THE ROLE OF PERCEIVED MOTIVATION AND ABILITY IN WARRANTING VALUE ASSESSMENTS AND IMPRESSION FORMATION: A COMPARISON OF WARRANTING THEORY AND TRUTH DEFAULT THEORY PREDICTIONS

By

Kelsey Earle

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Communication– Doctor of Philosophy

2023

ABSTRACT

Two studies were conducted to examine the hypothesized effects of both warranting theory and truth default theory in an online dating context. The variables of interest were compared using bivariate correlations and competing serial mediation models were presented. The 2x2 experiments examined the effects of perceived motivation and ability to manipulate. Perceived motivation and ability to manipulate were significantly correlated with warranting value. However, warranting value was not correlated with the impression outcomes (trustworthiness, content veracity, and warmth). Neither the warranting theory nor the truth default theory serial mediation models were significantly predicted by the data. This was partially explained by a failed manipulation of perceived motivation. Indirect effects of ability on impressions that excluded warranting value were found in Study 2. Together, the studies provide a first look into these relationships and exciting new avenues for future research.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

TDT Truth default theory

FtF Face-to-face

CMC Computer-mediated communication

PATM – Perceived ability to manipulate

PMTM – Perceived motivation to manipulate

Introduction

As the Internet continues to grow and change, people are becoming increasingly concerned with how to decipher truth from fiction. An important area of this research focuses on how online cues are interpreted in order to form impressions and make decisions about the content and the content provider. Indeed, these messages themselves contain clues that help people unravel online misinformation and deception. For example, social information processing theory (Walther, 1992) identified people's resourcefulness in cue utilization in forming interpersonal relationships. If individuals are given enough time, they are able to form bonds with others using mediated channels of communication, similar to bonds formed through face-to-face communication.

Research in this space has continued to grow and has emphasized how these processes can be used for impression formation and decision making in the modern Internet environment. Online dating is a clear example of how online relationships may develop and change into in-person relationships. Impression formation is a pivotal part of this process. Individuals must interpret and evaluate potential partners' profiles in order to find a compatible match. The perceived veracity of this information becomes incredibly important for users who are specifically seeking relationships. The daunting task of sifting through matches to find a compatible partner becomes even more challenging due to the proliferation of fake accounts, bots, scams, and even artificial intelligence (AI) generated information (Tao, 2022; Wu & Kelly, 2020). It becomes difficult to determine who is presenting themselves accurately when such misrepresentations are prevalent. The cues within the profiles may be the key for understanding how these users make veracity assessments and ultimately decide to match (or not match) with the other user. However, this begs the question – what online cues are useful in determining

perceived manipulability and trust? In other words, what cues help us determine what is true or false on the Internet in order to help us form impressions or relationships with other people?

Many different cues have been examined as part of these processes and several theories have gained traction as explanatory mechanisms in online deception detection and relationship formation. This paper uses two such theories, truth default theory (Levine, 2014; 2020) and warranting theory (DeAndrea, 2014; Walther & Parks, 2002) are used to understand how perceptions of manipulability drive online information processing and impression formation. Two specific cues that have been used, yet not concretely identified, are proposed as influential impression formation cues. Perceptions of another person's ability and motivation to manipulate online content are the mechanisms of interest. An online experiment was conducted to examine the relationships these factors have on perceived manipulability, trustworthiness, veracity, and affective impressions.

Individuals seeking connection through online dating platforms must decipher truth from fiction in order to form impressions and make matches, so understanding this process is exemplified in this context, making it ripe for study. A two-part study is presented. Study 1 utilized a college student sample and examined the study's variables within the context of an online dating website for college students. Lessons learned from Study 1 informed a new manipulation and dating profile for Study 2. Study 2 utilized an online sample of adults from across the United States and was presented as a new online dating website for adults. The studies are considered both separately and together to make sense of the various theoretical mechanisms. First, a review of the literature provides insights into the predicted relationships.

Truth Default Theory

Truth default theory (TDT; Levine, 2014) proposes that people automatically and unconsciously assume others are telling the truth, unless they have a reason to believe otherwise (Levine, 2014). A truth-default state is an adaptive feature of human communication as it allows people to reduce cognitive load, enables more efficient communication, and encourages cooperation (Levine, 2014; 2020). It simply takes less effort and makes things easier when people believe others are telling them the truth. Truth default theory describes lie expectations, lie presumptions, and truth and lie judgements while also specifying the conditions wherein deceptive or honest communication is likely to occur (Levine, 2014). Importantly, TDT is focused on credibility assessment and truth judgments rather than in discourse production (as seen in theories such as information manipulation theory 2; McCornack et al., 2014). As such, TDT has the potential to explain and predict online impression formation. We know not everything on the Internet is true, yet people still get taken advantage of and lied to. The truth default state may shed light on the manner in which this occurs.

Levine (2014; 2020) uses their framework to make several propositions about the truth-default state, how it is triggered, who lies, and when lies can be detected. Relevant to this paper are propositions six through nine. For full propositions see Figure 1. Essentially the propositions of interest propose the manner in which the truth default state occurs and when/how individuals abandon it. People understand that deception serves a purpose and ascribe motivations to others which helps them decide if someone is being dis/honest (Proposition six). A trigger event is needed for people to be knocked out of the truth-default state (Proposition seven). If the trigger event is strong enough, suspicion will be induced, and individuals will more carefully analyze the information (Proposition eight). There are a variety of places people can look to assess

dis/honesty. Depending on what is available or how they assess the information, they may judge the content to be dis/honest. They may even revert to the truth default state (Proposition nine; Levine, 2020).

Number	Proposition
6	"People understand that others' deception is usually purposive and are more likely to consider a message as potentially or actually deceptive under conditions where the truth may be inconsistent with a communicator's desired outcomes. That is, people project motive states on others, and this affects suspicion and judgments of honesty and deceit."
7	"The truth-default state requires a trigger event to abandon it. Trigger events include but are not limited to: (a) a projective motive for deception, (b) behavioral displays associated with dishonest demeanor, (c) a lack of coherence in message content, (d) a lack of correspondence between communication content and some knowledge of reality, or (e) information from a third-party warning of potential deception"
8	"If a trigger or set of triggers is sufficiently potent, a threshold is crossed, suspicion is generated, the truth-default is at least temporarily abandoned, the communication is scrutinized, and evidence is cognitively retrieved and/or sought to assess honesty-deceit."
9	"Based on information of a variety of types, an evidentiary threshold may be crossed, and a message may be actively judged to be deceptive. The information used to assess honesty and deceit includes but is not limited to: (a) contextualized communication content and motive, (b) sender demeanor, (c) information from third parties, (d) communication coherence, and (e) correspondence information. If the evidentiary threshold for a lie judgment is not crossed, an individual may continue to harbor suspicion or revert to the truth-default. If exculpatory evidence emerges, active judgments of honesty are made."

Figure 1. Truth default theory propositions (Levine, 2020, p. 98-99)

When someone is suspicious, they may regard the content with more scrutiny and spend more time determining the veracity of the content. As such, they are predicted to move away from the truth default. Certain cues and contexts may influence how suspicion is induced (Levine, 2020). For example, the ability for audience engagement (e.g., commenting on a review, liking a social media post, etc.) has been shown to produce more honest self-presentations (Toma et al., 2008; DeAndrea, 2014). If a comment is deleted by the poster, it removes the ability for audience engagement which in turn may elicit suspicion about the content regarding the reason

for the deletion. Indeed, Shin et al., (2020) found that justification for potentially deceiving acts altered impressions. Individuals were less concerned when the justification for the action seemed appropriate, even if it benefited the poster. Suspicion was alleviated when proper justification was given. Of course, while suspicion is a helpful tool for evaluating content, it is by no means a perfect indicator for evaluating the truthfulness of said content.

In the present study, trustworthiness and veracity are used as distinct, yet likely related, concepts. First, we emphasize that *perceived* trustworthiness and veracity are examined. This relies on individuals' reports of their attitudes and beliefs about these concepts. Additionally, there are conceptual differences between the two terms. In this study, perceived *trustworthiness* and *content veracity* are used. Perceived trustworthiness refers to the perceptions one has that a *person* is of good moral standing, is honest, and the statements they make can be relied on (Rotter, 1980). Trustworthiness is often utilized as a form of source credibility (McCroskey & Teven, 1999). Perceived content veracity on the other hand refers to the accuracy and correctness of the *content* provided by a person (Lozano et al., 2020). Indeed, the veracity effect (Levine et al., 1999) explains how and why people evaluate the accuracy of deceptive and truthful messages. The principle of veracity takes this a step further and identifies that perceptions of veracity often rest on the moral obligations of truth and deceptive judgements (Levine et al., 2010). As the authors describe, truth does not often require explanation, but deceit does. Content veracity then, refers to how truthful the content is believed to be, such that it requires no additional explanation. To this end, perceptions of trustworthiness of a source are often used in judging the content's veracity that the source provides.

The nature of trustworthiness and veracity can be difficult to unravel. Indeed, most people believe they can use perceptions of trustworthiness as indicators of content veracity. In

short, when people are prompted to make truth judgements (i.e., asked if their communication partner is lying or telling the truth), they are correct about 54% of the time (Bond & DePaulo, 2006). Interestingly, when unprompted, individuals often do not report on perceived trustworthiness, suspicion, or veracity; supporting the claims of TDT (Clare & Levine, 2019; Levine et al., 2020). As TDT proposes, people tend to default to the truth and simply do not consider the potential falsehoods of the given information unless prompted to do so. Once prompted, people are not good at detecting lies and truths. Suspicion alone is often not enough for individuals to make deceptive attributions regarding other people's attitudes or behaviors. Levine (2020) proposes that there is a second threshold that needs to be passed to move from active questioning to belief of deception. This means that unless strongly prompted, most people will assume the truth. Even if a triggering event is presented to make an individual reanalyze their trust in the content, they are incorrect in their truth judgement about half of the time.

However, Levine (2014) suggests that suspicion can be induced by diagnostically relevant information. This type of information may be more helpful in deciphering potential deception than the cues traditionally suggested by deception detection research (e.g., eye contact, hand gestures, or other nonverbal behaviors). People know that lies are told for a reason. Suspicion about deception will be present when there is perceived reason to lie (Levine et al., 2010; Levine, 2014). Therefore, diagnostic utility, or the extent to which information can be used to form a judgement or attitude (Levine, Blair, & Clare, 2014) is important to the formation of truth judgements. If information is diagnostically useful to truth assessment, then it can help the viewer assess the veracity of the information more easily than when the information is not diagnostically useful. Correspondent information (information that lines up with empirical facts often coming from other trusted sources) is highly diagnostically useful in deception detection

(Blair et al., 2018). Claims that can be tested or verified are more useful in making assessments. Blair et al., (2010) found that contextual information influences deception detection accuracy in mock crime scenarios. For example, information about a crime from forensic scientists was diagnostically useful when making deception judgements as it was verifiable information from an external source. As the external source was independent from the case, their information was deemed particularly trustworthy. In other words, the scientist was not skewed by things like emotions or involvement in the case, while a witness for example, may be biased by their involvement.

Deception detection and truth judgements are particularly important for individuals who are using cues as impression formation mechanisms. People adapt their communication by way of online cues to suit their interpersonal goals (Tidwell & Walther, 2002; Walther & Parks, 2002). Relevant diagnostic and correspondent information is used to make decisions about others, especially when communicating through mediated channels. For example, if two people are chatting online in order to get to know each other, they are likely to engage in more socially relevant question and answer behaviors. The communication behaviors they enact help them achieve their goal (in this case, interpersonal relationships). Additionally, people will use different communication strategies depending on the goals and the potential for relationship development. Ramirez (2007) found that anticipation of future interaction facilitated relationships, even over time. Individuals who knew they would interact in the future were more likely to ask more questions to get to know their communication partner. Their communication strategies differed from individuals who simply were paired to complete a task. In day-to-day life, there are instances where people are not encouraged, nor look for, social bonds. For example, when individuals seek out online reviews to make purchasing decisions, they likely do

not expect to form bonds with the review writers. Instead, they rely on different perceptions of the reviewer (by way of online cues) that help them make impressions and ultimately decisions (e.g., DeAndrea et al., 2018). Based on the goals of the individuals and the types of information available, certain types of information may be more or less useful. Online dating may see both of the effects seen in Ramirez's work on anticipated future interaction and of DeAndrea and colleagues' work regarding online reviews. Individuals are likely using the online dating website to form relationships, but they first have to sort through potential matches. The types of information imbued in the content may be helpful in deciphering potential deceptions and ultimately individuals' decisions to match or communicate with potential partners.

Correspondent information from other sources is a powerful diagnostically useful tool in online impression formation and deception detection (Blair et al., 2018; Levine, 2020). Depending on the goals and objectives of the individual, they may seek out alternative sources of information (i.e., correspondent information) to verify or solidify a claim. For example, a person might look over an online dating match's social media to confirm their identity or a person searching online for a dinner reservation may check a review website after viewing a menu to ensure the food will be good. The search for correspondent information aids individuals in making informed decisions about their impressions of other people and future behaviors.

In this way, there are certain contexts in which individuals are more likely to break their truth default via information seeking and impression formation mechanisms. By seeking out cues that help them ascertain the veracity of the presented content, individuals have more information from which to form their impressions. But what mechanisms drive information seeking, especially correspondent information seeking? Suspicion has been the main emphasis of study in this regard. People may break the truth default state when they are sufficiently suspicious and

start to look for corroborating or conflicting information about the claim or person of interest. However, as Levine (2020) notes, there are certainly other mechanisms that can knock people out of their truth default response. Perceptions of manipulability, by way of perceived warranting value, is one such mechanism that will be explored here.

Warranting Theory

Warranting theory (Walther & Parks, 2002) highlights the role of perceived manipulability in making assessments of online content and was introduced to help explain and predict impression formation from online self-presentations. A warrant is “any cue that authenticates or legitimizes an online self-presentation" (DeAndrea, 2014, p.187). When people evaluate online content and self-presentations, they assess its *warranting value* or “perceptions about the extent to which information is immune to manipulation by the source it describes” (DeAndrea, 2014, p. 186). In other words, people evaluate the likelihood that the content they are viewing has been altered or manipulated in some way and ascribe warranting value to that content based on their perceptions. The warranting principle, a central tenet of warranting theory, posits that perceived manipulability of online content carries substantial weight in impression formation (DeAndrea, 2014; Walther, et al., 2009). Warranting value has been proposed and utilized as having a moderating (DeAndrea et al., 2015) or mediating (DeAndrea & Carpenter, 2018) effect on impressions. If a cue has high warranting value (it is perceived to have low manipulability), the more likely individuals will believe the content to be true thus utilizing the information in impression formation or decision making. This has implications in a variety of settings where impression formation is an objective such as online dating (Toma et al., 2008) or when evaluating online content such as a social media post (Vendemia et al., 2018). **Taken together, warranting value is predicted by the warranting principle, warranting cues act as indicators of warranting value, and warranting value is predicted to influence the relationship between information and impressions of online content (Walther & Parks, 2002; DeAndrea 2014).**

Taking warranting theory into consideration, warranting cues can have their own diagnosticity wherein harder to alter information is more diagnostic (more useful as a manipulability cue) than easy to alter information (less useful as a manipulability cue). Online content that is difficult to modify (i.e., it is not easily falsified) should meet the correspondence criterion and lead individuals to assess low levels of perceived manipulability (warranting value) and be more likely to form accurate truth or deception judgements.

As an example, in the Walther et al., (2009) study, statements of extroversion or introversion and attractiveness were manipulated on a social media profile. Perceived manipulability did not exert a strong influence on perceptions of extroversion and introversion. It appears that people do not view introversion and extroversion claims as being extremely important to impression formation. In other words, it may not be diagnostically useful information as there is little to be gained or lost by expressing this characteristic. On the other hand, claims of attractiveness were more favorably received from a third party than from a self-claim. Correspondent information about attractiveness from a third party was perceived as more helpful in impression formation than when that same information was received from the person themselves. As such, third party claims were more diagnostically useful than first-party claims regarding attractiveness, aligning with the tenets of TDT.

Several elements of warranting theory emphasize the utility of correspondent information in impression formation. Walther and Parks (2002) identified the differences between impression formation cues in face-to-face, computer-mediated, and mixed communication models specifically in regard to perceived manipulability. The authors speculated that a link between one's corporeal self and one's self-presentation online could influence perceived warranting value. It is important to note that they primarily examined the literature regarding relationship

formation online, particularly when that relationship was expected to move offline at some point. Warranting then, could be derived from perceptions of the link between the online self and the offline person. They outlined several possible reasons for these occurrences including anticipation of future interaction, corroboration via other sources such as webpages and networks, and a lack of anonymity. Many of these features are present in an online dating context and may provide valuable insight into these relationships.

Anticipation of future interaction was predicted to act as a warranting cue. When people expect to have future conversations with their communication partner, they are more likely to share relational information (Ramirez, 2007; Walther, 1994). If they lie about themselves online, and then meet in person, they risk exposing that lie. As an example, it is fairly easy to manipulate or conceal one's weight on an online dating site, however this information is easy to verify upon meeting in person (Toma et al., 2008). This is not to say people do not lie on their online dating profiles. As seen in online dating research (e.g., Ellison et al., 2006; Toma et al., 2008) people do in fact lie or exaggerate things about themselves in an effort to look as attractive as possible. Walther (1994) utilized anticipation of future interaction and found that people generally formed more positive views of their group members when told to anticipate future interactions than those who did not anticipate further communication. Ramirez (2007) advanced this idea and considered the role of first impressions and impressions over time. Importantly, people need to be motivated to form and maintain these relationships. This motivation seems to drive perceptions of manipulability such that when future interactions are expected, individuals are expected to lie less frequently.

Similarly, information that can be corroborated or verified by *other sources* has also been identified as a potential influencer of warranting value. Walther and Parks (2002) explained that

links in a social network may be a source of information verification. For example, recommendations for a job may have more or less warranting value depending on who they come from. One study found that online recommendations about a person's skills (via the social media LinkedIn) were perceived as having higher warranting value, higher credibility, and viewed more positively when coming from a previous supervisor than coming from a previous subordinate (Rui, 2017). Information coming from a friend was also perceived as having more warranting value than when coming from the source themselves (Walther et al., 2009). Put differently, information about a person on its own is easily manipulable, but since a known other can verify this information, it is perceived to be less manipulable. A relevant example of this may be when individuals are set up on a blind date by a shared contact. While the daters themselves do not know each other, they can verify information about their date with members of their network. In this way, the daters likely are more willing to trust each other and the information provided by the shared source, as it can be corroborated within their network.

People are not the only source of corroboration or verification. Social media profiles and websites also provide useful information in this regard. Cues such as number of friends are generally system-generated and thus users do not have the ability to manipulate them (Tong et al., 2008; Walther & Jang, 2012). System generated cues should have high warranting value due to their verifiability and lack of manipulability by the individual to whom they are referring. In a series of studies examining warranting effects on photographs in online auctions (Johnson et al., 2015; Van Der Heide, et al., 2013), researchers compared user-generated and stock images in terms of number and price of auction sales. A user-generated photograph was found to have more warranting value than the stock image and led to higher and more confirmed sales.

Combined with a positive review score (a system generated cue) photographic warranting cues were shown to be an important part of purchasing decisions.

Online content offers some anonymity and thus we may resort to other types of warranting cues in our information processing and impression formation. Walther and Parks (2002) speculated that people are less likely to manipulate content when they cannot be anonymous. In this case, non-anonymous information can be linked to one's corporeal self. If anonymity is not an option, then other people have the ability to verify or corroborate the information. In the case of anonymous online reviews, where information about the poster is unavailable, people may resort to using other cues to decern warranting value and veracity (Ansari et al., 2018; DeAndrea et al., 2018). Online editing on wiki pages is a form of modification control that has been studied regarding warranting theory (DeAndrea & Carpenter, 2018). Wikis allow anyone to modify and change the content anonymously, therefore it is predicted that people will perceive the content to have lower warranting value. Research on anonymous sperm donors has examined how individuals signal credibility in their online profiles while still maintaining anonymity. In this case, they are leveraging different self-presentations to overcome the potential loss in warranting value that would have been provided by being identified (Bokek-Cohen, 2014). In sum, people perceive non-anonymous content as having higher warranting value than anonymous content but may seek out other cues to help form impressions. People who recognize that anonymity results in a loss of warranting value may also leverage other cues in an attempt to encourage more positive impressions, trust, or other goal-oriented objectives.

The aforementioned studies illustrate that perceived warranting value can be influenced by things like verifiability and corroboration, particularly when individuals intend to take online

relationships offline. However, "... the term warrant has been broadened to refer to any cue that authenticates or legitimizes an online self-presentation" (DeAndrea, 2014, p. 187). As an example, a person may look to online reviews before making a purchasing decision. They are not expecting to form any sort of relationship with the reviewer. Instead, they are forming impressions of the reviewer and of review content to make a decision. Perceived warranting value is predicted to influence this relationship wherein high perceived warranting value leads to more favorable impressions (DeAndrea, 2014).

As such, the more control a person has over the content, the more they are perceived to be able to manipulate the content. This decreases the perceived warranting value of the content which in turn influences impression formation. DeAndrea and Carpenter (2018) examined three types of information control: dissemination control, modification control, and source obfuscation. Alongside general warranting, the authors found that different types of information control could influence perceptions of manipulability. Modification control refers to the perceived ability a person has to modify or change content about themselves that is from a third-person (e.g., information on a wiki). Dissemination control is the extent to which someone is perceived to have control over the dissemination or presentation of online content (e.g., a person can pick what is shared online). Source obfuscation is the extent to which observers believe the true source of the content can be hidden or obscured (e.g., a person pretends to be someone else while posting; DeAndrea & Carpenter, 2018). Taken together, general warranting value, dissemination control, modification control, and source obfuscation have been proposed as ways to examine perceptions of manipulability in online content, particularly when impression formation is the end goal. DeAndrea and Carpenter's (2018) scale development has highlighted the dimensionality of perceived warranting value and provides a way to capture it.

To this point, impression formation has been discussed somewhat vaguely and broadly. Many of the predictions made by Walther and Parks (2002) allowed for a somewhat general interpretation of impressions. For example, one may use an impression such as perceived work ethic when viewing a resume, but work ethic is less important in cases such as online dating where interpersonal impressions such as attraction may be the goal. As such, this study emphasizes interpersonal impressions relevant to the experiment as the outcome of perceived warranting. Snyder et al., (1977) developed a measure of interpersonal impressions and used it to understand the relationship between physical attractiveness and impressions. Work from Snyder et al., (1977) has been instrumental in computer-mediated communication research involving behavioral confirmation. Other scholars of communication utilize other affective interpersonal impression scales such as the generalized attitude measure (McCroskey & Richmond, 1989) and Burgoon et al., 1978 measure of attitude used in persuasion research. Taken together, these scales of interpersonal attraction and attitude are helpful metrics for determining a person's affective warmth which is one potentially important outcome of perceived warranting value.

Most of the theory building surrounding the original formulation of warranting theory was centered around warrants between online and in-person attributes, while recent research has focused on any warrants that legitimize online self-presentations. One such application of the theory has emphasized how warranting value influences the impressions of user-generated content such as online reviews. In online reviews, individuals are specifically using the cues to make decisions both about the reviewer and the item they are reviewing. In this way, perceptions of manipulability become important. For example, positive reviews about a business on a third-party website (e.g., Google, Yelp, etc.) are more highly regarded than positive reviews on the business' own website (DeAndrea et al., 2018). In an online environment where reviews are

becoming commonplace, understanding how they are evaluated is critical. As such, the next section provides an overview of one of the most studied warranting theory claims regarding first- and third-person effects.

First- and Third-Person Effects

The assessment of warranting value in an online setting has been frequently studied within the context of online reviews. Early work identified the variance in perceived warranting between first and third person claims (Walther & Parks, 2002; Walther et al., 2009). A first-person claim (or a self-claim) includes statements made by a person about themselves (e.g., using pronouns such as I and me). A third-person claim (or other-claim) includes statements made by a person about others (e.g., using pronouns such as he/she/they or proper nouns such as a person's name). Per the warranting principle (Walther & Parks, 2002), people rely on the perception of manipulability when evaluating this type of online content.

It has been predicted that other-generated content provides more warranting value because it is more immune to manipulation from the source it describes. In other words, a third-person account has less chance of being altered or modified by the person or object being discussed. Third-party information is more diagnostically useful because it is coming from a different source that cannot (or at least should not) be able to be manipulated by the source (Blaire et al., 2018; DeAndrea, 2014). Impressions of social media profiles have been utilized as one way to examine first- and third-person warranting effects. Young adults form more positive impressions of an online social media profile when other people write comments and engage with the profile (Antheunis & Schouten, 2011). Utz (2010) also examined competing claims' impact on interpersonal impressions. They found that other-generated statements influenced popularity judgements, communal orientation, and social attraction more strongly than self-generated statements. Overall, first- and third-person effects of warranting have been an influential and interesting field of study. This is particularly true due to the proliferation of and dependence on these types of messages in online spaces. However, it is important to note that

this is only one way warranting value can be examined. It is discussed thoroughly here as it is relevant to truth and deception judgements in online spaces.

In summary, online content is rife with cues capable of helping individuals form impressions of perceived manipulability. In turn, these warranting judgements are theorized to influence impressions of the content and content provider. Much of this research identifies that third person information is perceived to have more warranting value, and thus is more likely to be perceived as accurate, than first-person information. Individuals tend to place a higher emphasis on third-person reports as compared to a first-person testimonial when forming impressions. While this argument has been upheld in a variety of studies, there are a several other factors that alter the way first- and third-person information is evaluated.

First- and Third-Person Inconsistencies

The warranting principle insinuates that first- and third-person online claims are influential and informative sources of information (Walther & Parks, 2002: Walther et al., 2009). However, previous work examining the first- and third-person effect have reported mixed findings. Indeed, research concluded that the difference in warranting value for self- and other-generated information is more complex than originally proposed (DeAndrea, 2014). A variety of studies have tested the underlying mechanisms for which first- and third-person warranting effects may emerge.

Walther and Parks (2002) proposed that there would be differences in perceived warranting value between first- and third- person accounts, particularly due to perceptions of what was “at stake” for the content poster. As an example, a restauranteur would likely write positive things about their restaurant in the first-person while a customer would be likely to write an honest review in the third-person regardless of their view (positive or negative). The

restauranteur wants to keep their business thriving, while a customer wants to help others make dining decisions. But this view neglects a third person review with factors such as potential self interest in the success of the restaurant, for example a competing business owner. In an online dating situation, individuals must make self-claims in their profile in order to attract potential matches. In this way, there is a lot "at stake" for the profile owner. Therefore, the types of claims and their warranting value matter far more than first- or third-person claims. This is especially true, as most online dating websites do not have any third-party communication available.

Previous research has noted that different types of information are more or less immune to first-and third-person effects of warranting value. Walther et al., (2009) concluded that people are less influenced by other's claims of introversion or extroversion than claims of attractiveness. Put simply, claims of introversion and extroversion regardless of who makes the claim are not particularly helpful in impression formation. However, claims of attractiveness are more influential coming from a third party than from the source. This study helped to set boundary conditions for first and third person effects by illustrating the utility of perceived importance of the content. The authors state, "Warranting data should be most useful to observers when discrepancies are great and where they are suspected to result from distortion or deception emanating from a target's motivation to appear more desirable" (Walther et al., 2009, p. 243). As mentioned previously, correspondent information is helpful in forming impressions, but correspondent information from a non-invested individual can be an influential and trustworthy source of information. Being able to "verify" information via correspondent information can be an important process in truth or deception judgements, particularly when individuals have more at stake in the reception of their claims.

To this end, DeAndrea and Vendemia (2019) explored the influence of perceived self-interest as a potential moderating variable. They found that particularly for positive reviews, perceived self-interest impacted impressions. When people perceived that the reviewer had a lot at stake by writing their review (high self-interest) people reported lower levels of trust and higher perceived manipulability on the review (when compared to low perceived self-interest). These findings may explain some of the seemingly inconsistent results regarding first person and third person warranting effects (e.g., DeAndrea et al., 2015; Walther et al., 2009, Utz, 2010).

There may be cues that subsume warranting information and influence impression formation in unexpected ways (DeAndrea, 2014). Indeed, Neo and Johnson (2020) challenged the idea that a third-person review will always be perceived as less manipulable than a first-person review. They contend that previous studies of warranting focused on one potential first-person source rather than other external third-person sources manipulating the information. As an example, previous studies recognized source obfuscation as a restaurant owner posing as a customer (i.e., fabricating an other-generated statement; Ansari et al., 2018). DeAndrea et al (2018) also note that high uncertainty regarding the third-person reviewer's identity is associated with lower levels of trust. However, it is unspecified by traditional warranting research, as to what would happen if a friend of the restaurant owner posted a review. They likely have an interest in the success of the restaurant and the information may not necessarily be truthful, but the correspondent information (from a third-person perspective) insinuates that it is relevant and trustworthy information.

Interestingly, there are times in which researchers have found such an effect. In online reviews about a politically (conservative or democratic) themed book, people perceived more manipulation when an in-party (i.e., matched the participant's political ideology) book was rated

negatively (Neo & Johnson, 2020). Even though the review is technically a third-person account, it was perceived as having low warranting value. Put differently, perceived manipulation (general warranting value) mediated the relationship between counter-attitudinal ratings (by others) and positive party affect (of the participant). The authors suggest that in this instance a third-party review is not as unbiased as once believed (Neo & Johnson, 2020). Indeed, even Walther and Parks (2002) noted, "Actors can undertake behaviors that may increase the apparent warrant of their performance for the receiver's benefit, through the presentation of warranting information" (p. 552).

An important consideration in this argument is that from the receiver's point of view, there must be a reason to suspect third-party manipulation. DeAndrea et al., (2018) explored the relationship between content deletion and doubt. They found that overall, people expressed more doubt about the website when content could be deleted. While people still reported high levels of doubt regardless of the condition they viewed in the experiment, the highest levels were reported when they were explicitly informed of the potential content deletion occurring on the website. The participants also reported that they would like to know whether or not content deletion could occur and to know the content providers' true identities (DeAndrea et al., 2018). As can be seen from this study, when explicitly told of potential manipulation, participants are likely to be more mindful of its effects.

Taken together, the boundary conditions of warranting theory are muddled and somewhat difficult to interpret. The mechanisms driving perceptions of manipulability are driven by a variety of individual and situational factors. However, an important conclusion can be drawn from this research: perceived warranting value matters in online impression formation. One thing these studies have in common is that the manner in which a receiver interprets the content is

important to perceptions of manipulability. That is, diagnostically important information in terms of the motivation, ability, and behaviors ascribed to the message sender influence impression formation and content evaluation. However, this specific relationship has yet to be explored in this way.

Diagnostically Useful Cues: Ability and Motivation

There are several factors that can influence how a person evaluates and forms attitudes about online content and the individuals who created the content. Research has identified impressions of a source matter in evaluating communication-related outcomes. For example, factors such as source credibility and attractiveness (e.g., Hovland & Weiss, 1951) have been shown to influence perceptions of communication effectiveness. Further, perceived source credibility can be determined via online sources and cues. For example, the recency of a social media post was shown to impact source credibility (Westerman et al., 2014). Less research has emphasized how and why and how people manipulate, alter, or control online cues to intentionally achieve their communication goals. As an example, individuals may selectively self-present when online, which could be considered a form of deception or misinformation. A person may lie about their body weight and hobbies on their online dating profile in order to get matches (Toma et al., 2008). Even though this information is inaccurate, the reasons for doing so are clear. People have the ability to manipulate the content (i.e., lie about their weight) and they are motivated to do so (i.e., they want to be seen as fit and attractive to potential matches).

Perceptions of warranting value are crucial to this process. As described by warranting theory, perceived manipulability influences message processing and impression formation. Certain factors (e.g., system generated cues, justification, self-interest, etc.) have been used to examine what influences a person's perceived manipulability. However, many of these ideas rest on the fact that the message sender has a reason (e.g., motivation) or the resources (e.g., ability) to alter the content in some capacity but few studies have specifically targeted this idea concretely.

Walther et al., (2009) found that people were less inclined to make decisions about claims of introversion and extraversion than claims of attractiveness while evaluating profiles and the authors identified the potential role of perceived sender motivation as a factor. As in the case of Neo and Johnson's (2020) study, perceived manipulability mattered more when a third-party posted a negative review against the participant's preferred party. These two examples highlight that it is not simply just the presence or absence of perceived manipulation that influences impression formation. It is proposed here that the perceptions of a message sender's *ability* and *motivation* to manipulate content are the driving factors influencing impression formation.

To this end, we first turn our attention to the manner in which warranting value is assessed within the empirical research done on this topic. Warranting value "reflects perceptions about the extent to which information is immune to manipulation by the source it describes" (DeAndrea, 2014, p. 187). DeAndrea later points out, "empirical evidence is needed that documents the salience perceptions of warranting value have during the evaluation of information online" (p. 200). The *salience* of warranting value perceptions as a determinant of impression formation is central to the argument thus far.

Given the importance of salience in these relationships, we turn to two potential drivers of salience: perceived ability and perceived motivation to manipulate. Perceived ability to manipulate (PATM) refers to the perceptions an individual has that a content provider can alter, change, or modify the content in some way. As seen in DeAndrea and Carpenter's (2018) work, information control is a central construct here. Perceived motivation to manipulate (PMTM) refers to the perceptions an individual has that a content provider has a reason to alter content to achieve some goal. For example, perceived self-interest has been shown to function in this manner (DeAndrea & Vendemia, 2019).

Both perceived ability and perceived motivation to manipulate hinge on the causal nature of evaluating another person's actions, thoughts, behaviors, or beliefs. Correspondent information is expected to be diagnostically useful in impression formation and deception detection. Simply put, individuals use the available cues to form opinions of trust or deception. The warranting principle also gets at the causal nature of these relationships such that perceived warranting value influences impression formation. In other words, perceptions of manipulability have an influence on the attitudes people form about others. In both cases, individuals are using the provided information to make causal judgements about another person's behavior.

To this end, we turn to the literature on attributions. Attributions are the perceptions people form to explain the causes of a given behavior (Heider, 1958). In a book chapter about attribution theory, Manusov and Spitzberg (2008) state "When attributions are informative of a person's nature or personality, they are considered 'correspondent'" (p. 39). They are specifically referring to the work on correspondent inference theory (Jones and Davis, 1965). Jones and Davis stated that their purpose was to "… construct a theory which systematically accounts for a perceiver's inferences about what an actor was trying to achieve by a particular action" (1965, p.222). They go on to describe how correspondence is the link between the attributions a person forms and the behavior they are trying to explain. The behavior *corresponds* to their perceptions. In this way, individuals are forming opinions of others based on the attitudes they have formed and if the behavior is associated with those perceptions.

If online content is perceived to be manipulable and the reasons for this manipulation are apparent, individuals are likely to form more negative impressions. For example, one study found that perceptions of digital modification (e.g., Photoshop) of images shared on social media were negatively associated with perceptions of intelligence and honesty (Vendemia & DeAndrea,

2018). Interestingly, modified photos posted by peers were rated more negatively than models. One possible explanation considers the perceived motivation for manipulating the photos. A model is expected to look attractive, and it is likely some photo modification is expected as part of their job. A peer is not held to these standards, so photo modification may be seen with a harsher lens. Here, we can see the influence perceived motivation to manipulate has on impression formation as a process of attribution.

PATM still emphasizes a person's perceptions but is focused on the manner in which information has the potential to be altered, rather than the reasons for doing so. Previous studies on warranting value have identified that one's ability to edit content has an influence on impressions. For example, wiki pages can be edited by all users. One study examined the influence ability to manipulate had in a wiki used for student evaluations. They found that when a user was perceived to have the ability to modify their own self-evaluation, more negative impressions were formed about the legitimacy of the content and of the individual themself (DeAndrea et al., 2015). As discussed previously, the Shin et al., (2020) study examined information control behaviors on social media. When comments were deleted without justification, more negative evaluations emerged than when the comments were deleted with justification. Taken together, PATM is another common way warranting value may be elicited in online communication.

Previous measures of warranting value fall short of understanding these mechanisms of perceived manipulability. The most recent scales of warranting value by DeAndrea & Carpenter (2018) measure warranting in terms of a behavior (e.g., The content provider manipulated the information that appeared on the website about themselves). However, a central tenet of warranting theory is that *perceptions* of manipulability are influential in impression formations.

Their scale instead captures whether or not participants believe a behavior has occurred. This difference is important when considering information control. Particularly in causal ordering, PATM will precede the actual manipulation behavior. In order to capture perceived *ability* to manipulate the scale item should instead be, "The content provider *has the ability to manipulate* the information that appeared on the site about themselves." The inclusion of "has the ability to" encourages a more faithful scale utility than the current DeAndrea and Carpenter (2018) version as it directly addresses the perception of immunity or ease in which content can be manipulated. Further, it directly assesses the dimensions of information manipulation potential, rather than actual behavior.

Taken together, perceptions of how and why people manipulate content have been central to work in interpersonal and CMC research broadly. However, specific to truth default and warranting theories, specific consideration to the influence of the perceived motivation and ability to manipulate the content has not been identified. The present study examines these elements in an online dating context. Online dating is an underexplored context for these lines of research, but also provides ample opportunities to examine the constructs of interest. The next section synthesizes the compiled literature into a series of hypotheses and research questions in order to test the proposed relationships.

Hypotheses and Research Questions

The literature review thus far has emphasized both warranting theory and truth default theory. The predictions related to veracity judgements and impressions in these theories share many common threads. The present study seeks to test how these theories explain and predict impression formation behaviors both independently and conjointly using an online dating context. As such, the hypotheses and research questions build on each other and end with a final model that combines the predictions of these bodies of literature. A brief summary will precede each hypothesis.

PATM and PMTM were discussed as the potential drivers of warranting value. The manner in which individuals form attributions about a specific behavior is related to their impressions. The reasons (i.e., attributions) individuals ascribe to behavior (e.g., Heider, 1958; Jones and Davis, 1965) are indicative of PMTM while the perceptions of information control are indicative of PATM (e.g., DeAndrea & Carpenter, 2018). Motivation and ability to manipulate are proposed to function as correspondent cues associated with perceptions of warranting value.

H1: Perceived motivation to manipulate is negatively associated with perceived warranting value

H2: Perceived ability to manipulate is negatively associated with perceived warranting value

The combined effects of motivation and ability are less understood. As mentioned previously, there are likely cases in which warranting can have more or less value depending on the evaluation of the perceptions of the content provider. Independent variables such as perceptions of ability and motivation to manipulate content can (and should) be tested as both independent variables and as moderators (Hayes 2018; Spencer et al., 2005). The combination of

both perceived ability and perceived motivation may have an interaction effect on perceived warranting value. It is likely that high PATM and high PMTM will be associated with the lowest perceived warranting value, and low PATM and low PMTM will be associated with the highest perceived warranting value. This is in part due to the causal attributions that can be made from the available information in online contexts such as online dating. When both PATM and PMTM are attributable to a specific behavior, the most positive or negative effects on warranting value will also be present.

On the other hand, the conditions with mixed levels of PATM and PMTM (high ability/low motivation or low ability/high motivation) are less clear. It is uncertain if these elements can be distinguished. Examining the interaction effects of these variables will help determine if and how PATM and PMTM are related to perceived warranting value and ultimately impression formation. As such a research question is posed to examine the interaction between motivation and ability on perceived warranting value.

RQ1: What effects do the interaction of PATM and PMTM have on perceived warranting value?

Warranting value has been predicted to be an influential part of impression formation, particularly regarding the veracity of content (DeAndrea, 2014; Walther & Parks, 2002). An increase in warranting value should therefore be associated with an increase in perceived veracity.

H3: There is a positive association between perceived warranting value and content veracity

Following this logic, it is expected that there is a relationship between perceived veracity and trustworthiness. If an individual believes the content is accurate, they should perceive higher

levels of trust. Warranting theory predicts that perceived manipulability has an effect on impressions. If that is true the outcome of content veracity should be associated with trustworthiness.

H4: There is a positive association between perceived veracity and trustworthiness

If an individual trusts the content and/or the content provider, they should form more positive impressions of the content provider. Affective impressions such as a person's perceived warmth are likely associated with perceptions of trust.

H5: There is a positive association between perceived trustworthiness and perceived warmth

Truth default theory states that a trigger event will knock people out of a truth default state. It is proposed here that warranting value can act as that trigger. As such, when content high warranting value, it should be associated with high levels of trust. People should maintain the truth default state in this instance. On the other hand, when warranting value is low, individuals should be knocked out of the truth default state which should be associated with low levels of trustworthiness.

H6: There is a positive association between perceived warranting and perceived trustworthiness

Affective interpersonal impressions of the individual are also predicted to contribute to perceived veracity. Individuals who are perceived as warm should also be perceived to provide true and accurate content.

H7: There is a positive association between perceived warmth and content veracity

Taken together, the preceding relationships were developed into competing models. The predicted relationships were based on the theories described thus far. The next set of hypotheses

predicts the serially mediated relationship between the variables from a warranting theory perspective (see Appendix Figure A1). It is predicted that perceived motivation and ability drive perceptions of warranting value which in turn develops impressions of the content itself. Warranting theory predicts that warranting assessments allow us to form impressions of the individual, so perceptions of content veracity should influence perceptions of trust which in turn should determine one's overall affective warmth. PATM and PMTM will be examined individually in the predicted serially mediated relationships.

H8: The relationship between perceived motivation to manipulate (PMTM) and perceived warmth is serially mediated wherein PMTM is negatively related to warranting value, warranting value is positively related to content veracity, content veracity is positively related to trustworthiness and trustworthiness is positively related to warmth.

H9: The relationship between perceived ability to manipulate (PATM) and perceived warmth is serially mediated wherein PATM is negatively related to warranting value, warranting value is positively related to content veracity, content veracity is positively related to trustworthiness and trustworthiness is positively related to warmth.

Hypotheses 10 and 11 utilize the truth default theory perspective in order to predict the serially mediated relationship between the variables (see Appendix Figure A2). Perceived motivation and ability are still predicted to influence warranting value. However, in the TDT model, warranting value is predicted to act as a potential catalyst for knocking an individual out of a truth default state. Therefore, warranting value is predicted to influence trustworthiness which in turn influences perceived warmth. These impressions should then encourage individuals to form overall assessments of content veracity. The truth default model also examines motivation and ability separately at first.

H10: The relationship between perceived motivation to manipulate (PMTM) and perceived content veracity is serially mediated wherein PMTM is negatively related to warranting value, warranting value is positively related to trustworthiness, trustworthiness is positively related to warmth, and warmth is positively related to content veracity

H11: The relationship between perceived ability to manipulate (PATM) and perceived content veracity is serially mediated wherein PATM is negatively related to warranting value, warranting value is positively related to trustworthiness, trustworthiness is positively related to warmth, and warmth is positively related to content veracity

Once again, the interaction effect of motivation and ability is less understood. Therefore, a research question is used to examine the serially mediated relationships. The interaction effects of motivation and ability will be tested in both the warranting theory model and the truth default theory model.

RQ2: What effects do the interaction of perceived motivation and ability have on the proposed serial mediation models?

An experiment was conducted in order to test these hypotheses and research questions. The next section provides an overview of the experiment and describes in detail the measures used to capture the variables of interest.

Study 1 Methods

A 2 (high/low ability to manipulate) x 2 (high/low motivation to manipulate) design was used to examine the effects on perceived warranting value and impression formation. An online dating scenario was used in this study wherein participants viewed a fictional online dating profile. A *profile viewer* is an individual who is examining and evaluating an online dating profile (e.g., the participants in this study). A *profile owner* is an individual whose online dating profile is being seen by the profile viewer. High PATM means that the profile owner has the power, control, and resources (i.e., the ability) to make changes on their online dating profile to things such as the picture or written information. Low PATM means the profile owner does not have power, control, nor resources to make changes on their online dating profile to things such as the picture or written information. High PMTM means that the profile owner has reason to or desires to alter their profile to achieve some goal. Low PMTM means the profile owner does not have reasons or desires to alter their profile to achieve some goal. PATM and PMTM were used as factors in this study.

An online dating website and the profile images used in the stimuli were created specifically for use in this study. The author created a logo and webpage design for a fictitious online dating website named *Spartan Hearts*. The online dating website was created for use at the author's university, so the school colors, mascot, and slogan were used in the profile design. Images of the people used in the online dating profile stimuli were created for this project using AI generated images (https://generated.photos/). The male and female profiles shared similar facial features, hair color, and eye color (Appendix Figures A15 and A16). The AI program allowed the researcher to use the same AI generated person with more masculine or feminine traits to create male and female presenting pairs of images with matching appearances. While

recognizing that individuals are not beholden to a sex binary, this study utilized a male-presenting and female-presenting stimuli. Moving forward, these stimuli will be referred to as *profile sex* and be described as either the male or female profile.

Participants were recruited through a large Midwestern university's research pool in the United States and were given course credit for their participation. After completing the informed consent, individuals were asked to indicate their preference to view a male or female online dating profile. Participants were told they should choose to view the profile they would prefer to see in an online dating situation as they would be rating these profiles. Participants in the study were able to report their own gender and sexual orientation in the demographic section of the study (which will be discussed later).

Several pairs of images were pre-tested for use in this study using the university's research pool. A total of 90 individuals participated in the pre-test. Individuals who participated in the pre-test were ineligible for participation in the main study. Participants were asked to choose to evaluate male or female images based on their preferences in a real online dating scenario. Participants in the pre-test evaluated four images using 18 items adapted from the Snyder et al., (1977) measure of warmth, the McCroskey and Richmond (1999) affect scale, and the Burgoon et al., (1978) interpersonal affect scale. Items were presented using a 7-point binary differential and included items such as *Fake- Genuine* and *Unkind – Kind.* Participants were also asked to respond to a one item measure of attractiveness (on a scale of 1-10, how attractive do you think this person is?). Warmth items were reliable across all conditions (α=.94), and each alpha was over .91 for each of the 8 rated images.

Ultimately, the pair chosen for use in this study were rated as comparatively and highly attractive (male M=6.29, SD=1.60; female M=6.53, SD=1.34) and were the most similar in

ratings of warmth (male *M*=5.01, *SD*=1.03; female *M*=5.23, *SD*=1.20). Other parings were either rated low on evaluations of warmth and attractiveness or had large discrepancies between the male and female evaluations of warmth and attractiveness.

After selecting their preferred profile sex, participants were randomly sorted into one of the four possible conditions (see Figure 2). Participants viewed their preference (male or female see Appendix Figures A15 and A16), but the name and content of the profile were held constant. Participants viewed an online dating profile on Spartan Hearts for a person named Alex. After viewing the profile, participants were asked to evaluate the profile through a series of questionnaire items (the dependent variables). A short demographic questionnaire concluded the study.

The two factors: *PATM* and *PMTM* were manipulated by the description of the online dating website. In the high PATM condition, participants were informed that users of Spartan Hearts could create their own profiles. In other words, daters had the ability to change the content of their profile. In the low PATM condition, participants were informed that the university would create the dater's profiles for them using the information gained from the university such as their student pictures and university profiles. In other words, the dater does not have the ability to change the content of their profile. High PMTM was induced by telling participants that the best profile of the month would be featured and win a prize ($20 Starbucks gift card). The low motivation condition did not have a prize. Instead, participants were told that a random account would be featured each month. In other words, participants should recognize that the potential to win a prize can motivate online daters to manipulate their profile and thus increase their odds of winning. Perceptions of Alex's PMTM would therefore be higher in the prize condition than in the random feature (no prize; low PMTM) condition.

	High PMTM	Low Motivation PMTM
High PATM	*Students can create their own profile and upload their own pictures. This will help ensure everyone's profile is unique so that matches can be made.* **Each month the best profile is featured based on the number of interactions and likes the profile receives. This person will win a $20 Starbucks gift card to go on a first date**	*Students can create their own profile and upload their own pictures. This will help ensure everyone's profile is unique so that matches can be made.* **Each month a random profile is featured. Only active accounts qualify.**
Low PATM	*The university will automatically generate the profiles using information from the student's MSU profile and their student ID picture. This will help ensure everyone's profile is consistent so that matches can be made.* **Each month the best profile is featured based on the number of interactions and likes the profile receives. This person will win a $20 Starbucks gift card to go on a first date**	*The university will automatically generate the profiles using information from the student's MSU profile and their student ID picture. This will help ensure everyone's profile is consistent so that matches can be made.* **Each month a random profile is featured. Only active accounts qualify.**

Figure 2. PATM by PMTM experimental conditions in Study 1

The online dating profiles were created in such a way as to maximize internal and external validity and participant believability. It was important that the profiles were similar in nature to what one would see on other dating sites. To this end, we evaluated other online dating websites such as Plenty of Fish, Tinder, Grindr, and Hinge to determine the types of information individuals are able to share on these platforms. These websites often include at least a person's first name, age, and either the city they reside in or the distance between the potential matches. The websites also encourage users to share information about themselves such as their employment status or occupation, if they use tobacco products, if they have or desire children, and the type of relationship they are looking for (e.g., serious, hookup, friendship, etc.). Often,

this type of information is located underneath the user's profile picture and uses small graphics. For example, a baby carriage may be next to one's desire to have children.

We modeled our profile design after other popular online dating websites' profiles. The profile design of the experiment also used small graphic images next to descriptors. However, we did not want these items to deter or detract individuals participating in the study. To this end, we also needed to consider the type of information that would be accessible for an online dating website on a college campus. Spartan Hearts was described as a dating website for only college students at the university, so Alex's location was listed as the same city in which the university resides. Alex was also listed as 21 years old since college students are often between the ages of 18 and 24. Lastly, Alex's profile stated they were a senior, business major, and involved in Greek life (see Appendix Figures 15 and 16 for the full images and descriptions).

However, the information contained within the profile itself needed to be logically and functionally possible for each of the manipulated variables of PATM and PMTM. For example, in the low PATM condition, participants were informed that the online dating profile was automatically generated from university information. As such, the content of the profile needed to contain information that could ostensibly be gleaned from a university system such as age, location, major, hometown, and affiliation in student organizations and clubs. This information is often contained in an online dating profile in which the dater has full creative control, so it worked in each of the experimental conditions. In Study 1, this information was bolstered by things like the use of the university mascot in the dating website name (Spartan) and the use of the school slogan in the profile (Go Green Go White!).

Taken together, the profile used in the study needed to be similar to what users would experience on other dating platforms (externally valid) and believable and consistent for within

each of the conditions (internally valid). By carefully crafting the design of the profile, both were achieved.

Study 1 Measures

PMTM and PATM. PMTM was measured using a scale developed for this study. The current scale adapted items from the situational motivation scale (SIMS; Guay et al., 2001). Items included statements such as "Alex wants to put their best foot forward on their dating profile" and "Alex would like to have a featured profile on Spartan Hearts." The PATM measure was adapted from DeAndrea and Carpenter's (2018) measure of dissemination control. Items included statements such as, "Alex had the ability to control what information appeared on their profile" and "Alex picked what information was presented on their profile." Both PMTM and ability were measured on a 7-point Likert-type scale from strongly disagree to strongly agree. The PMTM (α= .83) and ability (α= .95) scales were reliable.

Warranting. Perceived warranting value was modified from DeAndrea and Carpenter (2018). Items were measured on a 7-point Likert-type scale from strongly disagree to strongly agree. Items included statements such as, "Alex manipulated the information that appeared on their profile about themselves" and "Alex controlled the information about themselves appearing on their profile." The warranting items were reliable (α= .78).

Content veracity. The content veracity measure was derived from the source obfuscation measure by DeAndrea and Carpenter (2018) and was specifically worded to address if the content was a true representation of the profile owner. Items were measured on a 7-point Likert-type scale from strongly disagree to strongly agree. Items included statements such as "Alex's Spartan Hearts profile shows who Alex really is" and "Alex's Spartan Hearts profile is a good representation of who they really are." The veracity items were reliable (α= .85).

Trustworthiness. The McCroskey and Teven (1999) scale was used to measure trustworthiness. Items were measured on a 7-point bipolar differential with opposite pole items. Examples included ratings from *Honest – Dishonest* and *Untrustworthy – Trustworthy*. The trustworthiness items were reliable (α= .94).

Warmth. Items based upon those used by Burgoon et al., 1978, McCroskey and Richmond, 1999, and Snyder et al., (1977), were used to generate a measure of affective warmth. This measures overall affective impression. Items were measured on a 7-point bipolar differential with opposite pole items. Examples included ratings from *Unkind – Kind* and *Cold – Warm*. The warmth items were reliable (α= .89).

Study 1 Results

Sample Information

There were 231 people who participated in this study. Five individuals were removed for not completing the study. Individuals who responded with the same answer across all measures (i.e., straight lining) were removed from analysis. Straight lining was determined by identifying participants who had an overall standard deviation zero (0) for the summation of all analysis variables. Therefore, 11 participants were removed for straight lining. A total of 215 individuals were used in the analysis of this study. A majority of participants indicated that they identified as female (50.7%). Others identified as male (47%), non-binary or third gender (.5%), one individual preferred to self-describe (.5%), and three individuals declined to respond (1.4%). Individuals who participated were between 18 and 24 years old (M=20, SD=1.36). Participants reported their race as being white (80%), Asian (8.8%), Black or African American (6.5%), Latino/a/x (4.7%), Middle Eastern or Arabic (2.8%), Native American, Indian, or Alaska Native (1.4%), Hawaiian or Pacific Islander (1.4%). Individuals were able to select more than one race, if applicable. Three individuals preferred to self-describe their race (1.4%), and two individuals declined to respond (.9%).

Participants reported being heterosexual (86.5%), bisexual (7.9%), homosexual (2.3%), pansexual (.9%), and demisexual (.5%). Two participants chose to self-describe (.9%), and two participants declined to respond (.9%). A majority of the participants reported being single (45.11%), while others reported being in a committed relationship (21.9%), casually dating (17.2%), seriously dating (11.6%), and engaged (.5%). Two participants chose to self-describe their relationship status (.9%) and six participants declined to respond (2.8%). A majority of participants reported using an online dating website previously (54.9%).

Manipulation Checks

Manipulation checks were conducted to examine if the continuous level measures of PMTM and PATM reflected the categorical 2x2 manipulation induced by the dating website description.

PMTM. A t-test found no significant differences between the high (M=5.06, SD=1.02) and low (M=5.02, SD=.97) PMTM conditions and the continuous level measure of PMTM $t(214)=.32$, $p=.75$. The results of this test indicate that the PMTM induction was not successful.

PATM. A t-test found a significant difference between the high (M=5.95, SD=.99) and low (M=5.38, SD=1.28) PATM conditions and the continuous level measure of PATM $t(214)=.3.67$, $p<.001$. The results of this test indicate that the PATM induction was successful.

Profile Sex. Additionally, a comparison between the male and female presenting profiles was conducted. A t-test found significant differences between individuals who viewed the male and female conditions on measures of trust $t(214)=-2.47$, $p=.01$ and warmth $t(214)=-2.45$, $p=.01$. There were no significant differences between male and female profiles for measures of PMTM $t(214)=.28$, $p=.78$, PATM $t(214)=1.27$, $p=.21$, warranting $t(214)=-1.72$, $p=.09$ or veracity $t(214)=.26$, $p=.80$. These tests indicate that the male and female tests were evaluated differently.

Covariates

No significant differences in high and low PMTM were observed, but there were significant differences between high and low PATM for the manipulation check continuous-level variables. PMTM and PATM were both dummy coded (low=0, high=1) for use in the experimental analysis and were used as covariates in hypothesis testing to account for the effects of PATM and PMTM on all model variables.

Differences in warmth and trust were also observed between individuals who viewed the male and female profiles. The female profile was rated as more trustworthy (M= 5.11, SD=.94) and warm (M=5.04, SD=.65) than the male profile (trustworthiness M=4.79, SD=.95; warmth M=4.83, SD=.64). This is perhaps unsurprising given that women are often rated as affectively and physically more attractive than men (Chappetta & Barth, 2016; McGloin & Denes, 2018). Differences in attraction vary by sex, and change as people age (Whyte et al., 2021). However, these differences were important to account for in the model, so profile sex was added as a covariate. Profile sex was also dummy coded (male profile = 1; female profile = 2) in order to be used in the hypothesis tests.

It is also expected that age, relationship status, and use of online dating websites could potentially drive some of the impression formation variables of interest in the study. The fictional person, Alex, was designed to appear college aged. Both the male and female profiles appear to be a young adult. College students are often within the age range of 18-24, and indeed the sample reflected this. However, individual preferences in age of a potential partner should be taken into consideration (see Whyte et al.,2021), so age was added as a covariate in the model. A dummy coded variable was created for use in the analysis regarding participants' relationship status. Individuals who reported being single were dummy coded as 0, while individuals who reported being in any relationship were coded as 1 (e.g., married, dating, engaged, etc.). Lastly, a variable was dummy coded on previous experiences with dating apps. Those who had never used an online dating website were coded as 1, while those who had used an online dating website were coded as 2. The dummy coded variable was used as a covariate in the model.

Study 1 Hypotheses and Research Questions

Hypotheses 1 through 7 predicted relationships between variables that were examined using a bivariate correlation. The full correlation table can be found in Figure 3.

Variable	*M*	*SD*	1	2	3	4	5	6
1. Motivation (PMTM)	5.04	0.99	-					
2. Ability (PATM)	5.66	1.18	.29***	-				
3. Warranting Value	3.29	1.06	-.22***	-.47***	-			
4. Veracity	4.2	1	.35***	.17*	-0.05	-		
5. Trustworthiness	4.94	0.96	.21**	0.06	0.02	.32***	-	
6. Warmth	4.93	0.65	.42***	0.05	0.05	.27***	.51***	-

Figure 3. Descriptive Statistics and Correlations for Study Variables used in H1-H7

****p<.001, **p<.01, *p<.05*

The first hypothesis (H1) predicted a negative relationship between PMTM and perceived warranting value. A t-test found no significant difference in perceived warranting value between the high (M=3.24, SD=1.08) and low (M=3.34, SD=1.055) PMTM conditions $t(214)=-.67$, $p=.5$. However, the t-test findings are not surprising given the non-significant findings of the manipulation check. A bivariate correlation revealed a significant negative relationship between PMTM and perceived warranting value ($r(213)=-.22$, $p<.001$). Taken together, correlational data were in the predicted direction and significant, yet the conditions were not significantly different. Therefore, the data were partially consistent with H1.

H2 predicted a negative relationship between PATM and perceived warranting value. A t-test found no significant difference in perceived warranting value between the high (M=3.22, SD=1.00) and low (M=3.36, SD=1.12) PATM conditions $t(214)=-.94$, $p=.35$. A bivariate

correlation revealed a significant negative relationship between PATM and perceived warranting value $r(213)=-.47$, $p<.001$. Correlational data were in the predicted direction and significant, yet the conditions were not significantly different. Therefore, the data were partially consistent with H2.

H3 predicted a positive association between perceived warranting value and content veracity. There was no significant correlation between perceived warranting value and content veracity $r(213)=-.05$, $p=.44$. The data were not consistent with H3. H4 predicted a positive association between perceived veracity and trustworthiness. A bivariate correlation revealed a significant relationship between veracity and trustworthiness ($r(213)=.32$, $p<.001$). The data were consistent with H4. H5 predicted a positive association between perceived trustworthiness and perceived warmth. A bivariate correlation revealed a significant relationship between trustworthiness and warmth ($r(213)=.51$, $p<.001$). The data were consistent with H5. H6 predicted a positive association between perceived warranting and perceived trustworthiness. There was no significant correlation between perceived warranting value and trustworthiness ($r(213)=-.02$, $p=.82$). The data were not consistent with H6. H7 predicted a positive association between perceived warmth and content veracity. A bivariate correlation revealed a significant relationship between warmth and content veracity ($r(213)=.27$, $p<.001$). The data were consistent with H7.

Hypotheses 8 and 9 predicted serially mediated relationship from PMTM (H8) or PATM (H9) through warranting value, content veracity, and trustworthiness, with the outcome of warmth as predicted by warranting theory. Two separate models were run using the PROCESS macro in SPSS (Model 6, Hayes, 2022) with a 95% confidence interval and 5000 bootstrap samples. A dummy coded variable of PMTM or PATM was added as the independent variable

(X), warmth was added as the dependent variable (Y), serially mediated in order by warranting value (M1), content veracity (M2), and trustworthiness (M3). Profile sex, age, relationship status, and previous use of an online dating website were added as covariates in the models.

The serially mediated path from PMTM to warmth was not significant (β=-.13, *p*=.12, LLCI=-.31, ULCI=.04). The direct effect was not significant (β=.10, *p*=.20, LLCI=-.25, ULCI=.05). The data were not consistent with H8. The serially mediated path from PATM to warmth was not significant (β=-.15, *p*=.08, LLCI=-.27, ULCI =.02). The direct effect was not significant (β=-.12, *p*=.11, LLCI=-.27, ULCI=.03). The data were not consistent with H9.

Hypotheses 10 and 11 predicted a serially mediated relationship from PMTM (H10) or PATM (H11) through warranting value, trustworthiness, and warmth with the outcome of content veracity as predicted by truth default theory. Two separate models were run using the PROCESS macro in SPSS (Model 6, Hayes, 2022) with a 95% confidence interval and 5000 bootstrap samples. A dummy coded variable of PMTM or PATM was added as the independent variable (X), content veracity was added as the dependent variable (Y), serially mediated in order by warranting value (M1), trustworthiness (M2), and warmth (M3). Profile sex, age, relationship status, and previous use of an online dating website were added as covariates in the models.

The serially mediated path from PMTM to veracity was not significant (β=.08, *p*=.57, LLCI=-.19, ULCI=.35). The direct effect was not significant (β=.14, *p*=.28, LLCI=-.12, ULCI=.41). The data were not consistent with H10. The serially mediated path from PATM to veracity was not significant (β=-.07, *p*=.62, LLCI=-.34, ULCI=.21). The direct effect was not significant (β=-.02, *p*=.87, LLCI=-.28, ULCI=.24). The data were not consistent with H11.

The research questions sought to examine the interaction effects in the models. In order to examine these effects, a moderated mediation was run using the Hayes PROCESS macro

(Model 83, Hayes, 2022) with a 95% confidence interval and 5000 bootstrap samples. In these tests, PMTM was used as the independent variable (X), dummy coded PATM was used as the moderator (W), and warranting value, content veracity, trustworthiness, and warmth were used as mediators (M) or outcomes (Y) based on the predicted models.

Research question 1 was proposed to determine the interaction effects of PATM and PMTM on perceived warranting value. The interaction effect of PATM and PMTM was not significantly associated with warranting value (β=.06, *p*=.83, LLCI=-.52, ULCI=.64).

Research question 2 was proposed to determine the interaction effects of PMTM and PATM on the proposed serial mediation models. In the warranting theory model, the direct effect was not significant (β=-.10, t=-1.27, p=.20, LLCI=-.25, ULCI=.05). The overall moderated mediation model was not significant (β =0, LLCI=-.02, ULCI=.0). There was no significant interaction effect of PMTM and PATM in this model. In the truth default theory model the direct effect was not significant (β=.14, t=1.09, p=.28, LLCI=-.12, ULCI=.41). The overall moderated mediation model was not significant (β=0, LLCI=-.01, ULCI=.01). There was no significant interaction effect of PMTM and PATM in this model.

Study 1 Discussion

Study 1 provided insights into the ways in which ability, motivation, warranting, trustworthiness, veracity, and warmth are related. Measures of PMTM and PATM were both significantly correlated with perceived warranting value (H1 and H2), but t-tests revealed no significant differences between the groups, providing partial support for the hypotheses. The results were not surprising considering the PATM and PMTM inductions failed manipulation checks. Several other pairs including veracity and trustworthiness (H4), trustworthiness and warmth (H5) and warmth and veracity (H7) were also correlated. Serial mediation models (H8-H11) were non-significant. Neither the warranting theory model nor the truth default theory model were consistent with the proposed hypotheses. The interaction effects of PMTM and PATM were not significantly associated with warranting value (RQ1), nor with the proposed serial mediation models (RQ2).

First, we consider the results of the correlational hypothesis tests. Many of the predicted relationships between the study variables were significantly correlated, and a pattern emerged. PMTM and PATM were both significantly and negatively correlated warranting value. This is not necessarily surprising, especially considering DeAndrea and Carpenter's (2018) work on the variability in types of warranting value. The PMTM and PATM measures were adapted from their work, and it makes sense that these variables would be correlated. As argued earlier, careful wording of warranting value scales may be in order if specific types of manipulability are being examined. As such, these tests are not causal in nature. Instead, we can merely say that these perceptions are or are not correlated. In this way, we can say that ability, motivation and warranting value were rated similarly to each other in this study.

Interestingly, PMTM was correlated with all of the study variables (PATM, warranting value, veracity, trustworthiness, and warmth), while PATM was only correlated with PMTM, warranting value and veracity. In this way, the continuous level measure of PMTM may have captured valuable insight. In particular, there were significant and positive correlations between PMTM and the impression variables. Perceptions that the content was being manipulated was associated with more trustworthiness, content veracity, and warmth. This may be in part due to the nature of the experiment. People who use online dating applications are likely using them to achieve some goal (e.g., get a date, find a partner, etc.). Results showed no significant difference between the high and low PMTM conditions on the continuous level measure of PMTM. Perhaps, in this study, perceived motivation to manipulate *for* the potential prize was not the driver. Instead, perhaps it was the perceived motivation to manipulate in order to get a date that drove these relationships. In this way, it may have been seen as acceptable to manipulate the information as it helped individuals form their impressions.

On the other hand, PATM was not correlated with these variables. At least with this sample, PATM was not associated with impressions of trustworthiness and warmth. Regardless of the perceived ability to manipulate the content of the online dating profile, participants rated the trustworthiness and warmth of the profile owner relatively equally. Previous studies of warranting theory found differences between first-person and third person generated content (Antheunis & Schouten, 2011; Walther et al., 2009). This has been attributed to perceptions of another person's *ability* to manipulate the content. In this study, the ability induction emphasized different types of information control (i.e., DeAndrea & Carpenter, 2018) and one's ability to manipulate the content. If information control variants of warranting value exist as predicted, it is

surprising that PATM was not significantly related to impressions of trustworthiness and warmth.

Similarly, other interesting nonsignificant correlations were found between warranting value and content veracity and warranting value and trustworthiness. Although not hypothesized, warranting value was not correlated with warmth. It is particularly surprising given that the warranting principle (Walther, 2002) proposes the value in warranting assessments in impression formation. The correlations begin to reveal that perhaps warranting value was not a useful impression formation mechanism in this study.

This became even more apparent in the serial mediation tests. Neither the warranting theory model nor the truth default theory model fit the data. It seems that warranting value was not helpful in impression formation nor as a truth-default cue. Given the results of the correlations, this result was not particularly surprising. The interaction effects of PATM and PMTM did not add any clarity to the predicted models. Taken together, the proposed models did not fit the data. However, the correlational data provided some insights into the potential relationships especially considering the continuous level measures of PMTM and PATM. As such, the failed manipulation may be a driving force in the lack of significant and interpretable results.

PMTM and PATM were manipulated in a 2x2 experiment. Participants viewed a randomized condition of high or low PMTM and high or low PATM. A manipulation check was conducted to provide insight on the effectiveness of the manipulation. Continuous level measures of PATM and PMTM were collected in the survey and compared to the categorical level data in the study. There was no significant difference in the continuous level measure of PMTM between the high and low PMTM conditions. This indicates that PMTM was not sufficiently

manipulated in the experiment. There was a significant difference in the continuous level measure of PATM and the high and low PATM conditions. PATM was successfully manipulated in this experiment.

One potential explanation for the failed manipulation of PMTM is due to the difficulties in assessing the variable itself. Previous studies have shown that it is hard to experimentally induce or encourage motivation in study participants. Examples include motivating people to engage in civic duties such as voting (Bennett, 2007) or to exercise and engage in health behaviors (Ntoumanis et al., 2017). This study complicated motivation further by trying to induce perceptions of motivation of another person. In this way, individuals were asked to reflect on the thoughts and attitudes of another person. Attribution theories (e.g., Heider, 1958; Kelly, 1967; Weiner, 1982) may help explain this process. Attributions are explanations for behavior that are causal in nature. In other words, people attribute causality to the actions oneself or others make (Jones & Nisbett, 1971; Kelly, 1967). The actor-observer model in particular describes how individuals make sense of others' behaviors. Importantly the actor-partner (also referred to as actor-observer) model emphasizes how individuals ascribe causality to behavior, especially when success or failure is involved (Jones & Nisbett, 1971; Malle, 2006). Taken together, it appears as though individuals in this study were not attributing PMTM as a causal process as anticipated which may have driven a failed manipulation.

Participants were told that the best profile each month would win a $20 gift card and be featured on the online dating website. In this experiment, it did not appear as though individuals were attributing this motivation (the gift card and feature) as a driver of profile manipulation. We can see the attributional nature of motivation more clearly in other studies. Neo and Johnson (2020) found an interesting effect of perceived manipulation when politics were involved. More

negative impressions were formed when individuals were perceived as motivated to manipulate (i.e., post negative book reviews about their out party). Essentially, people ascribed attributions for behavior more readily when political cues were present. Walther et al., (2009) found that claims of attractiveness were more subject to warranting value assessments than claims of extroversion or introversion. Not unlike the results seen in the attribution literature, it seems as though perceived motivation to manipulate is only a causal explanation for behavior in certain cases. Indeed, it may depend on the types of causality people infer in their attributions (Malle, 2006).

Perhaps in this study, the prize was not large enough nor perceived to be motivating enough for Alex (the fictitious online dating profile owner) to intentionally manipulate their profile. In other words, participants in the study did not perceive Alex as motivated to manipulate when the prize for doing so was only $20 and a featured profile. Inducing perceived motivation is tricky. Neo and Johnson (2020) may have seen these effects more clearly due to the proliferation of political misinformation and relevant political commentary during the time of their data collection (before the 2020 US presidential election). Motivation to manipulate is likely an easy perception to make when a contentious and divisive political battle is present and continually discussed. Walther et al., (2009) were the first to note that different types of claims have more or less influence on warranting value and impressions depending on how central they were to the context and the effect that claim would have overall. In their case, claims of attractiveness had more influence on perceptions of warranting value than claims of intro/extroversion on social media. It seems as though in the case of online dating, a low-value monetary prize does not elicit attributions of motivation as predicted.

However, there was a significant negative correlation between perceived motivation and warranting value. This indicates that even if the manipulation check did not work as planned, there may still be a relationship between these variables. As predicted, the more motivation a person perceives, the lower warranting value is assessed. If a person is perceived to be motivated to manipulate, then the content they produce is seen as more manipulable. While the manipulation itself may not have been successful, nor the proposed serial mediation models, there is some data to suggest that the relationship between motivation/ability and perceived warranting value exists. As such, there may be other factors that are influencing the study variables and the overall success of the manipulations.

One potential contributor to the non-significant findings may be the college student sample and the university-sponsored dating application. While college-aged students make up a large portion of online dating users (Vogels & Mcclain, 2023), and thus make for a convenient and valid sample, college students may be satisfied with their current options for online dating. There are several existing dating applications such as Tinder, Grindr, Hinge, and Bumble that are widely used. College students may have been specifically disinterested in a dating application run through their university. There are potential risks to privacy and safety that would discourage use of a university-sponsored dating site, especially with the plethora of alternatives.

Some individuals may be worried about the potential repercussions of providing their personal information to individuals within close proximity. For example, they may be worried about things like sexual assault (Scannel, 2019) or dating violence (Cali et al., 2013) that could be experienced. A university-sponsored application may have been unwanted for individuals who are in minoritized groups of sexual orientation, such as members of the LGBTQ+ community (Albury & Byron, 2016). Their sexual orientation may be privately held information,

or they may value having online dating spaces specific to their sexual orientation (e.g., Grindr for gay men). The ability manipulation was designed so that either the university created the profile (low ability to manipulate) or the individual themselves created their profile (high ability to manipulate). However, regardless of the condition, the online dating profile was sponsored by the university. Regardless of the condition, participants may not have desired the university dating site which contributed to their disinterest. As such, other options for experimentally manipulating perceived ability to manipulate should be considered, specifically regarding the website's host. Overall, the results of this study indicate that more research could be conducted using a wider sample that utilizes a different type of ability manipulation in order to gain a better understanding of how individuals form impressions in online contexts such as online dating.

Conclusion

The results of Study 1 provided some information on the relationship between the variables of interest. However, analyses indicated that PMTM was not induced strongly enough or in ways that influenced impressions. Neither of the proposed serial mediation models supplied ample enough evidence to indicate the nature of the proposed relationships. Neither the warranting theory nor the truth default theory approaches fully explained the relationship between the variables. A lack of interest in a university-sponsored online dating site may also have exacerbated the lack of significant and interpretable findings. As such, a second study was conducted to strengthen the manipulations and investigate the effects in a non-student sample. A larger discussion of the study results is presented later in this manuscript.

Study 2 Methods

The same 2 (high/low PATM) x 2 (high/low PMTM) design was used to examine the effects on perceived warranting value and impression formation in Study 2 as in Study 1. However, the PATM and PMTM inductions were strengthened based on previously noted failed manipulations and edited for a wider sample. To do this, a different fictitious online dating website, *Magnet*, was designed as the stimulus material for use in Study 2. Participants were recruited through Prolific (https://www.prolific.co) and were paid $2 for completing the survey. Participants were required to be between the ages of 18-32, live in the United States, and have a study completion rating of at least 95% in order to be eligible. Age was restricted to participants aged 18-32 because the AI generated images were young adults. Alex's age was not specifically given in the experiment. Therefore, if Alex were 25, participants who were at least seven years younger or older were invited to participate, capturing the young adult age range.

Study 2 used the same AI generated images as Study 1 but were added to the new Magnet webpage (see Appendix Figures A17 and A18). Magnet shared many of the same features as the Study 1 website, Spartan Hearts. However, as this website was designed for a wider audience, some of the information contained in the profile was changed. The author created new logo, slogan, and color scheme for the fictitious Magnet webpage.

Study procedures were similar to those in Study 1. After selecting their preferred profile sex, participants were randomly sorted into one of the four possible conditions (see Figure 4). Participants viewed their preference (male or female see Appendix Figures A17 and A18). Participants viewed an online dating profile on Magnet for a person named Alex. After viewing the profile, they were asked to evaluate the profile through a series of questionnaire items (the dependent variables). A short demographic questionnaire concluded the study.

As in Study 1, PMTM and PATM were manipulated by way of the online dating website description (see Figure 4 for exact text). In the high PATM condition, participants were informed that daters had full creative control over the content in their profiles. In the low ability condition, participants were informed that dater's profiles were created using AI. Participants were instructed that information about the daters was gleaned from their social media profiles and all other publicly available data, these data were then fed into an AI tool and an online dating profile was, ostensibly, generated by the AI tool. The low PATM condition was described as allowing users full creative control of their profile.

In the high PMTM condition, participants were told that the best profile each month would be featured and would win a large prize ($1000 gift card for clothing and a year's subscription to a food delivery company Hello Fresh). The larger prize was expected to elicit more perceived motivation to favorably manipulate one's own data, as compared to the smaller prize in Study 1. In the low PMTM condition, a random profile would be featured. The low PMTM condition included no prize. As compared to the high PMTM condition, no perceived motivation to (favorably) manipulate one's profile should be elicited. The lack of incentive and randomness of selection should not encourage perceptions of motivation to manipulate.

The content of the profile differed from Study 1, however internal and external validity and participant believability were still considered for the wider United States sample. Once again, a careful design of the profile ensured that regardless of the condition participants viewed, the profile was similar to what they would see on a normal dating site and was believable. The content of the profile remained consistent in each condition, and therefore careful consideration was taken regarding profile design. ChatGPT (https://chat.openai.com/) was used to help write the content of the *About Me* section of the Magnet profile. A prompt was written to the AI,

asking for a short biography for an online dating profile that incorporated interests in travel, salsa dancing, and food. The bio written by ChatGPT was edited for clarity and used in the final stimuli. As before, the bio needed to make logical sense in each of the conditions.

	High PMTM	**Low PMTM**
High PATM	*We have made the process easy! Daters have complete creative control over everything that appears in their profile. Daters can write their own bio and select their own pictures using anything they want to post in order to create their profile. This will help daters show their best side so that matches can be made.* **The profile with the most interactions each month will be featured and win a $1000 gift card to the clothing retailer of their choice to purchase a new wardrobe for their future dates and free meals for a year from Hello Fresh (valued at $4160). Only active accounts apply.**	*We have made the process easy! Daters have complete creative control over everything that appears in their profile. Daters can write their own bio and select their own pictures using anything they want to post in order to create their profile. This will help daters show their best side so that matches can be made.* **A random profile is featured each month. Only active accounts apply.**
Low PATM	*We have made the process easy! Daters do not create their own profile or add their own pictures. Instead, AI (artificial intelligence) will automatically generate a bio and select pictures using the dater's social media presence and all publicly available information. This will help daters show their best side so that matches can be made.* **The profile with the most interactions each month will be featured and win a $1000 gift card to the clothing retailer of their choice to purchase a new wardrobe for their future dates and free meals for a year from Hello Fresh (valued at $4160). Only active accounts apply.**	*We have made the process easy! Daters do not create their own profile or add their own pictures. Instead, AI (artificial intelligence) will automatically generate a bio and select pictures using the dater's social media presence and all publicly available information. This will help daters show their best side so that matches can be made.* **A random profile is featured each month. Only active accounts apply.**

Figure 4. PATM by PMTM experimental conditions in Study 2

Additionally, since Study 2 focused on a United States sample rather than a college student sample, information about the dater included their relationship status (single), birthday, location, and employment status (as compared to information about their college degree in Study 1). The profile design was once again modeled after other popular online dating websites' profiles and used small graphic images next to descriptors. However, we did not want these items to deter or detract individuals participating in the study. For example, we did not want participants to rate Alex's profile negatively because they had a negative attitude or stereotype about Alex's career or if Alex was not located in the same city. To get around these issues, we used simplified and somewhat ambiguous versions in our experiment by showing an image of a briefcase and stating that Alex is employed, and that Alex was located "Within 50 miles of you." (see Appendix Figures A17 and A18 for the full images and descriptions). Once again, the profile designed for use in this study was believable, internally, and externally valid.

Study 2 Measures

All the measures from Study 1 were used in Study 2. Only the name of the dating website was changed within the questions. Otherwise, the content of the questionnaire was the same for Study 1 and Study 2. All variables in the study were reliable including PMTM (α= .85), PATM (α= .98), warranting value (α= .89), veracity (α= .95), trustworthiness (α= .95) and warmth (α=.95).

Study 2 Results

Sample Information

There were 303 individuals who participated in this study. One participant was removed for data quality and not adhering to the Prolific guidelines. One additional participant was removed for incomplete responses. Individuals who responded with the same answer across all measures (i.e., straight lining) were also removed from analysis. Straight lining was determined by identifying participants who had an overall standard deviation zero (0) for the summation of all analysis variables. Therefore, 1 participant was removed for straight lining. A total of 300 individuals were used in the analysis of this study. A majority of participants indicated that they identified as male (58%). Others identified as female (40%), non-binary or third gender (1.7%), and one individual declined to respond (.3%). Individuals who participated were between 18 and 32 years old (M=26.98, SD=3.75). Participants reported their race as being white (71.3%), Black or African American (11.7%), Asian (11.3%), Native American, Indian, or Alaska Native (1%), Hawaiian or Pacific Islander (1%), and Middle Eastern or Arabic (.3%). One participant chose to self-describe (.3%) and 5 participants declined to respond (1.7%). Individuals were able to select more than one race, if applicable.

Participants reported being heterosexual (73%), bisexual (14.7%), homosexual (6.3%), pansexual (2.7%), asexual, (2%), three individuals preferred to self-describe (1%), and one participant declined to respond (.3%). A majority of the participants reported being single (47%), while others reported being in a committed relationship (23%), married (17.3%), casually dating (4.7%), engaged (4%), or seriously dating (2.3%). One participant preferred to self-describe their relationship (.3%) and four participants preferred not to respond (1.3%). A majority of participants reported using an online dating website previously (71.3%).

Manipulation Checks

PMTM. A t-test found no significant differences between the high (*M*=5.77, *SD*=.81) and low (*M*=5.65, *SD*=.98) PMTM conditions and the continuous level measure of PMTM $t(299)=1.21$, $p=.23$. The results of this test indicate that the PMTM induction was not successful.

PATM. A t-test found a significant difference between the high (*M*=6.37, *SD*=.72) and low (*M*=3.27, *SD*=1.51) PATM conditions and the continuous level measure of PATM $t(299)=22.75$, $p<.001$. The results of this test indicate that the PATM induction was successful.

Profile Sex. A t-test found significant differences between individuals who viewed the male and female conditions on measures of trust $t(299)=-2.27$, $p=.02$ and warmth $t(299)=-2.04$, $p=.04$. There were no significant differences between male and female profiles for measures of PMTM $t(299)=.17$, $p=.86$, PATM $t(299)=-.20$, $p=.84$, warranting $t(299)=-1.51$, $p=.13$, or veracity $t(299)=-1.28$, $p=.20$.

Covariates

Once again, no significant differences in high and low PMTM were observed, but there were significant differences between high and low PATM for the manipulation check continuous-level variables. PMTM and PATM were both dummy coded (low=1, high=2) for use in the experimental analysis and were used as covariates in hypothesis testing to account for the effects of PATM and PMTM on all model variables.

Differences in warmth and trust were also observed between individuals who viewed the male and female profiles. The female profile was rated as more trustworthy (M= 5.09, SD=1.08) and warm (M=5.41, SD=.77) than the male profile (trustworthiness M=4.80, SD=1.12; warmth M=5.21, SD=.92). Profile sex was also dummy coded (male profile = 1; female profile = 2) in order to be used in the hypothesis tests.

As in Study 1, age, relationship status, and use of online dating websites could potentially drive some of the impression formation variables of interest in the study. Age was considered as a covariate in the study. Only 18–32-year-old individuals were allowed to participate. However, things like individual preferences in a partner's age, previous dating experience, and other effects of dating based on age should be taken into consideration, so age was added as a covariate in the model (Whyte et al., 2021). A dummy coded variable was created for use in the analysis regarding participants' relationship status. Individuals who reported being single were dummy coded as 1, while individuals who reported being in any relationship were coded as 2 (e.g., married, dating, engaged, etc.). Lastly, a variable was dummy coded on previous experiences with dating apps. Those who had never used an online dating website were coded as 1, while those who had used an online dating website were coded as 2. The dummy coded variable was used as a covariate in the model.

Study 2 Hypotheses and Research Questions

Hypotheses 1 through 7 predicted relationships between variables that were examined using a bivariate correlation. The full correlation table can be found in Figure 5. The first hypothesis (H1) predicted a negative relationship between PMTM and perceived warranting value. A t-test found no significant difference in perceived warranting value between the high (M=3.94, SD=1.61) and low (M=3.72, SD=1.45) PMTM conditions $t(299)=1.24$, $p=.21$. However, the t-test findings were again not surprising given the non-significant findings of the manipulation check. A bivariate correlation revealed a significant relationship between PMTM and perceived warranting value $r(298)=-.31$, $p<.001$. Taken together, correlational data were in the predicted direction and significant, yet the conditions were not significantly different. Therefore, the data were partially consistent with H1.

Variable	*M*	*SD*	1	2	3	4	5	6
1. Motivation (PMTM)	5.71	0.9	-					
2. Ability (PATM)	4.82	1.95	.41***	-				
3. Warranting Value	3.83	1.54	-.31***	-.73***	-			
4. Veracity	4.6	1.23	.40***	.21***	-0.04	-		
5. Trustworthiness	5	1.1	.32***	.20***	-0.04	.61***	-	
6. Warmth	5.33	0.84	.56***	.18**	-0.02	.54***	.65***	-

Figure 5. Descriptive Statistics and Correlations for Study Variables used in H1-H7

****p<.001, **p<.01, *p<.05*

H2 predicted a negative relationship between PATM and perceived warranting value. A t-test found a significant difference in perceived warranting value between the high (*M*=2.91, *SD*=1.12) and low (*M*=4.74, *SD*=1.34) PATM conditions *t*(299)=-12.87, *p*<.001. A bivariate correlation revealed a significant relationship between PATM and perceived warranting value *r*(298)=-.73, *p*<.001. The data were consistent with H2.

H3 predicted a positive association between perceived warranting value and content veracity. There was no significant correlation between perceived warranting value and content veracity *r*(298)=-.04, *p*=.49. The data were not consistent with H3.

H4 predicted a positive association between perceived veracity and trustworthiness. A bivariate correlation revealed a significant relationship between veracity and trustworthiness *r*(298)=.61, *p*<.001. The data were consistent with H4.

H5 predicted a positive association between perceived trustworthiness and perceived warmth. A bivariate correlation revealed a significant relationship between trustworthiness and warmth *r*(298)=.65, *p*<.001. The data were consistent with H5.

H6 predicted a positive association between perceived warranting and perceived trustworthiness. There was no significant correlation between perceived warranting value and trustworthiness $r(298)=-.04$, $p=.52$. The data were not consistent with H6.

H7 predicted a positive association between perceived warmth and content veracity. A bivariate correlation revealed a significant relationship between warmth and content veracity $r(298)=.54$, $p<.001$. The data were consistent with H7.

Hypotheses 8 and 9 predicted serially mediated relationship from PMTM (H8) or PATM (H9) through warranting value, content veracity, and trustworthiness, with the outcome of warmth as predicted by warranting theory. Two separate models were run using the PROCESS macro in SPSS (model 6, Hayes, 2022) with a 95% confidence interval and 5000 bootstrap samples. A dummy coded variable of PMTM or PATM was added as the independent variable (X), warmth was added as the dependent variable (Y), serially mediated in order by warranting value (M1), content veracity (M2), and trustworthiness (M3). Profile sex, age, relationship status, and previous use of an online dating website were added as covariates in the models.

The serially mediated path from PMTM to warmth was not significant (β=.01, $p=.89$, LLCI=-.18, ULCI=.20). The direct effect was not significant (β=.07, $p=.34$, LLCI=-.07, ULCI=.22). The data were not consistent with H8. The serially mediated path from PATM to warmth was significant (β=.25, p=.01, LLCI=.06, ULCI=.44). The proposed indirect effects (PATM – warranting value – content veracity – trustworthiness – warmth) were not significant (β=-.03, LLCI=-.08, ULCI=.02). However, the indirect effects excluding warranting value (PATM – content veracity – trustworthiness – warmth) were significant (β=.12, LLCI=.04, ULCI=.22). The indirect effect from PATM – veracity – warmth was also significant (β=.09,

LLCI=.01, ULCI=.18) The direct effect was not significant (β=.05, *p*=.62, LLCI=-.14, ULCI=.23). Therefore, the data were not consistent with H9.

Hypotheses 10 and 11 predicted a serially mediated relationship from PMTM (H10) or PATM (H11) through warranting value, trustworthiness, and warmth with the outcome of content veracity as predicted by truth default theory. Two separate models were run using the PROCESS macro in SPSS (model 6, Hayes, 2022) with a 95% confidence interval and 5000 bootstrap samples. A dummy coded variable of PMTM or PATM was added as the independent variable (X), content veracity was added as the dependent variable (Y), serially mediated in order by warranting value (M1), trustworthiness (M2), and warmth (M3). Profile sex, age, relationship status, and previous use of an online dating website were added as covariates in the models.

The serially mediated path from PMTM to veracity was not significant (β=-.01, *p*=.97, LLCI=-.28, ULCI =.27). The direct effect was not significant (β=.05, *p*=.65, LLCI=-.17, ULCI=.27). The data were not consistent with H10. The serially mediated path from PATM to veracity was significant (β=.43, p=.0025, LLCI=.15, ULCI=.71). The proposed indirect effect (PATM – warranting value – trustworthiness – warmth – content veracity) was not significant (β=-.03, LLCI=-.08, ULCI=.01). However, the indirect effect excluding warranting value (PATM – trustworthiness – warmth – content veracity) was significant (β=.09, LLCI=.01, ULCI=.21). The indirect effect from PATM – warmth – content veracity was also significant (β=.26, LLCI=.09, ULCI=.47). The direct effect was not significant (β=.22, *p*=.12, LLCI=-.06, ULCI=.50). Taken together, the data were not consistent with H11.

The research questions sought to examine the interaction effects in the models. In order to examine these effects, a moderated mediation was run using the Hayes PROCESS macro (model 83, Hayes, 2022) with a 95% confidence interval and 5000 bootstrap samples. In these

tests, motivation was used as the independent variable (X), dummy coded ability was used as the moderator (W), and warranting value, content veracity, trustworthiness, and warmth were used as mediators (M) or outcomes (Y) based on the predicted models.

Research question 1 was proposed to determine the interaction effects of PATM and PMTM on perceived warranting value. The interaction effect of PATM and PMTM was not significantly associated with warranting value (β=-.29, t=-1.00, p=.32, LLCI=-.85, ULCI=.28)

Research question 2 was proposed to determine the interaction effects of PMTM and PATM on the proposed serial mediation models. In the warranting theory model the direct effect was not significant (β=.07, *p*=.34, LLCI=-.07, ULCI=.22). The overall moderated mediation model was not significant (β =-.004, LLCI=-.03, ULCI=.01). There was no significant interaction effect of PMTM and PATM in this model. In the truth default theory model the direct effect was not significant (β=.05, *p*=.65, LLCI=-.17, ULCI=.27). The overall moderated mediation model was not significant (β=.01, LLCI=-.06, ULCI=.01). There was no significant interaction effect of PMTM and PATM in this model.

Study 2 Discussion

Study 2 provided new insights into the predicted relationships proposed by both warranting and truth default theories. First, a discussion specifically about Study 2 is presented, followed by an overall discussion of both studies and their theoretical contributions. The results of Study 2 found significant correlations between many of the variables. Measures of PMTM and PATM were both significantly correlated with perceived warranting value (H1 and H2), but t-tests revealed no significant differences between the groups for PMTM. There were significant differences between the high and low PATM groups, and data were overall consistent with H2 but not H1. Several other pairs including veracity and trustworthiness (H4), trustworthiness and warmth (H5) and warmth and veracity (H7) were also correlated. Serial mediation models (H8-11) were non-significant, although some indirect effects were present in the PATM paths. The interaction effects of PMTM and PATM were not significantly associated with warranting value (RQ1), nor with the proposed serial mediation models (RQ2).

As in Study 1, an interesting pattern of results emerged within the correlational data. As predicted, there were significant correlations between PMTM, PATM, and warranting value. PATM and warranting value were strongly negatively correlated. A negative correlation was also present between PMTM and warranting value. This is in line with the central proposition of warranting theory. As it becomes easier to manipulate a piece of information, lower levels of warranting value should be observed (DeAndrea, 2014; Walther and Parks, 2002). However, contrary to research in this area that states warranting value is associated with impressions (e.g., DeAndrea, 2014; Walther and Parks, 2002), warranting value was not associated with content veracity, trustworthiness, nor warmth. Continuous measures of PMTM and PATM were significantly correlated with all other measures in the study and with each other. In this way, the

correlational data may have captured the disconnect that occurred in the proceeding hypothesis and model tests. The only non-significant correlations occurred between warranting value and the impression variables. It appeared as though motivation, ability, and warranting were related to each other in one grouping, while veracity, trustworthiness, and warmth were related in another. Warranting value, which was predicted to be the link between them, was instead unrelated to the impression outcomes. Upon examining both correlation tables in Study 1 and 2, the non-significant relationships between warranting value and the impression outcomes even visually separates the correlation tables.

The serial mediation models were not significant in Study 2. Neither the truth default nor the warranting theory models fully predicted the relationships between the study variables. The results were perhaps unsurprising given the observable pattern in correlational data. However, there were significant indirect pathways in the ability models. The significant indirect pathways *excluded* warranting value. The correlational data provide additional insight to these results as warranting value was not associated with the impression variables.

The manipulation of motivation was problematic in Study 2. Even though continuous level measures of PATM and PMTM were in the predicted direction, the manipulation check for motivation failed as there was no significant difference between the high and low PMTM conditions. The prize was increased in this study to be of higher value (from $20 to over $4000), but this increase was still not enough to change perceptions of motivation in this study (as compared to the low motivation – no prize offered condition). Here again, we may turn to the causal nature of attributions to understand the results. The large prize was not significantly driving perceptions of motivation. In other words, it was not the perceived causal driver of perceived motivation. Generally speaking, scores for PMTM were high across all conditions.

Mean scores for both the high and low PMTM conditions were over 5.5. The ability induction on the other hand, was strengthened in Study 2. Individuals perceived a large difference between the high and low ability conditions with nearly a 3 point mean difference between the scores. Even with the strengthened ability induction, the models were not supported by the data.

Taken together, the results of Study 2 were similar to those of Study 1. Some correlational data supported the hypotheses, but the serial mediation models did not significantly fit as predicted. However, with the change in the manipulation and wider sample, a few new trends appeared and raised new questions about these relationships. Next, the two studies will be discussed together in order to provide a larger discussion regarding the study predictions, findings, and future directions.

Overall Discussion

The results of Study 1 and Study 2 were similar. Even though the induction was strengthened and a wider US sample was used in Study 2, the predicted relationships generally were not found. In this discussion, Study 1 and Study 2 are considered in tandem and theoretical implications are provided. Before discussing specific results of these studies, it is pertinent to consider why and how the two studies were complicated by the failed motivation manipulation.

There were no significant differences in PMTM between the high and low PMTM conditions in Study 1 nor Study 2. The first consideration is that there was a ceiling effect of perceived motivation. The mean values for Study 1 were approximately 5, and over 5.5 for Study 2. Considering the scale ranged from 1-7, these scores were relatively high. If we consider that even the low PMTM condition resulted in a mean score over 5, there was simply not a lot of room to move the needle in a high motivation condition. Study 2 was designed to increase perceptions of PMTM in order to win a prize. Therefore, the prize increased from $20 in Study 1 to over $4000 in Study 2. Regardless of the size of the prize, PMTM was relatively equal across conditions. This indicates that perhaps PMTM is not at all driven by potential monetary rewards, but by something else entirely. Even so, perceived PMTM is relatively high regardless of a potential prize.

One potential explanation of the ceiling effect lies in the context itself – online dating. In online dating, perceptions of motivation can come from the goals of the users themselves, which in most cases is to find a potential match. The type of match they are looking for (e.g., long term partnership, marriage, friendship, sexual relationships, etc.) may vary, but they are still motivated by their relational goal (Menkin et al., 2015). Perceptions of the motivations for manipulating content to make a match and go on a date may have driven up scores of perceived motivation in

these studies. In other words, people who use online dating websites are looking for matches. They all may be naturally motivated to manipulate their profiles in ways that make them look like an attractive partner and result in more matches. While not accounted for in this study, the reasons for using an online dating website could have had the effect of subjects attributing perceived motivation to manipulate their profile information favorably to all conditions equally.

It is important to note that PATM appeared to be successfully manipulated in both studies. There were significant differences in PATM when comparing the high and low PATM conditions. This effect came out more strongly in Study 2, wherein the low PATM condition participants were told the profile was AI generated. Interestingly, PATM was not correlated with trustworthiness nor warmth in Study 1. In Study 2, PATM was significantly correlated with all the study variables. Cues that are unable to be manipulated have been examined in the past as potentially informative warranting cues. System generated cues for example are automatically generated based on a users' activity (Antheunis & Schouten, 2011). In this way, having little to no ability to manipulate the content, as in the case of AI generated content, may function similarly to a system generated cue.

This paper emphasized the role of warranting value as an impression formation mechanism (i.e., warranting theory) and a cue that would cause individuals to abandon their truth default state (i.e., truth default theory). In both studies, warranting value was only significantly correlated with the continuous level measures of PATM and PMTM, but not the impression variables of trustworthiness, content veracity, and warmth. This disconnect is also present in the serial mediation tests. Therefore, there are several possible explanations for the significant correlations, yet nonsignificant serially mediated relationships. Each theory provided different explanatory mechanisms; data collected here were not consistent with the theories as predicted.

Warranting Theory

The serial mediation model predicted by warranting theory was not significant. In this experiment, it appeared as though perceived warranting value was not a useful predictor of content evaluations nor impression variables. In every model, the relationship between warranting value and veracity and trustworthiness was non-significant. Additionally, warranting value was not associated with the affective impression variable, warmth. However, trustworthiness, veracity, and warmth were all significantly and positively correlated with each other. While only PATM and PMTM were significantly correlated with warranting value.

As previously discussed, the lack of significant relationships between warranting value and the impression variables is surprising. The results of this study run contrary to previous findings of warranting theory research. Considering the warranting principle, it would be expected that warranting value would be associated with impressions, particularly affective warmth. However, these relationships were not found in this study. There are several potential explanatory mechanisms for the lack of significant relationships.

First, the measures used in this study were considered. Ability was a much better predictor in the present models. As ability was strongly correlated with warranting value, and the ability measure was adapted from DeAndrea and Carpenter (2018), it may simply be a better metric for examining warranting value in this instance. DeAndrea and Carpenter's (2018) work identified that there are several types of information control that can influence warranting value. Perhaps the types of information control identified in their study are limited to certain contexts or are missing valuable dimensions.

The wording of the warranting value scales should be considered. As mentioned previously, the DeAndrea and Carpenter (2018) scales of warranting value most often capture a

behavior that has already occurred (e.g., The content provider manipulated the information that appeared on the website about themselves). This phrasing fails to capture the true element of the theory's core predictions of the role of perceived manipulability. The added phrasing of "has the ability to" into the warranting items captures a more accurate theoretically driven measure of perceived warranting value. An attempt was made in this study to isolate perceptions of ability as a predictor of warranting value. While the correlations were significant for continuous level measures, the predicted relationships did not work when warranting value itself was added to the model.

The indirect effects that excluded warranting value in Study 2 were significant. This is an interesting finding, given the theoretical predictions hinged on the assessment of warranting value. It may be that the measure of general warranting value did not capture the manner in which PMTM and PATM were used in forming perceptions of warranting value. Instead, perhaps PMTM and PATM are their own sources of warranting value. Much like the different dimensions of information control outlined by DeAndrea and Parks (2018), there are likely more forms of warranting value that can be assessed. To this end, warranting value itself could be multidimensional. Rather than emphasizing the different types of information control, perhaps it would be pertinent to consider the ways we could assess warranting value more generally (i.e., applicability to other contexts) or with more specificity (i.e., determine how warranting value is assessed situationally). This study emphasized perceived PMTM and PATM as predictors of warranting value, but perhaps they are simply dimensions of warranting value. Perhaps it is not only information control that indicates warranting value assessments. Identifying and capturing these elements is challenging, but would provide added insight into the way in which warranting value impacts impression formation in online contexts.

Overall, warranting value was not associated with impression formation in this study. This runs contrary to the predictions of the warranting principle (e.g., DeAndrea, 2014). The serial mediation predictions were nonsignificant and were not consistent with the data. Moreover, even correlational data did not show relationships between warranting value and the impression outcomes. However, the truth default theory model fared no better.

Truth Default Theory

The serial mediation predicted by truth default theory was not significant. In this study, warranting value did not act as a cue that caused people to abandon a truth default state. Perceptions of manipulability were not associated with the processes proposed by TDT. There was no significant relationship between warranting value and trustworthiness. Correlational tests found significant relationships between ability or motivation and perceived warranting value. There were no significant correlations between warranting value and trustworthiness, warmth, or content veracity. The TDT model proposed here emphasized the role of warranting value in lieu of suspicion in deception detection. However, these predictions were not supported by the data. It is important to recognize that this is one of the first experimentally driven tests that utilizes both warranting theory and truth default theory. While the study predictions may not have been supported by the data, future exploration and collaboration of the two theories are encouraged.

Perhaps by asking individuals to assess warranting value, the effects were diminished. As Levine (2020) points out, asking individuals if they are suspicious can make people suspicious. When this happens, it is hard to differentiate between suspicion elicited by the induction or by the questions regarding suspicion. At the same time, how can you ask people to discuss their feelings (or lack thereof) of suspicion without asking them? It is possible that perceived manipulability may also fall victim to this issue. People may not actively ascribe warranting

value until prompted to do so. Alternatively, as these perceptions were not associated with impressions, warranting value may not have been a useful diagnostic tool.

Further, people may not be engaging in thoughtful processing in order to utilize warranting value in diagnostically useful ways. Many applications such as Tinder rely on quick judgements. The swiping mechanism makes it quick and easy to match or unmatch with potential partners (van der Zanden et al., 2022). Swiping on Tinder, for example, is when an individual drags their finger across their phone to the right if they like the profile, and to the left if they do not like the profile (see https://tinder.com/faq). Matching in online dating means that both individuals have indicated their positive interest in getting to know the other person. Many dating applications also require that individuals both agree to match with one another before they are allowed to communicate. As such, people may utilize heuristic cues located within the pictures or information provided by the profile owner in their bios (van der Zanden et al., 2022). People may not be considering the warranting value of a potential match's profile at first glance, nor using it to form impressions at the initial matchmaking stage. Especially, as in the case of some apps, they risk not being able to communicate unless both agree to communicate (e.g., both swipe right on Tinder) resulting in a match. Perceptions of manipulability likely become more important in impression formation and ultimately content veracity once both partners can communicate and dedicate time to more systematic processing and relationship development. In this way, the timing of warranting value evaluations becomes important. Finding a way to examine an individual's perceived warranting value throughout the online dating process could be one way to disentangle these relationships.

Causal attributions were described as a potential explanation for why and how perceived PMTM and PATM to manipulate were related to warranting value and impression formation.

Considering the results of the models, it is possible that individuals in these studies were making attributions that were not predicted.

A related term in the attribution literature is the correspondence bias. A correspondence bias is a person's tendency to form impressions about a person's disposition from an event or situation that itself explains the behavior (Gilbert & Malone, 1995).

A related line of research in attributions lies in actor-observer or effects (Jones & Nisbett, 1971; Malle, 2006). Actor-observer effects researchers have observed a difference in how individuals attribute meaning and causality to another individual as compared to themselves. A meta-analysis found that actor-observer effects are not quite as ubiquitous nor strongly supported in research as once believed (Malle, 2006). Instead, Malle posits that a self-serving bias may describe these effects more thoroughly. In other words, people may form attributions about themselves and of others to achieve a self-serving goal (e.g., to make themselves feel good about a choice or to downplay other's accomplishments when jealous). In this way, individuals on an online dating website that are perceived to have manipulated their content in some way are likely perceived as untrustworthy especially if the reason for the manipulation is perceived to be for their own benefit. This study attempted to elicit these attributions by way of a large prize. In the high PMTM condition, the potential to win a large prize would be correspondent information related to perceptions of trust, content veracity and warmth. However, as these relationships were generally not found in this study, the correspondent information was not associated with the outcomes.

Post Hoc Tests

The results of the hypothesis tests of the warranting and truth default theory models that used dummy coded variables were generally not consistent with the predictions. However, given

that the correlations were significant using the continuous level data, the models were run again, but with the continuous level measures of PMTM and PATM.

No significant serial mediation was present in the Study 1 tests. Results showed a very similar pattern of results as the dummy coded variable models such that the relationships between warranting value and the impression variables were not significant. As before, PMTM was also not a helpful predictor in the model, but PATM was somewhat explanatory. The post-hoc tests for Study 2 showed more significance, especially for models that included PATM. As can be seen in the models below, H9 and H11 were consistent with the data when using the continuous level measure of PATM. This pattern of results was not seen in the tests using PMTM or the interaction of PATM and PMTM.

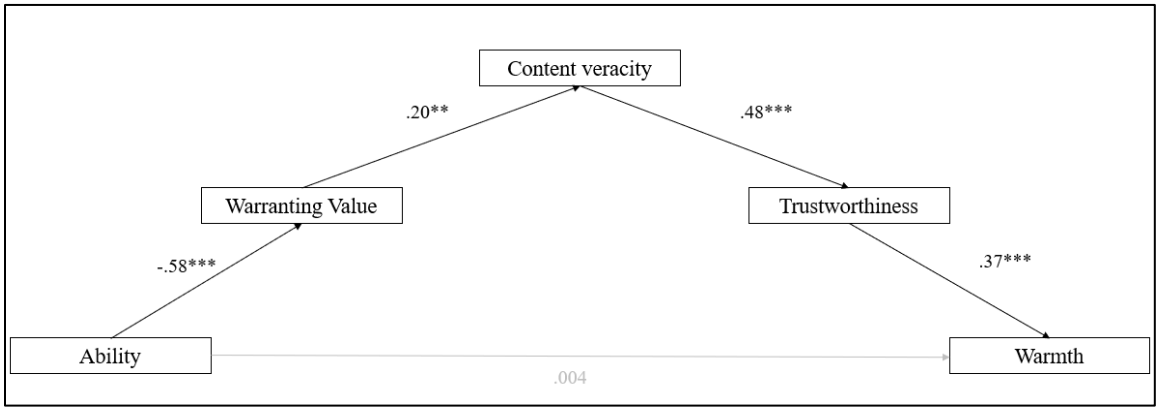

Figure 6. The truth default theory model test using a continuous-level measure of PATM

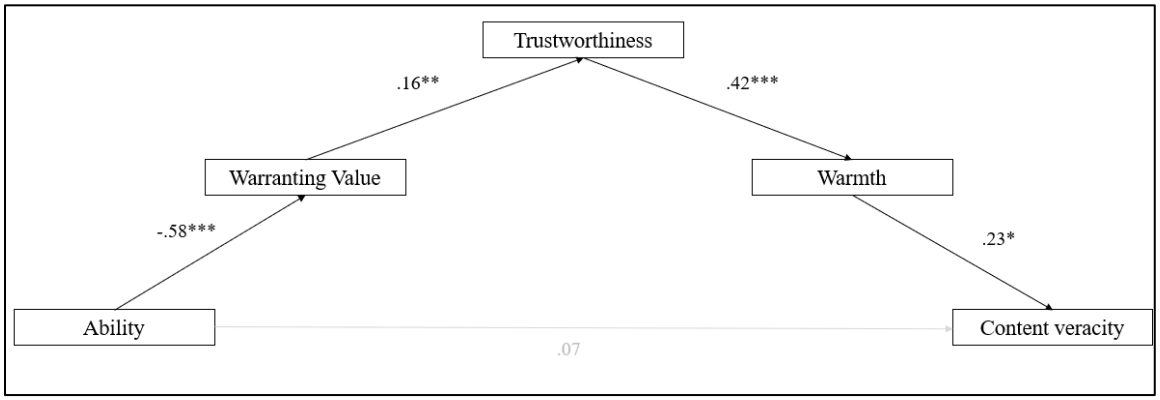

Figure 7. The warranting theory model test using a continuous-level measure of PATM

The use of measured items may provide the added context needed to adequately make sense of the data. Given the correlation data and the strengthened manipulations in Study 2, it is promising to see the predicted relationships start to appear. While the ability induction was successful in Study 2, only people's perceptions of PATM were significantly explanatory in the serial mediation model. These effects will be explored in more detail in the next section, with particular consideration for the future of these lines of research.

Taken together, the results of Study 1 and Study 2 were inconclusive yet informative. These innovative studies sought to explore a variety of new relationships. While no theory or model wholly explained the relationships between the variables, valuable insight was gleaned. The use of post-hoc tests may have provided useful insight for next steps in these lines of research. The next section discusses some of the limitations in this study and provides recommendations for researchers moving forward.

Limitations and Future Directions

The discussion thus far has examined the theoretical contributions of Study 1 and Study 2 but has yet to discuss the potential for future research. This section outlines several of the limitations from the studies and several suggestions for future empirical work to examine these phenomena.

Perceptions of motivation proved difficult to manipulate effectively. Two different attempts were made in Study 1 and Study 2 but were generally unsuccessful. This was complicated by a ceiling effect and other potential sources of perceived motivation stemming from the online dating context. Perceptions of warranting value are also difficult to assess. In some cases, without being prompted, these feelings of perceived manipulability would not exist. First, PMTM and PATM should be examined in other contexts beyond online dating. Many of these effects were observed in politically oriented content (e.g., DeAndrea & Vendemia, 2019; Neo and Johnson, 2020). Specifically, perceptions of motivation between in and out-party groups seemed to drive impressions. As such, this context may highlight the relationships between the study variables more clearly. There are other communicative opportunities to test these relationships in contexts such as social media posts or online reviews.

Perceived motivation to manipulate may come from a variety of sources. As mentioned previously, PMTM in online dating may come from perceptions of motivation to manipulate in order *to get a date.* This study emphasized perceived motivation to manipulate in order to *win a prize*, which was unsuccessful. Future researchers could leverage the motivation to get a date in future research and thus examine its effect on factors such as warranting value, trustworthiness, content veracity, and warmth. A good first step may be a study that reflects the data found in Toma and colleagues' (2008) research on online dating deception. Knowing how and why

individuals lie on their online dating profiles can inform the experimental manipulations in future work. For example, since men lie more often about their height, PMTM about one's height on an online dating profile could be experimentally induced (e.g., with a vignette). Outside of the online dating context, more salient motivations to manipulate may be used. As seen in work by Neo and Johnson (2020) and Vendemia et al., (2019) political information appears to influence impressions, specifically when out-party information is presented. To examine the effects of PMTM on warranting and truth default theories' processes, a political context could be useful.

Beyond the context of the study, experimenters should consider how they will design their studies and manipulate the independent variables of interest. Careful design of experiments related to warranting theory and truth default theory should be curated in order to make small yet meaningful strides. This study applied two theories in new ways and predicted competing hypotheses while also placing the research in a relatively underexplored context. As such, several suggestions are presented. Specifically for warranting value, it may be of some benefit to isolate specific elements of perceived manipulability as a factor rather than an outcome. For example, instead of asking a person to report their perceived warranting value regarding online content, researchers can evaluate the effects varying levels of manipulability have on impressions. In this way, direct tests of the warranting principle can be ascertained and evaluated. This may also be valuable in ascertaining the validity of the warranting instruments. If, as seen in this study, a continuous level measure of warranting value does not accurately represent the independent variable manipulation, the measures may need to be reworked or reimagined,

A significant serial mediation model occurred using continuous level measure of PATM, but not the dummy coded experimental condition variable. These results were only found in

Study 2, where a wider population of US adults participated, and the experimental manipulation was strengthened. It was clear, especially in Study 2, that the PATM induction was successful. However, the continuous level PATM item provides the variance needed to explore these relationships further. In this way, we can explore how the differing values of PATM explain or predict the other study variables. The manipulation alone was not enough to explain the predicted serial mediation model. Instead, the individual perceptions people reported were much stronger predictors. The nuances captured in a perceptive measure of PATM are exciting for future research, particularly in terms of warranting theory and CMC. Results of the post hoc tests indicate that the *perceptions* of ability to manipulate are important. Capturing people's perceptions in a measure held more predictive power in the model. This provides support for arguments made earlier in this paper: there is a difference in the actual behavior and a perception (i.e., ability or motivation to manipulate). A cross-sectional survey may provide extremely useful information. For example, perceptions of PMTM and PATM could be evaluated wholistically regarding an entire genre of mediated communication (e.g., social media or online dating sites), or within platform (e.g., Twitter or Tinder). Doing so may also supplement the research on truth default theory and the relationship between warranting theory. For example, if people evaluate high levels of PATM overall on online dating websites, we could evaluate that relationship with perceptions of trustworthiness and content veracity. Insights from this type of work would be influential in experimental design of future work in this area.

Additionally, research considering warranting theory has been somewhat limited in scope. Much of the research conducted on warranting theory has been in the online review setting and utilizes a first- and third-person approach to perceptions of manipulability. As such, many of the predictions in this study heavily relied on the effects presented in online review

research. Perhaps, warranting effects seen in these studies are beholden to the context. One potential avenue to explore is to first examine PATM and PMTM in the online review setting. This context allows for malleable stimuli that can explore how small changes in result in differing interpretations and impressions. However, researchers are encouraged to move out of the online review setting and into interpersonal communication, social media, AI generated information, and other emerging online communication in order to study warranting theory. Many of the initial propositions from Walther and Parks (2002) can be incorporated into these studies in order to gain a more comprehensive understanding of the effects.

The warranting principle states that warranting value is important in impression formation. In this study, an overall affective impression measure of warmth was used to capture a comprehensive range of impressions of a source. This has had success in previous studies (Snyder et al., 1977). Interestingly, the link between warranting value and warmth was not found in this study in either the correlation test nor in the full serial mediation model. Perhaps this measure was not an important impression as a function of warranting value. Instead, it could be other relevant impressions. In the case of online dating, perhaps attractiveness would be more strongly influenced by warranting value. Researchers should be cognizant that an overall measure of affect may not capture the effect of warranting value. Instead, a specific outcome of interest should be examined.

The link between warranting and truth default theory is also worthy of future exploration. This study did not find the predicted relationships, but it that does not mean they do not exist. One suggestion is to examine the relationship between suspicion and warranting value. This study proposed that warranting value could act as a deception detection cue and be useful in impression formation and decisions about content veracity. However, a first step may be to more

thoroughly examine the relationship between feelings of suspicion and feelings of perceived manipulability. As warranting value was not sufficient enough in this study to encourage individuals to abandon their truth default state, perhaps even for warranting a suspicion threshold must be crossed before it is used as an impression formation mechanism.

There is one major consideration of truth default theory that is worthy of further exploration – are there instances in which the truth default does not apply? For example, it could be, especially in contexts such as online dating, a "lie default" state exists. Perhaps instead of assuming the truth, people begin their assessments from a state of deception assumptions. Particularly in cases wherein an individual's goal is to form impressions, assuming that everything in an online dating profile is (at least somewhat) deceiving may still aid users in impression formation. Similarly, on social media it may behoove users to assume a post is mis or dis-information and then go on to seek additional information as needed. From this perspective, a deceptive default state can still be heuristically simple. Instead of suspicion being a triggering event, perhaps desire for information may motivate future information seeking behaviors. We may then draw upon information seeking topics and theories such as expectancy violation (see Burgoon, 2015), motivation information management (see Afifi & Weiner, 2004), and support seeking (e.g., Collins & Feeny, 2000; High & Scharp, 2015) to examine these processes further.

It is also imperative to consider other potential competing influences on people's evaluation of online content, specifically in the online dating context. Perceived attractiveness is one potentially important factor that was not explored in this study. While the images themselves were pre-tested for attractiveness and warmth, perceptions of attractiveness were not accounted for in the models. This was intentional in this study, as it was not part of the proposed theoretical models. However, perhaps attractiveness has a "halo or horns" effect on impression formation

such that if a person is perceived to be more or less attractive, it influences other impressions. Indeed, attractiveness has been shown to influence deception detection for both the person of interest and the participant of the study (DePaulo et al., 1987). Attractive people appear to have different deception detection mechanisms than unattractive people. Exaggerating one's features, telling white lies, or being outright deceptive is commonplace in dating (Rowatt et al., 1999; Toma et al., 2008), as individuals use deception to attempt to secure a date. Perhaps in this case, attractiveness had explanatory power that was not accounted for. However, it is important to consider where in a theoretical model attractiveness asserts influence. Attractiveness could be an impression variable (and therefore an outcome of warranting value assessments), but as described above it could also be a driver of causal processes (a predictor of veracity judgements). To this end, careful theoretical rationale should be used in determining the causal ordering of attractiveness in research that involves truth default and warranting theory.

This study, as in previous research (e.g., Whyte et al. 2021) found that male and female individuals were rated differently in terms of warmth and trustworthiness. Women are often rated as more attractive, warmer, and more trustworthy than men, as was seen here. The differences seen in this study are particularly interesting given that the images were AI generated and were similar in face shape and features. Use patterns by men and women also differ in online dating. Kreager et al., (2014) for example, found that women send far fewer messages than men in online dating applications. They also found that men and women have varying success in finding a desirable partner. McGloin and Denes (2016) found differences in how men and women evaluated online dating profiles. For men, highly attractive female profiles were untrustworthy, while for women, highly attractive male profiles were trustworthy. Gender differences should be explored particularly in the online dating context. Due to the aforementioned differences in

online dating use and the observed effects in this study, specific groups of people may have relevant cues that help them assess trustworthiness, content veracity, and affective evaluations such as warmth or attractiveness. This study used only a male and female presenting stimuli that catered to heterosexual norms. This was done intentionally, as the goal of the study was not to examine gender or demographic differences, but to examine warranting theory and TDT mechanisms in online dating. Future research may desire to test the differences in minoritized groups based on their sexual orientation, gender, and race. Doing so will also strengthen our understanding of the theories themselves and how they apply to wider populations while also gaining more intimate knowledge of impression formation processes.

To accomplish this, future research can utilize AI generated images to their advantage. Specifically, as in this study, AI images of different genders can be created with similar face shapes, hair and eye colors, skin tones, and features in order to maintain internal validity as was done in this study. Programs such as Generated Faces could be a very valuable resource moving forward, and researchers are encouraged to utilize AI experimentally. Not only can AI be used in the creation of stimuli, but the impressions of AI and AI-generated content can be used within future research. In this study, AI-generated information was used as a function of information control (i.e., PATM). However, the way people evaluate AI itself should be carefully considered. With the proliferation of AI, particular attention should be given to perceptions of truth and deception. Perceptions of ability was one way of examining trust and veracity judgements, but as AI changes and grows research should continue to explore the way individuals form impressions. This is extremely apparent in online dating. For example, how do people tell if an image is AI generated? Does it matter if the “about me” section is written by AI? Impressions of AI as it

develops should be carefully monitored and evaluated in order to understand the effects it has, especially in an online dating context.

Overall, the ideas and outcomes from the two studies provide numerous opportunities to enhance our understanding of seemingly well-established research (e.g., attribution), theoretically novel research (e.g., warranting theory or truth default theory), as well as new and innovative uses of online contexts (e.g., online dating and AI). The ambitious theoretical predictions and attempts to apply the research to understudied contexts proved challenging. However, in these challenges, a large number of opportunities for future research were revealed.

Conclusion

The two studies presented here sought to understand the relationships between warranting value and the truth default state. It was predicted that PMTM and PATM were drivers of these processes. Bivariate correlations revealed significant relationships between the variables in both Study 1 and Study 2. However, the predicted serial mediation models utilizing warranting theory and truth default theory were not significant. The proposed models were theoretically driven and promising, but the data were not consistent with the predictions. Future research should seek to isolate the factors presented in the study in order to better understand the manner in which these mechanisms persist. These results notwithstanding, this study provided new and interesting avenues for future research in CMC, specifically in online dating.

REFERENCES

Afifi, W. A., & Weiner, J. L. (2004). Toward a Theory of Motivated Information Management. Communication Theory, 14(2), 167–190. https://doi.org/10.1111/j.1468-2885.2004.tb00310.x

Ajzen, I. (1991). The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2006). Constructing a theory of planned behavior questionnaire. Retrieved from: http://people.umass.edu/aizen/pdf/tpb.measurement.pdf

Albury, K., & Byron, P. (2016). Safe on My Phone? Same-Sex Attracted Young People's Negotiations of Intimacy, Visibility, and Risk on Digital Hook-Up Apps. Social Media + Society, 2(4), 2056305116672887. https://doi.org/10.1177/2056305116672887

Ansari, S., Gupta, S., & Dewangan, J. (2018). Do Customers Perceive Reviews as Manipulated? A Warranting Theory Perspective. 16.

Antheunis, M. L., & Schouten, A. P. (2011). The Effects of Other-Generated and System-Generated Cues on Adolescents' Perceived Attractiveness on Social Network Sites. Journal of Computer-Mediated Communication, 16(3), 391–406. https://doi.org/10.1111/j.1083-6101.2011.01545.x

Bennett, L. W. (2007). Civic Learning in Changing Democracies. In P. Dahlgren (Ed.), *Young citizens and new media: Learning for democratic participation* (pp. 59-78). Routledge.

Blair, J. P., Reimer, T. O., & Levine, T. R. (2018). The role of consistency in detecting deception: The superiority of correspondence over coherence. Communication Studies, 69(5), 483-498. https://doi.org/10.1080/10510974.2018.1447492

Bokek-Cohen, Y. (2015). How do anonymous sperm donors signal credibility through their self-presentations? Health Sociology Review, 24(1), 81–93. https://doi.org/10.1080/14461242.2015.1004800

Brandt, A. C., Vonk, R., & van Knippenberg, A. (2011). Augmentation and Discounting in Impressions of Targets Described by Third Parties with Ulterior Motives. Social Cognition, 29(2), 210–220. https://doi.org/10.1521/soco.2011.29.2.210

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. Personality and Individual Differences, 67, 97–102. https://doi.org/10.1016/j.paid.2014.01.016

Burgoon, J. K. (2015). Expectancy Violations Theory. In The International Encyclopedia of Interpersonal Communication (pp. 1–9). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118540190.wbeic102

Burgoon, M., Miller, M. D., Cohen, M., & Montgomery, C. L. (1978). An empirical test of a model of resistance to persuasion. Human Communication Research, 5(1), 27-39. https://doi.org/10.1111/j.1468-2958.1978.tb00620.x

Cali, B. E., Coleman, J. M., & Campbell, C. (2013). Stranger Danger? Women's Self-Protection Intent and the Continuing Stigma of Online Dating. Cyberpsychology, Behavior, and Social Networking, 16(12), 853–857. https://doi.org/10.1089/cyber.2012.0512

Chadwick, A., & Vaccari, C. (n.d.). News Sharing on UK Social Media: Misinformation, Disinformation, and Correction. 32.

Clare, D. D., & Levine, T. R. (2019). Documenting the Truth-Default: The Low Frequency of Spontaneous Unprompted Veracity Assessments in Deception Detection. Human Communication Research, 45(3), 286–308. https://doi.org/10.1093/hcr/hqz001

Coles, B. A., & West, M. (2016). Trolling the trolls: Online forum users constructions of the nature and properties of trolling. Computers in Human Behavior, 60, 233–244. https://doi.org/10.1016/j.chb.2016.02.070

Collins, N. L., & Feeney, B. C. (n.d.). A Safe Haven: An Attachment Theory Perspective on Support Seeking and Caregiving in Intimate Relationships.

DeAndrea, D. C. (2014). Advancing Warranting Theory: Advancing Warranting Theory. Communication Theory, 24(2), 186–204. https://doi.org/10.1111/comt.12033

DeAndrea, D. C., & Carpenter, C. J. (2018). Measuring the Construct of Warranting Value and Testing Warranting Theory. Communication Research, 45(8), 1193–1215. https://doi.org/10.1177/0093650216644022

DeAndrea, D. C., Tong, S. T., & Lim, Y. (2018). What causes more mistrust: Profile owners deleting user-generated content or website contributors masking their identities? Information, Communication & Society, 21(8), 1068–1080. https://doi.org/10.1080/1369118X.2017.1301523

DeAndrea, D. C., Van Der Heide, B., & Easley, N. (2015). How Modifying Third-Party Information Affects Interpersonal Impressions and the Evaluation of Collaborative Online Media: Influence of Third-Party Content on Impressions. Journal of Communication, 65(1), 62–78. https://doi.org/10.1111/jcom.12139

DeAndrea, D. C., Van Der Heide, B., Vendemia, M. A., & Vang, M. H. (2018). How People Evaluate Online Reviews. Communication Research, 45(5), 719–736. https://doi.org/10.1177/0093650215573862

DeAndrea, D. C., & Vendemia, M. A. (2019). The Influence of Self-Generated and Third-Party Claims Online: Perceived Self-Interest as an Explanatory Mechanism. Journal of Computer-Mediated Communication, 24(5), 223–239. https://doi.org/10.1093/jcmc/zmz011

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The Truth About the Truth: A Meta-Analytic Review of the Truth Effect. Personality and Social Psychology Review, 14(2), 238–257. https://doi.org/10.1177/1088868309352251

DePaulo, B. M., Tang, J., & Stone, J. I. (1987). Physical Attractiveness and Skill at Detecting Deception. Personality and Social Psychology Bulletin, 13(2), 177–187. https://doi.org/10.1177/014616728713200

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. Journal of Computer-Mediated Communication, 11(2), 415–441. https://doi.org/10.1111/j.1083-6101.2006.00020.x

Flynn, D. j., Nyhan, B., & Reifler, J. (2017). The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics. Political Psychology, 38(S1), 127–150. https://doi.org/10.1111/pops.12394

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. Psychological Bulletin, 117(1), 21–38. https://doi.org/10.1037/0033-2909.117.1.21

Hanel, P. H. P., & Vione, K. C. (2016). Do Student Samples Provide an Accurate Estimate of the General Public? PLoS ONE, 11(12), https://doi.org/10.1371/journal.pone.0168354

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. Communication Monographs, 85(1), 4–40. https://doi.org/10.1080/03637751.2017.1352100

Heider, F. (1958). The Psychology of Interpersonal Relationships. New York: Wiley.

High, A. C., & Scharp, K. M. (2015). Examining Family Communication Patterns and Seeking Social Support Direct and Indirect Effects Through Ability and Motivation. Human Communication Research, 41(4), 459–479. https://doi.org/10.1111/hcre.12061

Hovland, C. I., & Weiss, W. (1951). The Influence of Source Credibility on Communication Effectiveness*. Public Opinion Quarterly, 15(4), 635–650. https://doi.org/10.1086/266350

Jalbert, M., Newman, E., & Schwarz, N. (2020). Only Half of What I'll Tell You is True: Expecting to Encounter Falsehoods Reduces Illusory Truth. Journal of Applied Research in Memory and Cognition, 9(4), 602–613. https://doi.org/10.1016/j.jarmac.2020.08.010

Jones, E. E., & Davis, K. E. (1965). From Acts to Dispositions the Attribution Process in Person Perception11Much of the research reported herein was supported by National Science Foundation Grants 8857 and 21955 to the first author. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 2, pp. 219–266). Academic Press. https://doi.org/10.1016/S0065-2601(08)60107-0

Jones, E. E., & Nisbett, R. E. (1971). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), Attribution: Perceiving the causes of behavior (pp.79–94). Morristown, NJ: General Learning Press.

Kelley, H.H. (1967) Attribution theory in social psychology. In D. Levine (ed.), Nebraska Symposium on Motivation, pp. 192–238. Lincoln: University of Nebraska Press.

Kreager, D. A., Cavanagh, S. E., Yen, J., & Yu, M. (2014). "Where Have All the Good Men Gone?" Gendered Interactions in Online Dating: Gendered Interactions in Online Dating. Journal of Marriage and Family, 76(2), 387–410. https://doi.org/10.1111/jomf.12072

Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the Currency of Democratic Citizenship. Journal of Politics, 62(3), 790–816. https://doi.org/10.1111/0022-3816.00033

Kunda, Z. (1990). The case for motivated reasoning. Psychological Bulletin, 108(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Levine, T. R., Kim, R. K., & Hamel, L. M. (2010). People Lie for a Reason: Three Experiments Documenting the Principle of Veracity. Communication Research Reports, 27(4), 271–285. https://doi.org/10.1080/08824096.2010.496334

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." Communication Monographs, 66(2), 125–144. https://doi.org/10.1080/03637759909376468

Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. Psychological Bulletin, 132(6), 895–919. https://doi.org/10.1037/0033-2909.132.6.895

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

Manusov, V., & Spitzberg, B. (2008). Attribution theory. In Engaging theories in interpersonal communication: Multiple perspectives (pp. 37–49). Sage Publications, Inc. https://doi.org/10.4135/9781483329529.n3

McCroskey, J. C., & Richmond, V. P. (1989). Bipolar scales. In P. Emmert & L. L. Barker (Eds.), Measurement of Communication Behavior (pp. 154-167). New York: Longman.

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. Communications Monographs, 66(1), 90-103. https://doi.org/10.1080/03637759909376464

Menkin, J. A., Robles, T. F., Wiley, J. F., & Gonzaga, G. C. (2015). Online dating across the life span: Users' relationship goals. Psychology and Aging, 30(4), 987–993. https://doi.org/10.1037/a0039722

Morosoli, S., Van Aelst, P., Humprecht, E., Staender, A., & Esser, F. (2022). Identifying the Drivers Behind the Dissemination of Online Misinformation: A Study on Political Attitudes and Individual Characteristics in the Context of Engaging With Misinformation on Social Media. American Behavioral Scientist, 000276422211183. https://doi.org/10.1177/00027642221118300

Neo, R. L., & Johnson, B. K. (2020). Online products and consumers: Partisan ratings and mechanisms for affective polarization. Telematics and Informatics, 54, 101467. https://doi.org/10.1016/j.tele.2020.101467

Parks, M. R. (2011). Boundary Conditions for the Application of Three Theories of Computer-Mediated Communication to MySpace. Journal of Communication, 61(4), 557–574. https://doi.org/10.1111/j.1460-2466.2011.01569.x

Porter, S. R., Whitcomb, M. E., & Weitzer, W. H. (2004). Multiple surveys of students and survey fatigue. New Directions for Institutional Research, 2004(121), 63–73. https://doi.org/10.1002/ir.101

Ramirez, A. (2007). The Effect of Anticipated Future Interaction and Initial Impression Valence on Relational Communication in Computer-Mediated Interaction. *Communication Studies*, *58*(1), 53–70. https://doi.org/10.1080/10510970601168699

Rotter, J. B. (1980). Interpersonal trust, trustworthiness, and gullibility. *American Psychologist*, *35*(1), 1–7. https://doi.org/10.1037/0003-066X.35.1.1

Rui, J. R. (2018). Source–Target Relationship and Information Specificity: Applying Warranting Theory to Online Information Credibility Assessment and Impression Formation. *Social Science Computer Review*, *36*(3), 331–348. https://doi.org/10.1177/0894439317717196

Scannell, M. J. (2019). Online Dating and the Risk of Sexual Assault to College Students. Building Healthy Academic Communities Journal, 3(1), 34. https://doi.org/10.18061/bhac.v3i1.6688

Scherer, L. D., & Pennycook, G. (2020). Who Is Susceptible to Online Health Misinformation? American Journal of Public Health, 110(S3), S276–S277. https://doi.org/10.2105/AJPH.2020.305908

Shin, S. Y., Dai, Y. (Nancy), Beyea, D., Prchal, B., Makki, T. W., Schlafhauser, K., & Van Der Heide, B. (2020). Curbing Negativity: Influence of Providing Justifications About Control Over User-Generated Comments on Social Media. Communication Research, 47(6), 838–859. https://doi.org/10.1177/0093650218794853

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. Journal of Personality and Social Psychology, 35(9), 656–666. https://doi.org/10.1037/0022-3514.35.9.656

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. Journal of Personality and Social Psychology, 89(6), 845–851. https://doi.org/10.1037/0022-3514.89.6.845

Tao, H. (2022). Loving strangers, avoiding risks: Online dating practices and scams among Chinese lesbian (lala) women. Media, Culture & Society, 44(6), 1199–1214. https://doi.org/10.1177/01634437221088952

Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating Fact From Fiction: An Examination of Deceptive Self-Presentation in Online Dating Profiles. Personality and Social Psychology Bulletin, 34(8), 1023–1036. https://doi.org/10.1177/0146167208318067

Utz, S. (2010). Show me your friends and I will tell you what type of person you are: How one's profile, number of friends, and type of friends influence impression formation on social network sites. Journal of Computer-Mediated Communication, 15(2), 314–335. https://doi.org/10.1111/j.1083-6101.2010.01522.x

Van Der Heide, B., Johnson, B. K., & Vang, M. H. (2013). The effects of product photographs and reputation systems on consumer behavior and product cost on eBay. Computers in Human Behavior, 29(3), 570–576. https://doi.org/10.1016/j.chb.2012.11.002

Van Der Zanden, T., Mos, M. B. J., Schouten, A. P., & Krahmer, E. J. (2022). What People Look at in Multimodal Online Dating Profiles: How Pictorial and Textual Cues Affect Impression Formation. Communication Research, 49(6), 863–890. https://doi.org/10.1177/0093650221995316

Vendemia, M. A., Bond, R. M., & DeAndrea, D. C. (2019). The strategic presentation of user comments affects how political messages are evaluated on social media sites: Evidence for robust effects across party lines. Computers in Human Behavior, 91, 279–289. https://doi.org/10.1016/j.chb.2018.10.007

Vendemia, M. A., Coduto, K. D., & Lee-Won, R. J. (2019). When are emotional online product reviews persuasive? Exploring the effects of reviewer status and valence in consumer purchase decisions. Communication Research Reports, 36(5), 437–448. https://doi.org/10.1080/08824096.2019.1683532

Vogels, E. A, & Mcclain, C. (2023). Key findings about online dating in the U.S. Pew Research Center. Retrieved July 4, 2023, from https://www.pewresearch.org/short-reads/2023/02/02/key-findings-about-online-dating-in-the-u-s/

Walther, J. B. (1992). Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective. Communication Research, 19(1), 52–90. https://doi.org/10.1177/009365092019001003

Walther, J. B. (1994). Anticipated Ongoing Interaction Versus Channel Effects on Relational Communication in Computer-Mediated Interaction. Human Communication Research, 20(4), 473–501. https://doi.org/10.1111/j.1468-2958.1994.tb00332.x

Walther, J. B. (1996). Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction. Communication Research, 23(1), 3–43. https://doi.org/10.1177/009365096023001001

Walther, J. B., & Jang, J. (2012). Communication Processes in Participatory Websites. Journal of Computer-Mediated Communication, 18(1), 2–15. https://doi.org/10.1111/j.1083-6101.2012.01592.x

Walther, J. B., Van Der Heide, B., Hamel, L. M., & Shulman, H. C. (2009). Self-Generated Versus Other-Generated Statements and Impressions in Computer-Mediated Communication: A Test of Warranting Theory Using Facebook. Communication Research, 36(2), 229–253. https://doi.org/10.1177/0093650208330251

Walther, J. B., Van Der Heide, B., Kim, S.-Y., Westerman, D., & Tong, S. T. (2008). The Role of Friends' Appearance and Behavior on Evaluations of Individuals on Facebook: Are We Known by the Company We Keep? Human Communication Research, 34(1), 28–49. https://doi.org/10.1111/j.1468-2958.2007.00312.x

Walther, J., & Parks, M. (2002). Cues Filtered Out, Cues Filtered In: Computer-Mediated Communication and Relationships.

Weiner, B. (1982). An attribution theory of motivation and emotion. Series in Clinical & Community Psychology: Achievement, Stress, & Anxiety, 223–245.

Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social Media as Information Source: Recency of Updates and Credibility of Information. Journal of Computer-Mediated Communication, 19(2), 171–183. https://doi.org/10.1111/jcc4.12041

Whyte, S., Brooks, R. C., Chan, H. F., & Torgler, B. (2021). Sex differences in sexual attraction for aesthetics, resources and personality across age. PLoS ONE, 16(5), e0250151. https://doi.org/10.1371/journal.pone.0250151

Wu, Y., & Kelly, R. M. (2020). Online Dating Meets Artificial Intelligence: How the Perception of Algorithmically Generated Profile Text Impacts Attractiveness and Trust. 32nd Australian Conference on Human-Computer Interaction, 444–453. https://doi.org/10.1145/3441000.3441074

APPENDIX

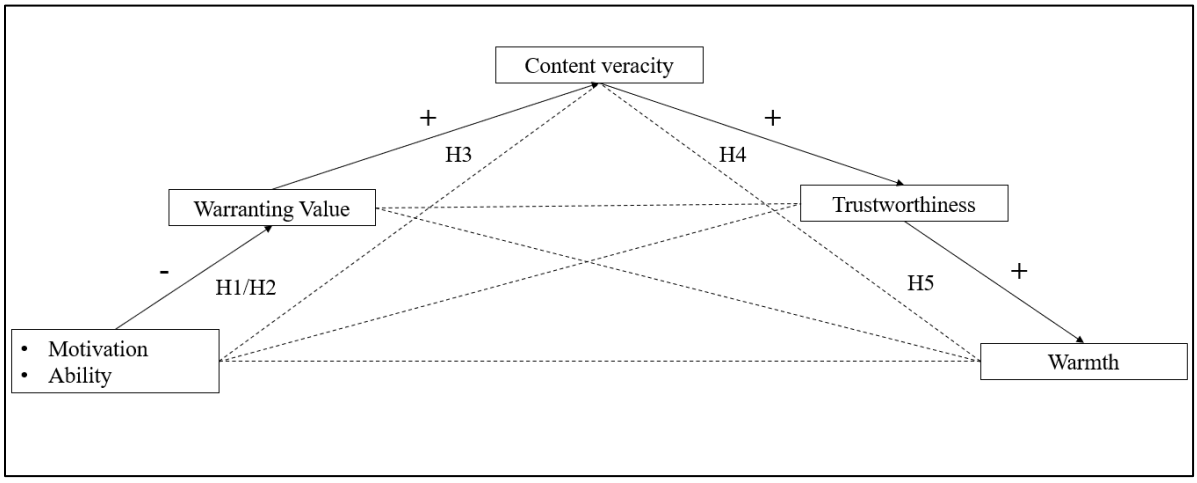

Figure A1. Warranting theory approach. Moderated serial mediation model, Study 1

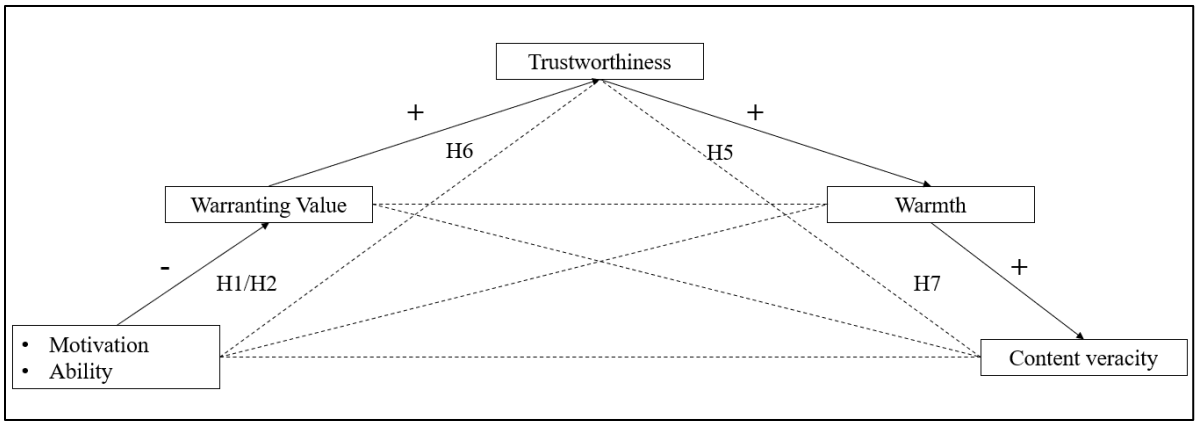

Figure A2. Truth default theory approach. Moderated serial mediation model, Study 1

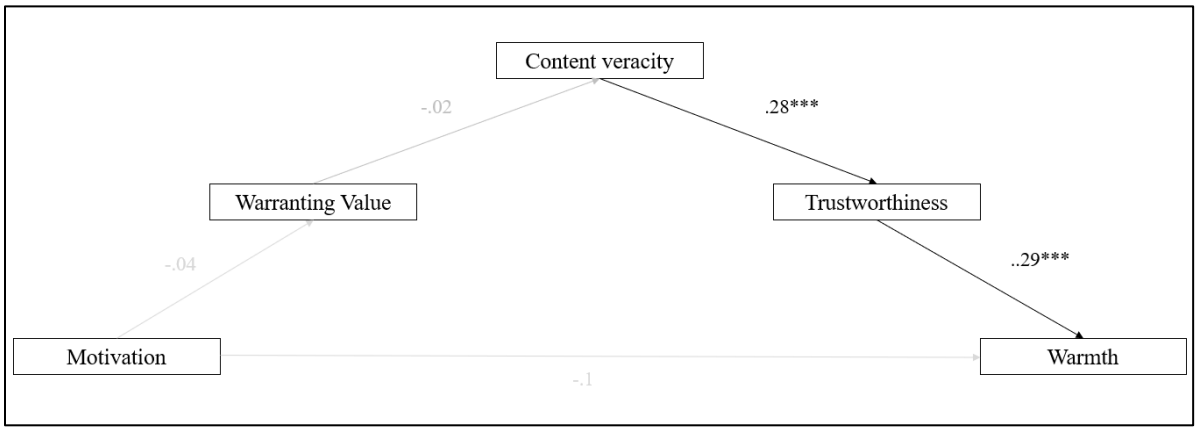

Figure A3. Significant and nonsignificant pathways in H8, Study 1

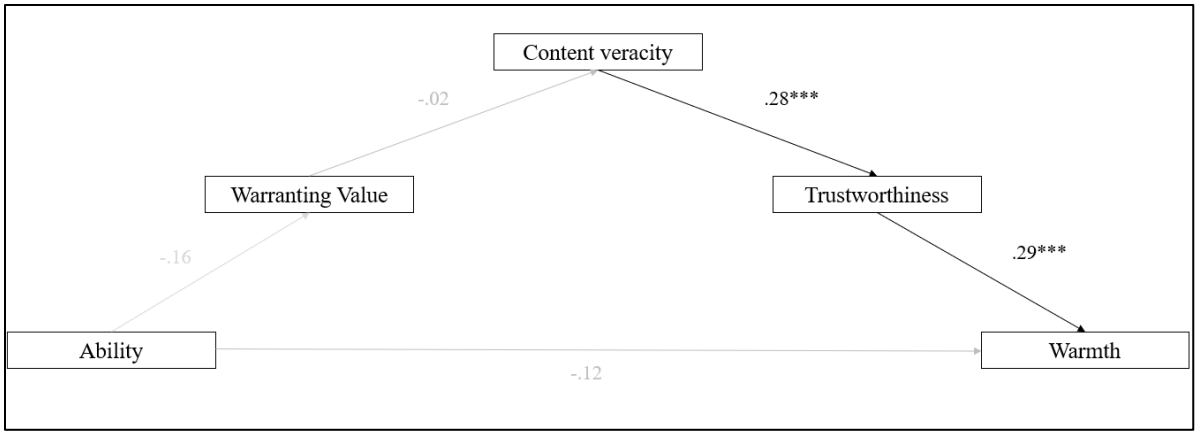

Figure A4. Significant and nonsignificant pathways in H9, Study 1

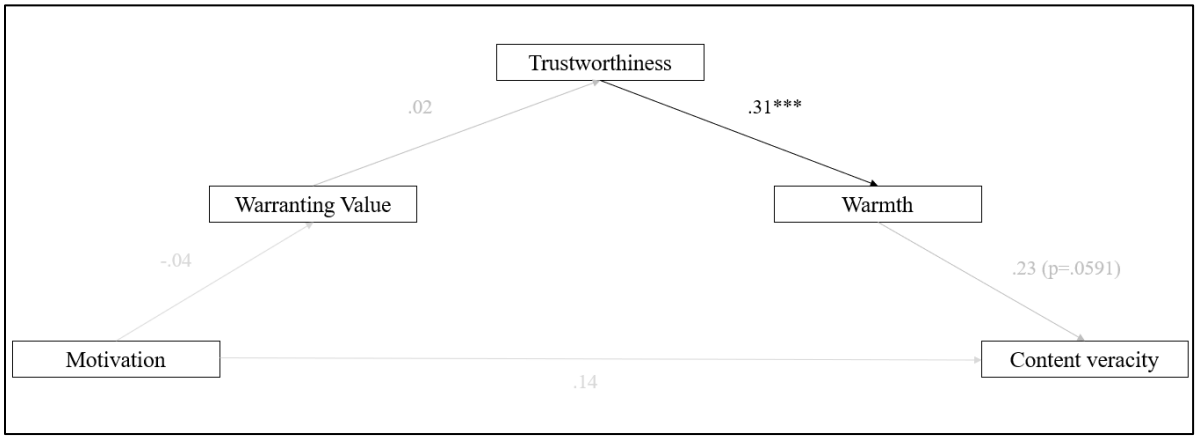

Figure A5. Significant and nonsignificant pathways in H10, Study 1

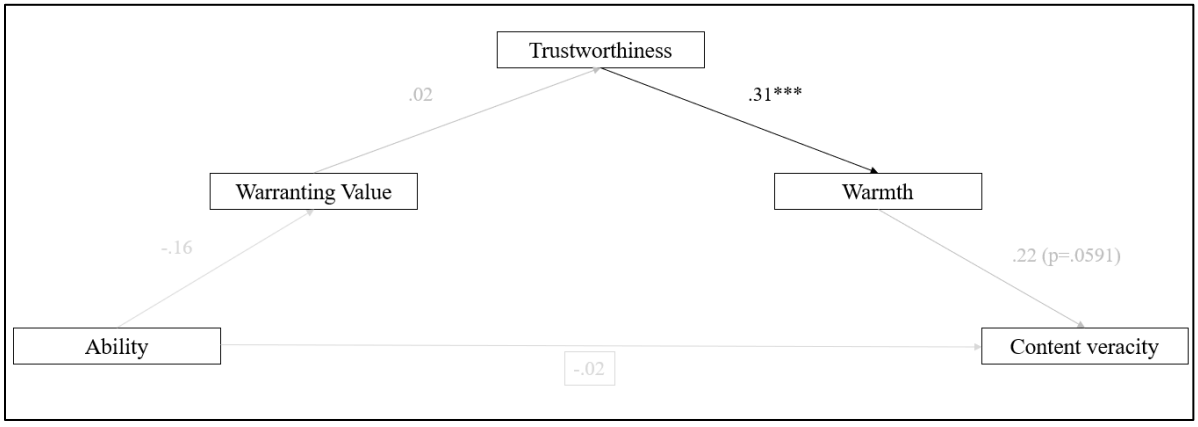

Figure A6. Significant and nonsignificant pathways in H11, Study 1

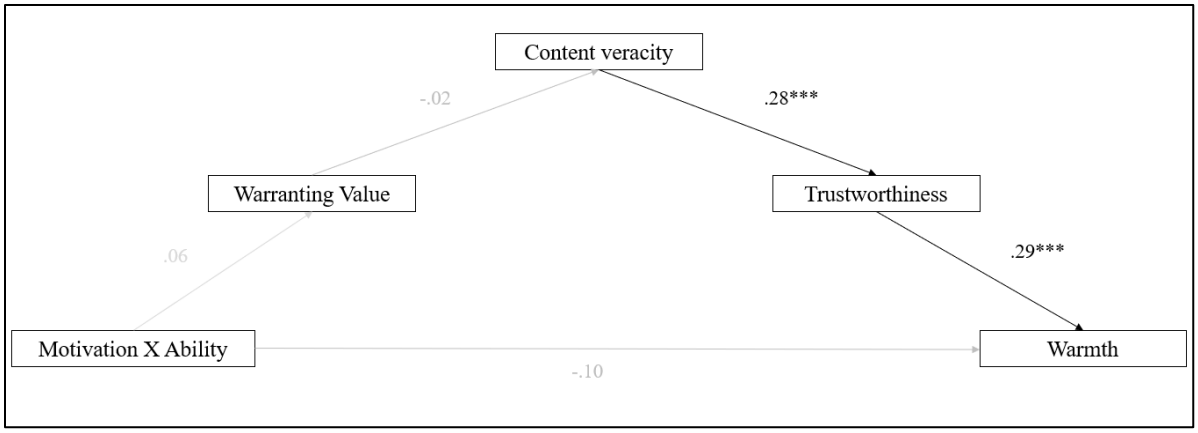

Figure A7. Significant and nonsignificant pathways in RQ2 (warranting theory model), Study 1

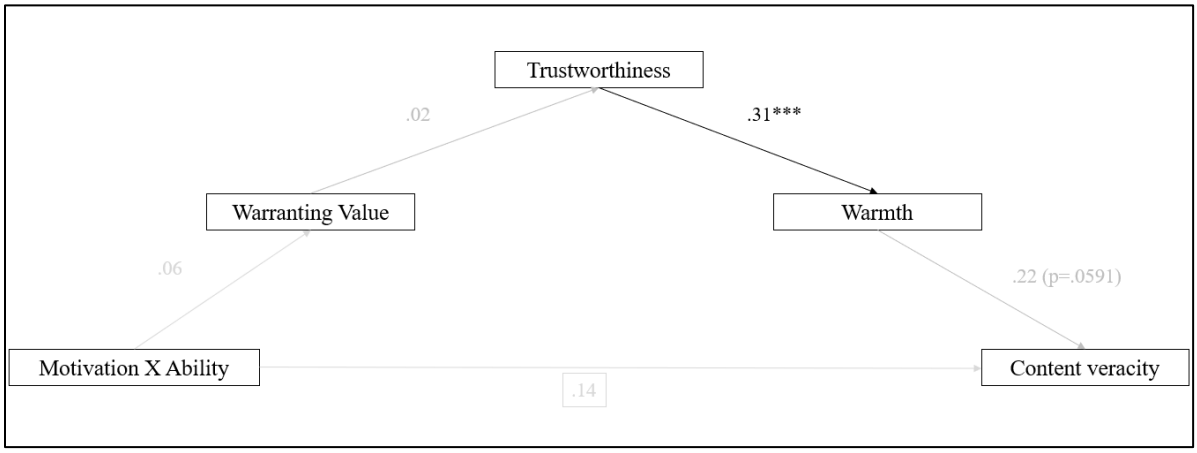

Figure A8. Significant and nonsignificant pathways in RQ2 (TDT model), Study 1

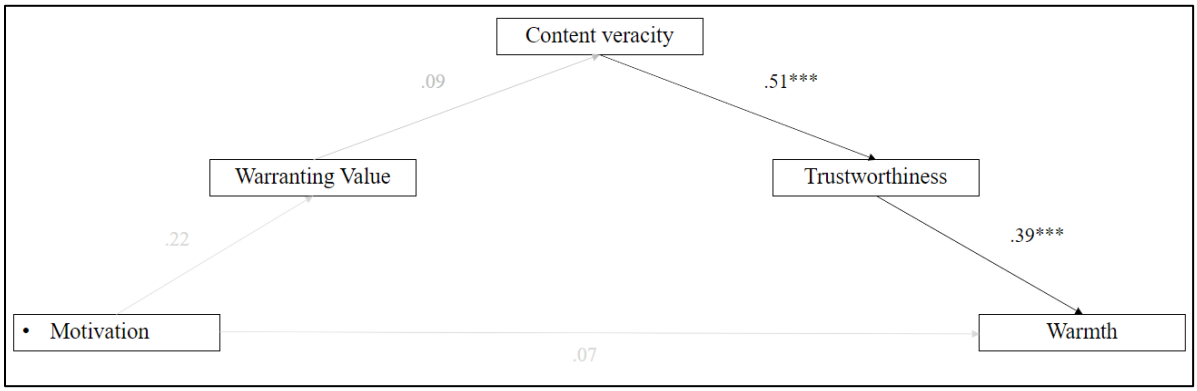

Figure A9. Significant and nonsignificant pathways in H8, Study 2

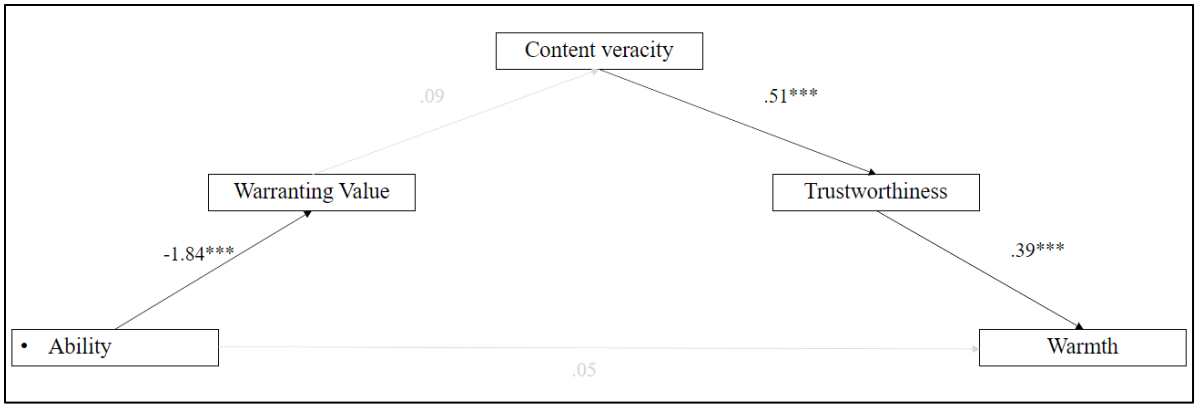

Figure A10. Significant and nonsignificant pathways in H9, Study 2

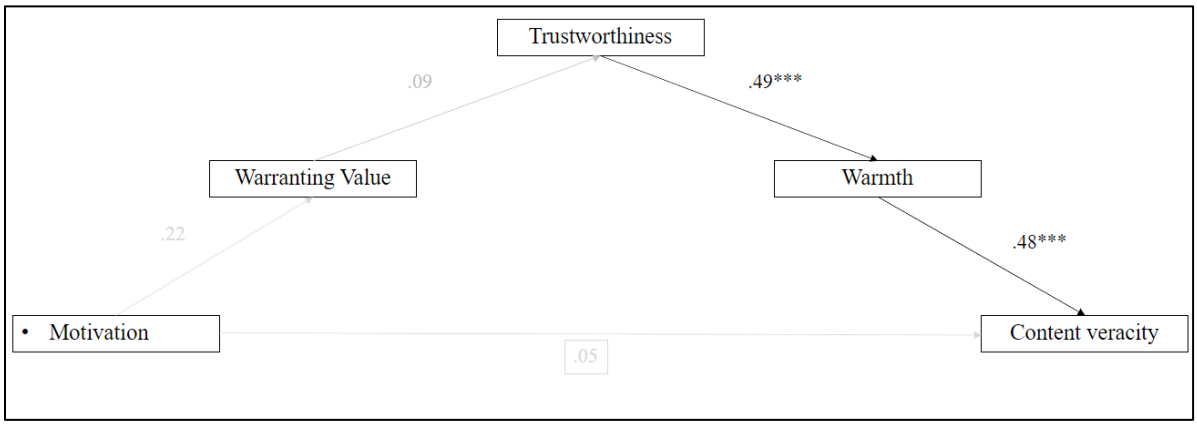

Figure A11. Significant and nonsignificant pathways in H10, Study 2

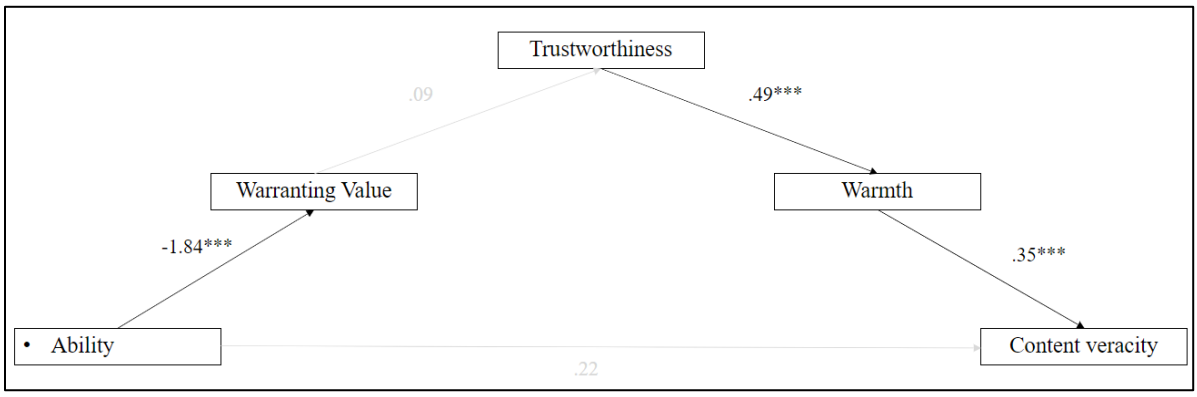

Figure A12. Significant and nonsignificant pathways in H11, Study 2,

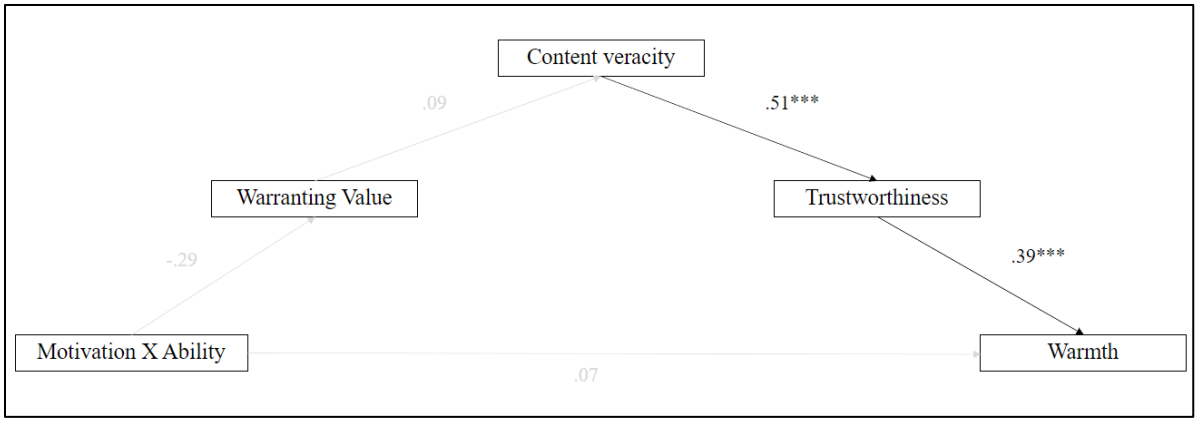

Figure A13. Significant and nonsignificant pathways in RQ2 (warranting theory model), Study 2

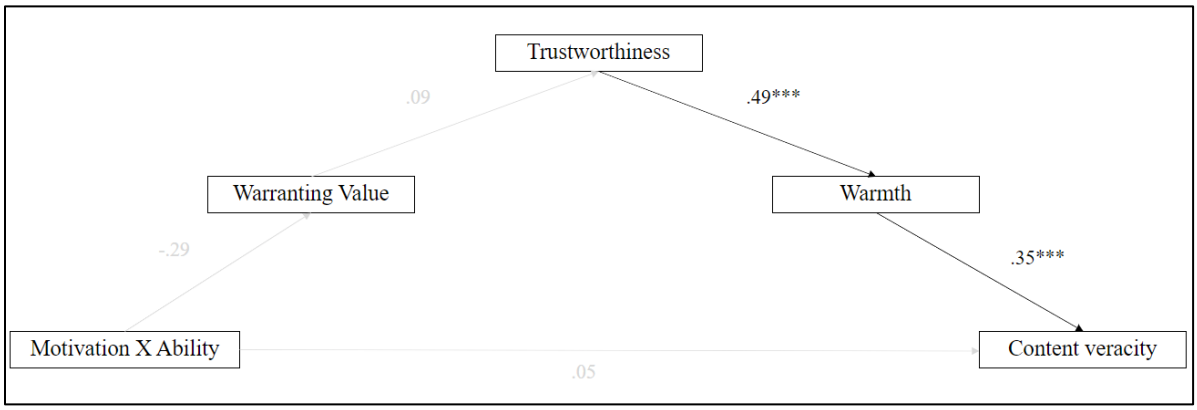

Figure A14. Significant and nonsignificant pathways in RQ2 (TDT model), Study 2

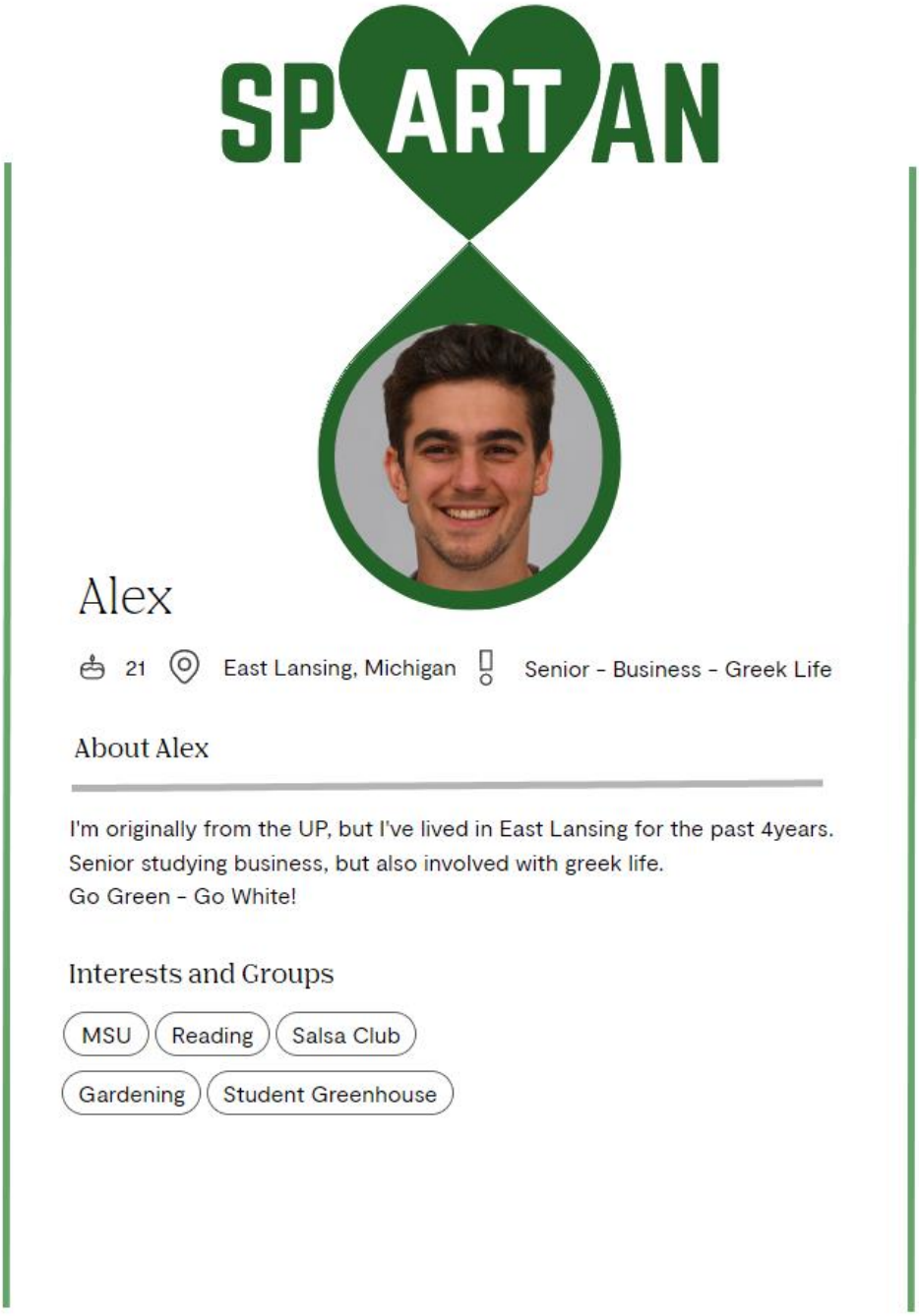

Figure A15. Study 1 male profile

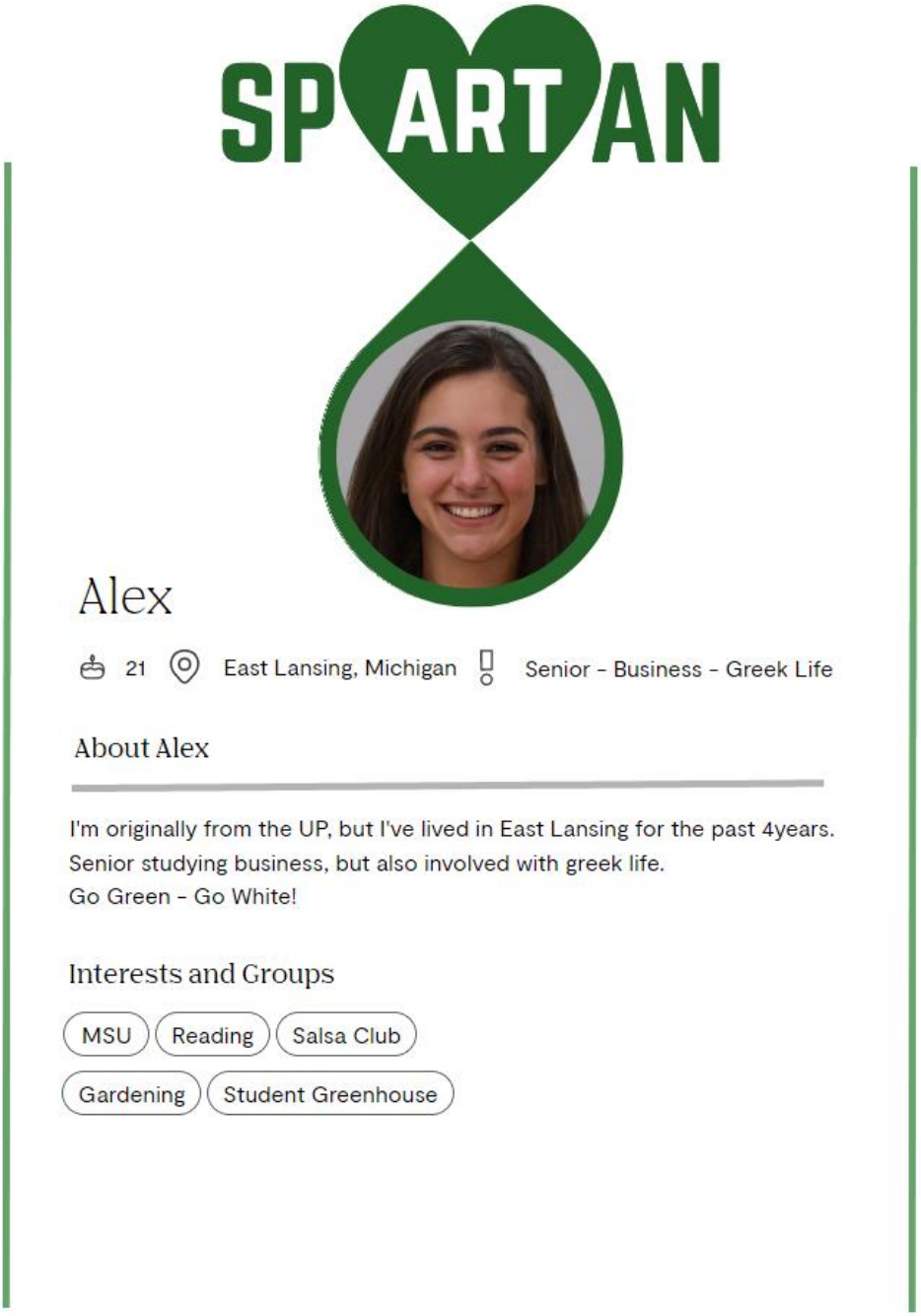

Figure A16. Study 1 female profile

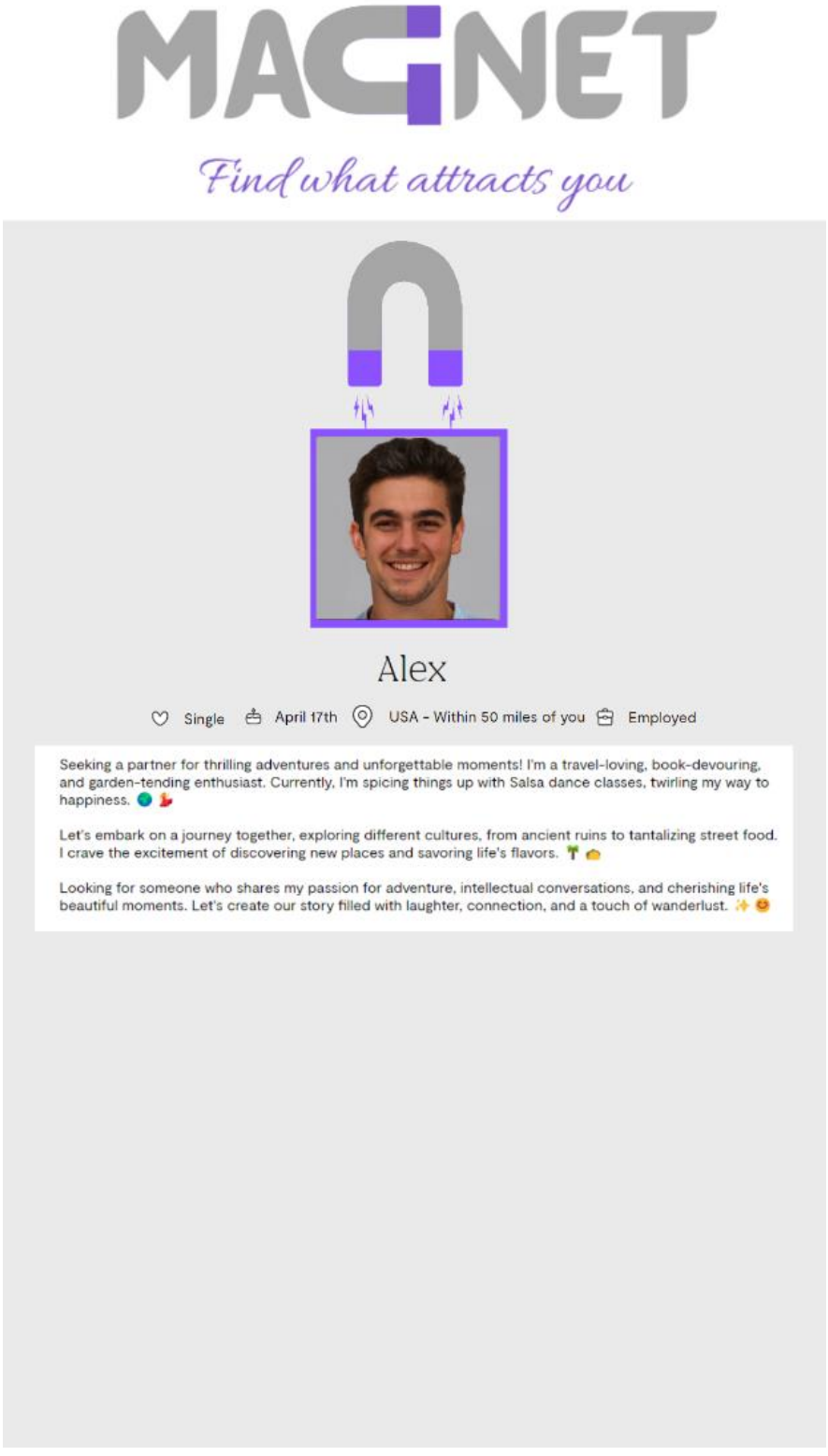

Figure A17. Study 2 male profile

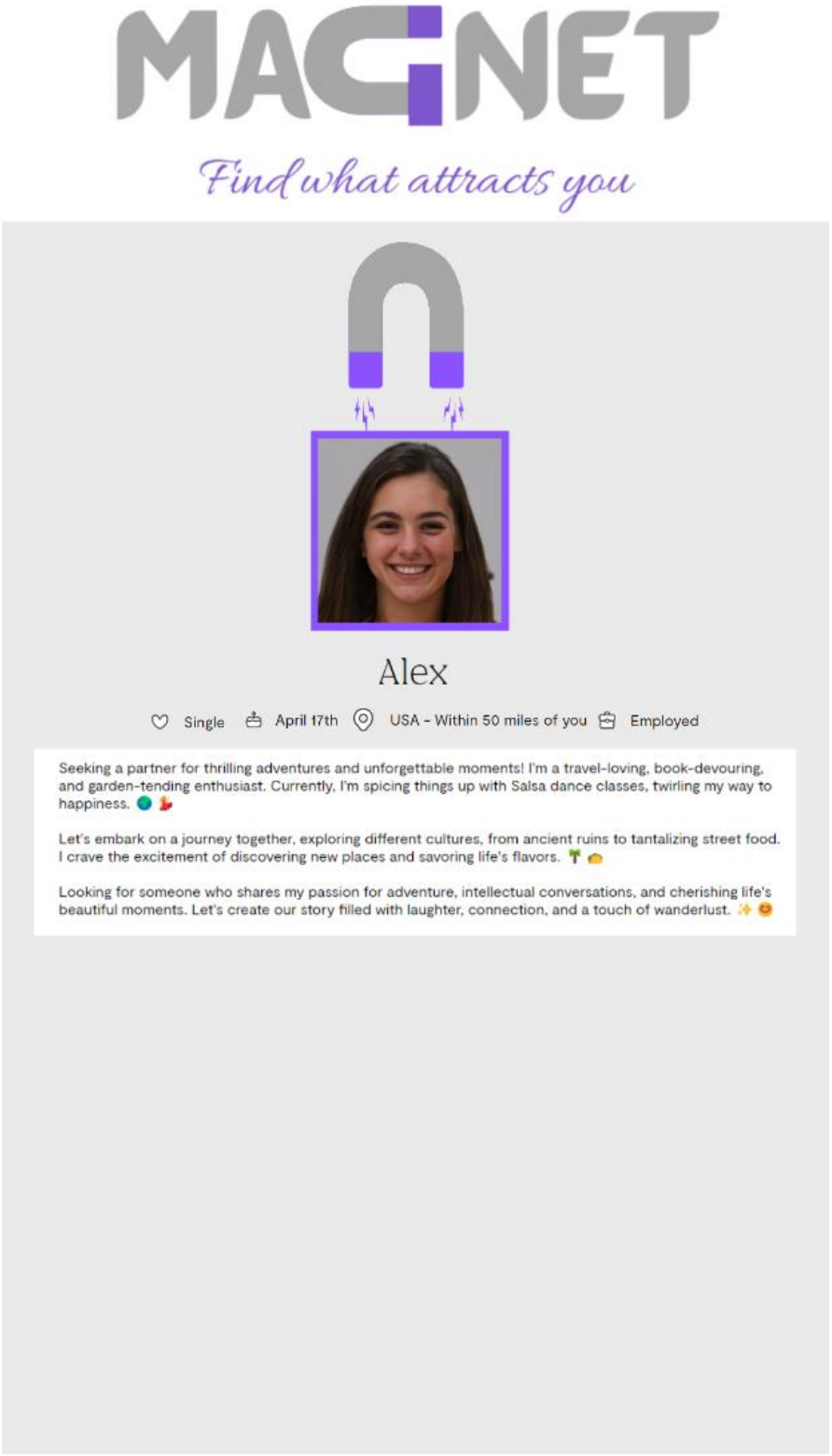

Figure A18. Study 2 female profile