how disclosures change as TikTok's rules and guidelines evolve.

Additionally, this research could be replicated beyond the beauty niche to determine whether the patterns observed in this thesis are consistent across other domains like food, gaming, fitness, travel, or any other popular niche. It is possible that different niches have different norms for disclosure and authenticity, so user expectations regarding commercial content may be different and some niches may have best practices that could be applied

elsewhere. In a similar vein, future research could stay within the beauty niche but expand to compare cross-platform differences to determine best practices for sponsorship disclosures across a range of social media types.

Finally, interviewing influencers or viewers could expand the work done in this thesis by adding a deeper qualitative understanding of how people interpret the less traditional pieces of commercial content, such as PR products and affiliate codes. This may help explain why affiliate codes were deemed more acceptable, boosting engagement while other commercial content lowers engagement: do viewers not understand the commercial nature of affiliate codes, or is there something about the way they are used that could be extrapolated onto other sponsorship forms?

Overall, any of these extensions would help to refine society's understanding of transparency, authenticity, and commercial intent in influencer marketing.

6 Conclusions

6.1 Answers to RQ1 and RQ2

The findings of this thesis provide answers to both research questions.

RQ1, which investigated how different TikTok ad disclosure styles affect consumer trust, used engagement as a proxy in line with previous research. The results showed a reliable reduction in engagement for sponsored posts, suggesting there is a trust penalty with traditional sponsorships. However, regarding non-traditional advertising strategies there were more mixed results. Explicit disclosures regarding free gifted PR products from brands did not significantly impact engagement, while ambiguous PR product disclosures reduced engagement and affiliate codes increased it. Although there was a directional pattern discovered toward stronger engagement penalties with clearer sponsorship disclosure, multicollinearity due to overlap in clarity tier categories mean that the effects were not statistically significant and results were inconclusive. Consequently, H1 was not supported as more explicit ad disclosure types did not lead to greater penalties in engagement.

RQ2, which investigated whether an influencer's typical tone and frequency as it related to reviewing products moderated the sponsorship penalty, ultimately found that the relationship between the terms was additive and not interactive with no statistically significant relationship. While having more positive reviews did lead to higher engagement in general, H2 was not supported.

6.2 Theoretical Contributions

This thesis expanded upon existing research and contributed to the theoretical understanding of influencer marketing, as well as digital persuasion as a whole, in several ways.

First, it extends the PKM into the context of TikTok by showing that recognition of commercial intent continues to trigger scepticism and reduce engagement even in informal, short-form video environments. Second, it applies PIT to the same context and to explain how sponsorship lowers engagement by disrupting the parasocial relationships that fuel influencers. Lastly, by incorporating the AMF this study has helped clarify how different commercial cues, both traditional ones like sponsorships and newer commercial cues like affiliate codes and PR products, map onto different dimensions of authenticity in practice. Implied PR disclosures appear to decrease the strategy of transparency dimension and violate viewers' expectations for influencers, while affiliate codes create value for the audience and align more closely with the "absolute authenticity" quadrant by

creating value. Together, the results of this thesis help to deepen the understanding of how commercial practices influence credibility and trust in the realm of TikTok and influencer marketing.

Overall, this thesis advances a number of key frameworks in the marketing and persuasion spheres. It demonstrates that different types of commercial signals lead to different reactions on TikTok, affirms that sponsorships do come with a trust penalty, and that an influencer's typical tone does not moderate the sponsorship penalty despite impacting engagement as a whole.

6.3 Final Remarks

Ad disclosures and how they should be treated is a highly relevant topic to everyone to maintain a transparent digital ecosystem that protects the consumer. Particularly as influencer marketing continues to grow and evolve, it will become increasingly important to expand upon this thesis as it has shown that different commercial cues on TikTok shape engagement and trust. Improving both best practices for ad disclosures and enforcing those practices will help promote an ethical online landscape, especially in sensitive spaces that are prone to affecting self-image. By providing empirical insight into these dynamics, this thesis has contributed to a more informed understanding of influencer advertising and its implications for both industry and society as a whole.

7 Disclosures

The author has no known affiliation with any of the brands or influencers involved with this study.

No text in this thesis was generated with AI, nor was the analysis done using AI. AI was only used as one of several methods for discovering sources for the literature review, troubleshooting issues on Excel, WEKA, and JASP, and organizing the author's notes for herself regarding desired thesis content and structure. All thoughts and analysis are those of the author.

References

Advertising Standards Authority. (n.d.). Advertising codes. https://www.asa.org.uk/codes-and-rulings/advertising-codes.html

Audrezet, A., de Kerviler, G., & Moulard, J. G. (2018). Authenticity under threat: When social media influencers need to go beyond self-presentation. https://www.sciencedirect.com/science/article/pii/S0148296318303229

Barari, M., Eisend, M., & Jain, S. P. (2025). A meta-analysis of the effectiveness of social media influencers: Mechanisms and moderation. Journal of the Academy of Marketing Science. https://link.springer.com/article/10.1007/s11747-025-01107-3

Byrne, B. M. (2010). Structural equation modeling with AMOS: Basic concepts, applications, and programming. New York, NY: Routledge.

Collins. (n.d.). For the culture. https://www.marctothec.com/fortheculture

Dandona, I., Shikha, S., Ashwini T. S., Dani, M., Sangeetha, P., & Patra, S. G. (2025). *The rise of influencer marketing: Consumer trust and brand engagement in the digital era*. Advances in Consumer Research, 2(4), 3054-3063.

Dupuis, R., Musicus, A. A., Edghill, B., et al. (2025). How TikTok influencers disclose food and beverage brand partnerships: Descriptive study. Journal of Medical Internet Research, 27(1), e60891. https://www.jmir.org/2025/1/e60891

Eisend, M. (2006). Source credibility dimensions in marketing communication: A generalized solution. https://www.researchgate.net/publication/228649951_Source_Credibility_Dimensions_in_Marketing_Communication-A_Generalized_Solution

European Commission. (2023). Influencer legal hub. https://commission.europa.eu/live-work-travel-eu/consumer-rights-and-complaints/influencer-legal-hub_en

Federal Trade Commission. (2019). Disclosures 101 for social media influencers. https://www.ftc.gov/business-guidance/resources/disclosures-101-social-media-influencers

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. https://www.jstor.org/stable/2489738

Gant, M. (2025). Representation of mature beauty influencers. https://www.byrdie.com/mature-beauty-influencers-interview-11704871

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis (7th ed.). Upper Saddle River, NJ: Pearson Educational International.

Horton, D., & Wohl, R. (1956). Mass communication and para-social interaction. https://www.tandfonline.com/doi/abs/10.1080/00332747.1956.11023049

Jin, S. V., & Ryu, E. (2021). "I trust what she's #endorsing on Instagram": Moderating effects of parasocial interaction and social presence in fashion influencer marketing. Journal of Fashion Marketing and Management. https://www.emerald.com/insight/content/doi/10.1108/JFMM-04-2020-0059/full/html

Ki, C.-W. (Chris), Cuevas, L. M., Chong, S. M., & Lim, H. (2020). Influencer marketing: Social media influencers as human brands attaching to followers and yielding positive marketing results by fulfilling needs. Journal of Retailing and Consumer Services, 55, 102133. https://ideas.repec.org/a/eee/joreco/v55y2020ics0969698920300059.html

Konovalova, V. (2024). How social media platforms fuel extreme opinions and hate speech. https://www.wbs.ac.uk/news/how-social-media-platforms-fuel-extreme-opinions-and-hate-speech/

Krzymowski, M. (2024). The link between social media and body image issues among youth in the United States. https://ballardbrief.byu.edu/issue-briefs/the-link-between-social-media-and-body-image-issues-among-youth-in-the-united-states

Kuras, A. (2024). What's inside a PR package? https://prowly.com/magazine/pr-package/

Library of Congress. (n.d.). Influencer marketing metrics and costs. https://guides.loc.gov/influencer-marketing/metrics-and-costs

McCue, T. J. (2019). Verizon Media says 69 percent of consumers watch video with sound off. https://www.forbes.com/sites/tjmccue/2019/07/31/verizon-media-says-69-percent-of-consumers-watching-video-with-sound-off/

Mental Health Foundation. (2019). Body image report: Executive summary. https://www.mentalhealth.org.uk/explore-mental-health/articles/body-image-report-executive-summary

Molenaar, A. (2025). Influencer niches. https://influencermarketinghub.com/influencer-niches/

MSL Staff. (2023). Influencer equity: Best practices and data. https://mslgroup.com/news/influencer-equity-best-practices-da-ta

Pan, M., Blut, M., Ghiassaleh, A., & Lee, Z. W. Y. (2024). Influencer marketing effectiveness: A meta-analytic review. Journal of the Academy of Marketing Science. https://link.springer.com/article/10.1007/s11747-024-01052-7

Salesforce. (2024). Affiliate marketing overview. https://www.salesforce.com/au/blog/affiliate-market-ing/

Sharethrough. (2021). 75 percent of people watch mobile videos on mute. https://digiday.com/sponsored/75-percent-of-people-watch-mobile-videos-on-mute/

Tafradzhiyski, S. (2025). TikTok statistics. https://www.businessofapps.com/data/tik-tok-statistics/

Thorpe, E. (2023). 7 stats that show women dominate influencer marketing. https://archive.fohr.co/articles/7-stats-that-show-women-dominate-influencer-market-ing

Tian, Y., et al. (2025). The impact of TikTok social media influencers on consumer behavior: A systematic literature review. https://www.researchgate.net/publication/395085885_The_Impact_of_TikTok_Social_Media_Influencers_on_Consumer_Behavior_A_Systematic_Literature_Review

TikTok. (n.d.). Accessibility. https://support.tiktok.com/en/using-tiktok/creating-videos/accessibility

TikTok. (n.d.). Camera tools. https://support.tiktok.com/en/using-tiktok/creating-videos/camera-tools

TikTok. (n.d.). Promoting a brand, product or service. https://support.tiktok.com/en/business-and-creator/creator-and-business-accounts/promoting-a-brand-product-or-service

Yltava, R. (2023). Main products sold by influencers (U.S.). https://www.statista.com/statistics/1426507/main-products-sold-by-influencers-us/