7-29-2024 10:00 AM

The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Mackenzie C. Smith, *Western University*

Supervisor: Tremblay, Paul, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Psychology

Abstract

Interpersonal conflict in ambiguous social situations can instigate aggressive responses in individuals with the hostile attribution bias (HAB). However, the nature of the psychological properties of these situations needs to be explored more extensively, including the specific social-contextual properties. One individual difference that few studies have explored with the HAB is dispositional trust, which is proposed to be the opposing side of HAB. The current vignette study aimed to explore this, along with improving previous measures and creating a new measure of HAB. Factor analyses and multilevel modeling were used for establishing a hostile composite and exploring both individual differences, social-contextual factors, and their interaction on hostile responses. Exploratory factor analysis confirmed the factor structure proposed for a hostile composite (i.e., HAB score), but did not confirm our categorization of general vignette scenarios. Further, agreeableness and trust were found to have significant negative relationships with the HAB, supporting one objective of the study. Honesty-Humility was also explored and had a significant negative relationship with HAB. Finally, perpetrator power was significant, where hostility was found more in scenarios where power dynamics were equal. Limitations and suggestions for future directions are explored. The current study contributes to the understanding of the HAB in different social situations, and the complexities that individual differences and social-contextual factor have on perceptions of hostility.

Keywords

Hostile Attribution Bias; Dispositional Trust; Anger; Personality; Agreeableness; Social-Contextual Factors; Power Dynamics

Summary for Lay Audience

The Hostile Attribution Bias (HAB) is the tendency to attribute hostile intent onto others in ambiguous situations. Interpersonal conflict in unclear social situations is known to increase the risk of aggression in people with HAB. However, less information is known about certain psychological characteristics and social factors that may influence responses in people with HAB. For instance, trait levels of trust have not been explored previously, but is proposed to be on the opposite side of the spectrum to HAB. This study explored trust, along with other personality variables and environmental/social influences on HAB. Further, this study looked to build on and improve previous measures of HAB. Written scenarios were given to participants that varied in terms of intentionality and social context. The results supported our items measuring the HAB being grouped together to form a total HAB score, but they did not support our categorization of the vignette scenarios. Further, trust and agreeableness decreased the likelihood that someone would report hostile responses. Honesty-Humility traits also decreased the likelihood someone would report hostile responses, whereas having equal power to the individual depicted in the scenario increased the likelihood of hostility. Limitations of the study and future directions for research were explored. As aggression has a large impact on society, it is important to continue research that tries to understand what influences these behavioral responses.

Contribution Statement

Dr. Paul Tremblay assisted at every stage of this project, especially with analyses and conceptualization of the factor analyses and multilevel modeling results. Further, he also contributed with editing throughout the duration of the project.

Acknowledgments

I would like to thank my supervisor Dr. Paul Tremblay for his continued support, professional advice, and help throughout my time working on this project. I would also like to thank my thesis committee, Dr. Rachel Calogero and Dr. Alex Benson, for their guidance and feedback throughout this process. Thank you to the Department of Psychology and Western University for giving me this opportunity to learn and pursue research in an area I'm very passionate about. I would like to thank my parents, Greg and Christina Smith, for being my number one supporters throughout my education journey. Finally, I would like to thank those closest to me who helped either with thesis specific support or personal support at any point throughout the last two years, specifically Michelle and Dave - I could not have done it without you!

Table of Contents

List of Tables

List of Appendices

Introduction

The hostile attribution bias (HAB) is an individual difference variable in the tendency to perceive another person's actions as being either hostile or malicious when the other person's intention is actually benign or ambiguous (Epps & Kendall, 1995; Gagnon & Rochat, 2017; Wilkowski & Robinson, 2010). This erroneous interpretation by the observing individual may then lead them to exhibit reactive aggressive responses (Epps & Kendall, 1995; Wilkowski & Robinson, 2010; Gagnon & Rochat, 2017). Research has shown that aggressive dispositions in children involve difficulty in understanding ambiguous intentions of others (Nasby et al., 1980; Dodge, 1980; Nelson & Perry, 2015). Further, research in adults has assessed their perceptions of hostility in various hypothetical situations which involve varying degrees of provocation and malicious intentions by the other party (e.g., Bowen et al., 2016; Tremblay & Belchevski, 2004).

Interpersonal conflict in ambiguous social situations can instigate aggressive responses in individuals with HAB including children (e.g., meta-analysis by De Castro et al., 2003) or adults (Bailey & Ostrov, 2008; Holtzworth-Munroe & Hutchinson, 1993), but the nature of the psychological properties of these situations need to be explored more extensively. Conflicts that occur in everyday life are inescapable and can have negative impacts on interpersonal interactions and mental health. Trait activation theory proposes that certain traits are activated or otherwise manifest themselves under specific situational properties (e.g., Tett et al., 2021). It would therefore be valuable for researchers to develop specific knowledge about the particular social-contextual properties that might trigger the HAB. Other research has also called for more attention to ways of mapping

social-contextual situations (e.g., Aunger, 2020). In line with these recommendations, the current study aimed to build on and improve previous measures to assess the HAB across a wider set of daily conflict situations.

Hostile Attribution Bias Research in Children

Researchers have investigated the possible mechanisms underlying the HAB and have theorized that a potential impairment may occur in the social information processing steps of interpersonal communication or relationships. Six steps of social information processing have been proposed (McFall & Dodge, 1982). The theory was developed in the context of aggression in children (Dodge & Crick, 1990), but research has expanded this conception to adult populations (Ping-l et al., 2016). The six steps, in order of execution, are: (1) accurately encoding information based on context or environmental cues, (2) accurately interpreting the encoded information, (3) deciding on a goal for the interaction, (4) generating potential responses, (5) evaluating responses and choosing the best response, and (6) enacting on chosen response. An impairment in any of these six steps (with an emphasis on step two) may lead to antisocial or aggressive behavior (McFall & Dodge, 1982).

Research has found that children who behave aggressively do so as a result of interpreting the intentions of others with hostility, especially in ambiguous situations (De Castro et al., 2003; Dodge et al., 2003; Hiemstra et al., 2019). Years later, Hiemstra and colleagues (2019) investigated the strength of interventions for aggressive children, specifically cognitive bias modification training, to reduce hostile interpretations of facial cues. The theory supporting this intervention was also social information processing, where hostile interpretation of others during ambiguous situations stems from

maladaptive processing. Hiemstra et al. (2019) found that preteen boys could be trained to better read facial expressions and thus reduce interpretations of hostility, but this training did not directly reduce the reported levels of anger or aggression. Therefore, training directly impacted the cognitive component, but did not effect the affective or behavioral outcome.

The HAB has also been found to be positively related to traditional and cyber-aggression victimization rather than peer aggression, due to the reactive nature of the behavior (Pornari & Wood, 2009). In other words, the HAB has a positive relationship with the reactive aggression seen in retaliation to bullying (whether cyber or in real life) compared with the premeditated bullying itself. Research has also investigated protective factors for bias in children and has found that higher IQ, superior emotional understanding, and "advanced theory-of-mind" decreased the likelihood that children would have or develop the HAB (Choe et al., 2013). Lower self-control has been linked to the HAB, both in cross-sectionally and longitudinally in children (Nelson & Perry, 2015). The HAB has also been used for predicting adult criminal thinking in juvenile delinquents, where it was found that having a high levels of HAB was related to adult criminal thinking, which in turn predicted delinquency and adult criminal behavior (Walters, 2022).

Hostile Attribution Bias Research in Adults

Research on the HAB in children often used experimental methods in which participants may have been exposed to artificial ambiguous social situations set up by the experimenter. In contrast, research with adults has focused mostly on asking participants how they think they would react in various hypothetical scenarios depicting interpersonal

provocation. A systematic review by Tuente and colleagues (2019) looked at twenty-five studies that investigated the relationship between the HAB and aggression in different adult populations (e.g., general, forensic psychiatric, traumatic brain injury). They found small to medium positive relationships in most studies between the two variables, supporting the notion that aggressive behaviors in adults are linked with the inappropriate interpretation of ambiguous stimuli within social contexts. This study was the first systematic review looking at the relationship between aggression and HAB in adults.

Some research on HAB has also focused on specific interpersonal contexts. For example, in scenarios involving driving motor vehicles, those with higher trait aggression displayed higher HAB when intentions of other drivers were ambiguous (Matthews & Norris, 2006). Interestingly, researchers have found that when a male authority figure is present, the hostile ratio ("a weekly hostility incidence rate which is the equivalent of the average number of reported hostile incidents per 100 persons per week") decreases (Archer, 1961, p. 563). Archer also found a gender difference within the sample, where the hostility ratio was significantly higher for males than females, indicating the influence of context on gender differences.

Kim and colleagues (1998) looked at the impact of power imbalances on seeking revenge following unjust harm. When an instigator causes harm, the person on the receiving end was less likely to seek revenge when they had less power than the instigator. However, the recipient was more likely to seek revenge when they had more power than the instigator. Further, the authors found that when a third party intervened, these trends reversed. In a context involving police, the HAB was not significantly higher in those seeking a law enforcement career, contrary to the hypothesis where it was

expected that those interested in law enforcement would be more likely to have a HAB due to recent and prevalent depictions of police engaging in violence, namely against certain demographics (Ferraccio, 2018). Another study looked at HAB in prisoners (measured using the Hostile Interpretations Questionnaire; Mamuza & Simourd, 1997) and the impact of feeling disrespected by authority figures on aggressive responses, where being reminded of times they were disrespected increased risks of violence through justification of actions (Butler & Maruna, 2009).

A meta-analysis investigating exposure to violent media found an overall positive correlation of $r = .20$ between exposure and hostile appraisals (i.e., HAB; Bushman, 2016). Hostile appraisals and world views can increase the likelihood of aggression (Bushman, 2016). Hostile appraisals or world views or schemas are not exactly the same as the HAB, but there is certainly some overlap between these concepts, and we would imagine a positive correlation between the two. With increases in social media use, researchers have also seen an increase in cyberbullying at various ages, including at the college level. Within this group in particular, the HAB has been seen to be a mediator between covert narcissism and cyberbullying (Fang et al., 2023).

Childhood experiences may have some links to the HAB in adulthood. Studies have found that punitive childhood experiences (i.e., childhood punishments) or maltreatment increase the likelihood of hostility bias in adulthood (Johansson et al., 2021; Milburn et al., 2013; Li et al., 2022). Angry rumination has also been connected to the HAB in a longitudinal study, where the variables were seen to influence each other over time (Wang et al., 2019). Further, negative emotional responding and the HAB have both been found to predict different forms of aggression, including general aggression,

physical aggression, verbal aggression, and relational aggression (Chen et al., 2014). Similarly, negative urgency ("a tendency to react rashly when faced with intense negative emotions…") is another psychological concept that has been associated with the HAB in terms of reactive aggression, where negative urgency was a mediator between the relationship of HAB and reactive aggression (Gagnon & Rochat, 2017, p. 211).

Context Cues

Social contextual cues seem to be important for (mis)interpreting the attributions of another person within an ambiguous social situation (Laue et al., 2018; Schönenberg & Jusyte, 2014; Zajenkowska et al., 2023). Overall, those with HAB seem to be oversensitive to interpersonal conflict situations, leading to an "over-interpretation" of other people's behaviour as potentially hostile (Laue et al., 2018; Zajenkowska et al., 2023). Further, it has been found that perceived intentionality of the 'harm-doer' (i.e., how intentional their actions appeared) mediates the relationship between sensitivity to provocation and reported levels of anger (Zajenkowska et al., 2021).

A study of the offender population found that facial cues are an important component, as it has been found that high HAB individuals are more likely to interpret "ambiguous stimuli containing proportions of an angry expression as hostile", which is associated with a likelihood to act aggressively (Schönenberg & Jusyte, 2014, p. 61). In other words, when cues are ambiguous and not overtly depicting hostility, but also not obviously non-hostile, those with the HAB are more likely to interpret the cues as hostile. In turn, this increases the likeliness of reactive aggression. Further, violent offenders seem to be more susceptible to aggressive-impulsive behavior, which may underlie the

connection between the hostile interpretation of facial cues and aggression (Schönenberg & Jusyte, 2014).

Researchers have also explored provocation sensitivity through eye gaze testing, as well as other methods, to understand the direct focus of people's interpretations. For instance, aggression was more closely associated with eye gaze fixation time on non-hostile cues, suggesting these individuals are fixating on ambiguous cues and then interpreting them as hostile (Lin et al., 2016). In the adolescence population, it was found that aggressive children were more likely to fixate on more hostile cues, relating back to potential errors with the encoding step within social information processing (Laue et al., 2018). Further, it was found that people adjust their attributions "across conflict escalation," indicating that hostility ratings of situations increase when severity of the conflict increase (Godleski & Murray-Close, 2023). This reflects both relational and instrumental conflict trajectories, once more showing the importance of situational/context cues with hostility interpretations.

Overall, contextual cues, including facial cues and expressions, are important for nonverbal means of communication within social interactions. For individuals higher in HAB, ambiguous cues relating to hostility may be over-interpreted, leading to potential instances of aggression, as seen with eye gaze fixation times. Further, other contextual factors are important to consider when understanding the impact of HAB on behavior, such as power imbalances and virtual environments.

Self-Report HAB Measurement

Vignettes depicting hypothetical scenarios have been used in research for decades when trying to better understand people's judgments and attitudes. One of the original

vignette studies by Higgins and colleagues (1975) used depictions of different personality traits of one individual within multiple contexts to understand peoples' responses and judgments based on written descriptions. That approach has been used often in the context of an experimental design that allowed manipulation of various properties of the particular vignette (e.g., the impacts of adding a weapon; D'Costa, 2021).

When studying the HAB, it is important to note the difference between hostility more generally and the bias specifically. To be hostile is 'to be or act in an unfriendly manner,' whereas the HAB is the perception of hostility in situations where the intent of the instigator is ambiguous.

In addition, based on the principle that stable and cross-situationally consistent individual differences or traits, the HAB should be assessed across several situations. Several researchers have developed self-report measures of the HAB by using a sample of short vignettes depicting interpersonal conflicts, realistic of people's encounters in daily life. These questionnaires usually include questions asking the participants about their perceptions of whether the instigators in the conflict were acting with hostile intent (i.e., intended to provoke or harm them). Ideally, the vignettes include some scenarios that are somewhat ambiguous as to the intentionality of the instigator. When scoring these measures, the researchers do not typically indicate what level of a score is associated with HAB in an objective way. Instead, it is assumed that the overall scores aggregated across all scenarios produce a relatively normal distribution, and that those who obtain higher scores are assumed to have stronger perceptions of hostility in the actions of others. In a sense it may be more accurate to refer to these measures as simply individual differences in perception of hostile intent in the actions of others. Therefore,

although we will continue to use the term HAB in this study, we are referring to individual differences in perceptions of hostile intent based on a sample of vignettes that vary in the salience of the hostile intent clues they provide.

Among some of these measures are the Hostile Interpretations Questionnaire (HIQ; Mamuza & Simourd, 1997), the Social Information Processing – Attribution Bias Questionnaire (SIP-ABQ; Coccaro et al., 2009), The Hostile Expectancy Violation Paradigm (HEVP; Gagnon et al., 2016), and the Ambiguous Intentions Hostility Questionnaire (AIHQ; Combs et al., 2007). The HIQ consists of seven vignette scenarios, where each vignette has five specific follow-up questions relating to the specific vignette looking at hostility and components of the social situations. The SIP-ABQ has a subscale for the HAB and has descriptions of scenarios (eight vignettes) that vary in terms of intentionality and ambiguity where participants complete four follow up questions per vignette regarding intentionality. The HEVP contains 320 scenarios where the initial sentence establishes either a hostile or non-hostile intent, the second sentence consists of a target that creates an ambiguous provocation, and the final sentence shows whether intent was present. Scenarios are then coded as being hostile or non-hostile, and the specific coded context either matched or mismatched with the target's intention in the final sentence. Finally, the AIHQ includes 15 vignettes varying in social situations which include negative situations and are rated by perceived intentionality, hostility, blame, and potential reactions.

The measure by Tremblay and Belchevski (2004) consisted of 24 vignettes describing different potentially aggravating situations, where some are classified as being accidental, some as ambiguous, and some as intentional. This measure is rated based on

perceptions of self-anger and aggression. Finally, Lakey and colleagues (2005) created a measure called the Hostile Attribution Bias Scale which consists of 14 short vignettes that vary in terms of ambiguity and levels of potential hostility. Items are rated using Likert scales for intent, anger, and desire to retaliate. These two materials, along with the AIHQ, were used for the current study due to their variety in social situations, along with their easy adaptability to incorporate more characteristics important to the present study, such as target gender.

Overall, no one single measure includes expansive variety in terms of social situations, including modern social situations, such as social media use, along with the categorization of vignette types and characteristics. The current study aimed to explore these short comings of previous measures and develop a more up-to-date conceptualization of the HAB.

Gender and Sex Differences and Personality

One study looked into gender differences in regard to sensitivity to provocation (Zajenkowska et al., 2023). The authors found that men with higher sensitivity to provocation perceived anger in others more often, however the inverse was seen in women where they perceived anger in others less often. Another study found that gender and impulsivity were moderating factors for the HAB and negative emotional responding on aggression (Chen et al., 2014). These studies highlight the importance of researching gender differences in variables that are well established predictors of aggression.

Research has investigated relations between aggression and personality trait inventories based on the Big Five (Barlett & Anderson, 2012). Specifically, agreeableness (negative relationship) and openness to experience (positive relationship)

were related with physical aggression (Barlett & Anderson, 2012). Neuroticism (negative relationship) and extroversion (positive relationship) were also related to physical aggression. However, when looking at violent behavior, the strongest indirect effect found was through aggressive attitudes (Barlett & Anderson, 2012). In a similar manner, aggressive emotions seemed to be a mediator between neuroticism and physical aggression (Barlett & Anderson, 2012).

Ashton and Lee (2014) also explored personality, where they looked at the Big Five along with the dark triad, and their HEXACO six-factor model. Those who were low in honesty-humility were equivalent in the opposite to the core of the dark triad traits (Lee & Ashton, 2005; Lee et al., 2013). Moreover, it was found that Machiavellianism overlapped with both low agreeableness and extraversion, psychopathy overlapped with both low emotionality and conscientiousness, and finally, narcissism overlapped with high extraversion (Lee et al., 2013). The dark triad has been correlated with anger, aggression, and even violence (Lee et al., 2013). Due to the overlap between HEXACO factors and the dark triad, the HEXACO model may then be a beneficial resource for research on hostility.

Other areas of personality have also been investigated in relation to the HAB. This includes studies investigating the HAB as a potential mediator between psychopathy and aggression (Kastner, 2012), a mediator between narcissism and aggression within intimate partner violence (Fields, 2013), and having a strong relationship with vulnerable narcissists (characterized as less “socially confident”) but not grandiose narcissists (Hansen-Brown & Freis, 2021). Looking at the outcome of counterproductive work behavior, gender has been seen to mediate the relationship between stressors and

personality. In particular, men displayed more counterproductive work behaviors together with higher levels of certain personality traits (including trait aggression and HAB), whilst also displaying lower levels of other personality traits (including, agreeableness, neuroticism, and conscientiousness; Spector & Zhou, 2014). However, the same results were not seen for women (Spector & Zhou, 2014).

The Other Side of Hostility? Trust

At the time of the current study, little to no research has directly looked at the relationship between the HAB and generalized or dispositional trust. As aforesaid, the HAB has been extensively connected to aggression (Bowen et al., 2016). In contrast, trust has been found to be negatively correlated with aggression/hostility (Rotenberg et al., 2021; Mattarozzi et al., 2015). Thus, it would be interesting and beneficial to better understand the relationship between the HAB and trust. Conceptually, is seems that trust and perceptions of hostile intent could be at opposite poles of a continuum. Trust can be seen as a general core belief that in uncertain or ambiguous situations, one is willing to give the instigator the benefit of the doubt that their actions were benign or otherwise not intended to cause harm.

Dispositional trust is ‘trait level’ trust that is generalizable and applies to all hypothetical areas of life (Shambare, 2016). This is also referred to as Global trust, which refers to the orientation towards people in general (Rotter, 1967; Wrightsman, 1974). As the current project aims to understand the relationship between HAB and trust, the above definition allows for understanding of the overlap of intentionality between the HAB and trust.

Trust is potentially one of the most important variables influencing interpersonal behaviors (Golembiewski & McConkie, 1975). Trust is an essential component of everyday life, and therefore, warrants interest in research looking at different topics, such as on the HAB. Previous research has found that individuals tend to be initially a bit more distrusting of strangers (Dunning et al., 2019), indicating they may either have a poor trusting attitude, a weak faith in humanity, or both (Shambare, 2016). Further, trust is mainly influenced by interpersonal factors that solidify general views of other people through different interactions throughout childhood and into adulthood (Stamos et al., 2019). Trait levels of trust specifically, can be impacted by a person's first impression as well as previous experiences (Yu et al., 2014).

The Trust Inventory (Couch, 1994; Couch et al., 1996) has three different sections that allow for three different definitions of trust: partner trust (romantic relationships), network trust (family and friends), and generalized trust (people in general). Other earlier measures of 'global trust' include the Interpersonal Trust Scale (Rotter, 1967) and the Trustworthiness subscale from the Philosophies of Human Nature Scale (Wrightsman, 1974). Beyond older self-report measures, there was also a trust game created by Berg et al. (1995), which was then critiqued and amended by Ermisch and colleagues (2009), due to a lack of internal validity. This alteration created the Binary Trust game where those included within the game only have two options to choose from (such as keeping or giving money). More recent measures of trust include the development of two questions called the "SOEP-trust" (Naef & Schupp, 2009), which allows the researchers to understand whether trust in strangers (or global trust) may "load as an independent component" (Naef & Schupp, 2009). A common measure developed around the same

time is the General Social Survey Index (Brehm & Rahn, 1997; Zmerli & Newton, 2008) which measures generalized trust.

When looking at interpersonal hostility and trust beliefs, it was found that the two variables were negatively correlated, indicating that when one expressed more trusting beliefs, less hostile intent was reported (Rotenberg et al., 2021). Although the authors found a linear relationship, a more complex quadric relationship was also seen, where those with both very low and high trusting beliefs demonstrated more hostile intentions during the hypothetical and acquaintance scenarios. Friendly gestures (such as smiling) are seen to be important for building trust, however they may influence perceptions of hostility when seen during interactions with people we are already not fond of (Menon et al., 2014). These studies show the complex relationship that trust and hostility have, and their influence on everyday interactions and cognitions.

On the opposite side of trust, there is distrust. Researchers have also looked at this relationship with hostility and found that at the within person level, interpersonal distrust can predict social aggression at a later timepoint (Yang et al., 2024). Further, the authors found that the HAB mediated this relationship, further supporting a connection between trust and the HAB.

Similar to the HAB, context cues are very important in understanding trust — specifically trust in interpersonal scenarios. Research with first impressions found that ‘happy’ faces (i.e., smiling and open eyes) were deemed the most trustworthy, even when the expressions were subtle, and angry facial expressions were considered the least trustworthy (Thierry et al., 2021). Further, those deemed more trustworthy were given more support in different real-world scenarios (Thierry et al., 2021), leading to the

potential for more positive life outcomes. Previous research has established that power dynamics do influence the development of trust, but it differs depending on the situation, and it is not clear in terms of the influence that power has on trust (du Plessis et al., 2023). For instance, it has been found that when there is an unequal exchange of power (i.e., a power imbalance) between individuals, less trust is seen to be present within an interpersonal exchange (du Plessis et al., 2023). Further, trust tends to be reciprocated more when an individual perceives the initial truster as being high power, and a similar trend is seen with distrust (Mooijman, 2023). Dynamic situations also seem to influence trust, as research has found that when one gains or loses power, changes in trust follow (Brion et al., 2019). This suggests a potential influence for different power dynamics in everyday social situations on levels of trustworthiness, although more research is needed on this topic. Overall, it has been suggested that individuals who possess more power behave more based on their dispositions compared to those who possess less power (Guinote et al., 2012). Because of the influence that power dynamics have on trust, the relationship between dispositional trust and power dynamics may be a bit more complex than other areas of trust.

When looking at the "Big Five," both hostility and trust were found within the different domains (Bronchain, 2023) with trust being a facet of agreeableness (Crew et al., 2018). One study found that interpersonal trust mediated the relationship between agreeableness and ostracism (Hales et al., 2016). When looking at anti-social behaviours at the country level, those who resided in countries where more chronic threats of violence were present were more likely to be agreeable, with cooperative behaviors being limited to those considered familiar or within the same group (i.e., an ingroup bias; White

et al., 2012). With gender differences and trust, women have been found to judge faces classified as “trustworthy” as being being more trustworthy in comparison with men, particularly trustworthy women’s faces (Mattarozzi et al., 2015). However, no gender differences were seen with either neutral or “untrustworthy” faces, suggesting a potential intergroup bias for trust in women. Agreeableness and trait aggression also influenced the trustworthiness ratings, where those low in agreeableness but high in aggression were less trusting in general (Mattarozzi et al., 2015). Therefore, both gender of the participant and the target, along with personality, influence ratings of trustworthiness on first impression judgments of faces.

Overall, trust is strongly influenced by situational factors in the moment, as well as from past interactions. Contextual factors, personality, together with multiple other factors have been seen to influence trust and trust behaviors. The current project aims to further the understanding of the connections trust has with other psychological concepts, such as the HAB, and how it may influence people’s behavioral responses.

The Present Study

In this study, we have developed a new measure of HAB to provide a way to assess not only individual differences in HAB but also the influence of social-contextual differences. While previous individual difference measures of HAB have converged on the need to sample conflict situations across various social contexts (e.g., Tremblay & Belchevski, 2004), the present study uses a more formal approach to treating the sample of conflict situations as a random factor, which provides a stronger case for generalizations beyond that sample of vignettes and also provides a way to quantify the source of variances in HAB attributable to the variation in those social contexts. This new

measure includes domains of work, interactions with public/strangers, family and friends, driving and police, drinking environments including parties, academic contexts, and intimate relationships. Each of these domains includes two to five short vignettes depicting common social interpersonal situations where one person provokes another one. This survey also builds on prior surveys (e.g., Combs et al, 2007; Lakey et al., 2005; Tremblay & Belchevski, 2004) by including questions with rating scales assessing different exemplars of hostile attribution bias such as perceptions or intent of maliciousness, provocation, harm, and rudeness, and self-assessments of one's anger and likelihood to express disapproval in those encounters.

The **first purpose** of this study was to investigate the factor structure of this new measure. Our first aim as part of this purpose was to evaluate the dimensionality of the six questions on each vignette. Specifically, we ***hypothesized*** that the first four items which were meant to measure perceptions of hostility would form one factor, and that the items assessing anger and expressing disapproval would remain as single variables. Support from these analyses would provide justification for creating the HAB composite four-item score for each vignette. The second aim was to use the HAB composite scores (i.e., the first four items) for each of the 24 vignettes and investigate the extent to which the factor structure ***would map onto*** our categorization of the seven social contexts. A multilevel modeling approach was implemented to separate the individual trait-like HAB influence from the social contextual influences.

The **second purpose** of this study was to ***assess the extent*** to which participants' responses to a situation are due to (1) their individual characteristics (manifested as cross-situational stability, more specifically cross-vignette stability, of their responses), (2) the

social psychological effect of particular contextual factors, and (3) the interaction between these two sources (i.e., a person x situation interaction; Furr & Funder, 2021). To accomplish this, we used a multilevel modeling framework (MLM) with two crossed-random factors (i.e., participants and vignettes). We expected that this method would reveal that participants' responses to any question on any vignette would depend largely on individual differences, but also on the particular context or even the particular vignette. We were able to identify the proportion of variation attributable to each source but did not hypothesize which level would contribute more variation.

The **third purpose** of this study was to expand our understanding of how the HAB relates to hypothesized theoretically linked constructs such as dispositional trust and more general personality dimensions such as agreeableness. We ***hypothesized*** that people with high levels of HAB would have low levels of trust and agreeableness, as previous research has found a negative relationship between aggression/hostility and trust (Rotenberg et al., 2021; Mattarozzi et al., 2015). Other common personality traits, such as openness to experience, extraversion, neuroticism, and conscientiousness were explored, as previous research has mainly explored their connection with different types of aggression (Barlett & Anderson, 2012). As part of this research, we also ***explored*** the role of power differential (e.g., instigator of the vignettes has power such as police officers and professors) and gender (e.g., when the instigator is male) in HAB responses. Some research has looked at power imbalances, where it was found that perceived power imbalances decrease the likelihood of facilitating trust (du Plessis et al., 2023); however, power dynamics are a broader topic where it is not necessary for imbalances to be

present. Therefore, power dynamics, stranger status, and gender were analyzed in an exploratory fashion with no formal hypotheses created.

Methods

Participants

The study consisted of a cross-sectional design with online data collection using a survey questionnaire administered on Qualtrics upon two samples of participants.

The first sample was recruited through SONA — an online participant pool management platform — in Dec 2023 to Jan 2024 and consisted of 279 university students (170 women, 104 men, 2 non-binary, 1 trans, 1 queer, and 1 prefer not to say). The mean age of the sample was 18.7 years (SD = 1.7) with a range from 17 to 33. These participants were enrolled in a first-year Psychology course (or other Psychology courses using the Psychology Research Participant Pool) at Western University. The initial sample size consisted of 364 participants, but 85 were removed due to the following reasons: 16 duplicates, 2 with too much missing data, 22 completed in under 5 minutes, 15 completed between 5 to 7.5 minutes that failed one or more of the three attention checks, and 30 that failed 2 or 3 attention checks. Inclusion criteria consisted of students being aged 17 years or older and fluent in English.

The second sample consisted of 201 adults (138 women, 60 men, 2 non-binary, and 1 preferring not to say), recruited from Prolific on Feb 2, 2024. The mean age of the sample was 39.0 years (SD = 18.7) with a range from 18 to 79. The initial sample consisted of 215 participants, but 14 were removed (with 5 completing the survey under five minutes, and 9 who failed 2 or more of the 4 attention checks).

Procedure

Students in the SONA sample viewed the recruitment information briefly describing the survey study of thirty minutes and were directed to the Qualtrics online

survey. The survey first presented the Information Letter. Participants then indicated consent online by selecting a button to indicate yes. Next, they completed the survey consisting of the following measures (also described in the Measures section): a brief demographics questionnaire, the HAB vignettes with questions and rating scales assessing their perceptions, the General Trust Scale (Yamagishi & Yamagishi, 1994), and the Aggression Questionnaire (Buss & Perry 1992). Participants were then presented with debriefing information and received 0.5 credits for their participation. Participants were notified that there were no penalties for leaving unanswered questions or for exiting the study at any time.

The procedure for the Prolific sample was very similar to the SONA sample with three minor differences. First, in terms of compensation, participants in the Prolific sample received a code at the end of the Qualtrics survey to enter into the Prolific application to receive compensation for completing the study. Participants who completed the survey were debriefed, thanked, and compensated with £4.50 (approximately $5.63 USD) for their participation, a rate of £9.00 ($11.25 USD) per hour, through their anonymous Prolific account. Participants were notified that they could leave any questions unanswered without penalty. The second and third minor differences were that some of the demographic questions differed in the two samples as outlined in the next section, and the Prolific participants completed the HEXACO-PI-R 60 item version (Lee and Ashton, 2009) instead of the Buss-Perry Aggression Questionnaire.

Materials

All materials presented below are included in the Appendix with the exception of the measures of trust, trait aggression, and HEXACO-PI-R.

Demographic Questionnaire

The demographic questions included: age, gender, socioeconomic status, ethnicity (open-ended question later coded into categories; see results section) anticipated study program (SONA sample) current occupation (Prolific sample). When coding the open-ended demographic questions, if the participants gave more than one response, the first or the more specific was coded. For example, if the participant indicated they were "White, Canadian" they were included within the white/European category, however if they did not use a comma to separate the two identifiers, such as "White Canadian" they were included in the specific identifying group, which would be Canadian/North American. Further, regarding ethnicity, if the participants indicated they were "Middle Eastern," with no other specifiers to country of origin, they were included within the West Asian category.

Hostile Attribution Bias Vignettes

All participants received the same 24 vignettes in different random orders: "*In this section you will be presented with 24 short written scenarios depicting social interactions that people may encounter in their lives. You are asked to imagine yourself in those situations and respond to the questions asking about your thoughts, feelings, and actions*."

An example of a vignette is the following: "*You are carrying a heavy load of groceries up to a check-out line at the grocery store and just as you are about to enter in line, a woman cuts in front of you. You end up dropping some items on the floor*."

Six questions were presented (in the same order) following each vignette along with 7-point Likert scales as follows:

1. *How malicious or mean do you perceive this individual to have been?* (not malicious at all = 0, extremely malicious = 6);
2. *How likely is it that this person meant to provoke you?* (not at all likely = 0, extremely likely = 6);
3. *How likely is it that this person meant to harm you?* (not at all likely = 0, extremely likely 6);
4. *How rude do you perceive this person to have been?* (not at all rude = 0, extremely rude = 6);
5. *How angry would you feel in this situation?* (not at all angry = 0, extremely angry = 6);
6. *How likely is it that you would express disapproval in words or gestures to this person?* (not at all likely = 0, extremely likely = 6).

Although the above question rating scales were presented to participants on 0-6 rating scales, the responses were coded for analyses on a 1-7 scale.

Approximately half of the vignettes were adapted from previous studies (Combs et al, 2007; Lakey et al., 2005; Tremblay & Belchevski, 2004), and the other half consisted of new vignettes were created in order to incorporate more modern scenarios, including the use of social media technology, and to more specifically target some of the variables of interest such as power dynamics and new domains including the police and workplace. The overall selection also include variation in terms of the gender of the instigator (i.e., male, female, unknown), whether the instigator was known or a stranger, and whether the instigator was in a position of power or not. All vignettes are presented in the results section by content domain. The seven content domains included: work (4

vignettes), public (5 vignettes), family/friends (2 vignettes), driving/police (4 vignettes), drinking (3 vignettes), academics (3 vignettes), and relationships (3 vignettes).

General Trust Scale (Yamagishi & Yamagishi, 1994)

The General Trust Scale (Yamagishi & Yamagishi, 1994) was originally created to access general beliefs about honesty and trustworthiness in others. Participants were asked to indicate how much they agree or disagree with 6 items, such as "Most people are trustworthy" using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). The six items were presented in a random order. The Cronbach alpha reliability for this scale and subsequent scales are based on our combined samples. For General Trust, $\alpha = .79$.

The Aggression Questionnaire (Buss & Perry, 1992)

The Buss-Perry Aggression Questionnaire (Buss & Perry, 1992) includes four subscales that assess dispositional physical aggression (9 items; $\alpha = .85$), verbal aggression (5 items; $\alpha = .82$), anger (7 items; $\alpha = .84$), and hostility (8 items; $\alpha = .84$), and a total score based on the 29 items. The items are usually measured on a 5-point scale, but we measured them on a 7-point Likert scale using the same instructions of how characteristic the statements are of the individual (1 = extremely uncharacteristic of me, 7 = extremely characteristic of me). An example of an item is: "Once in a while, I can't control the urge to strike another person."

HEXACO-PI-R (Lee and Ashton, 2009)

The HEXACO-PI-R: 60 item version was developed by Lee and Ashton (2009). The six dimensions include: Honesty-humility ($\alpha = .77$), emotionality ($\alpha = .78$), extraversion ($\alpha = .84$), agreeableness ($\alpha = .80$), conscientiousness ($\alpha = .78$), and openness to experience ($\alpha = .80$). All 6 domains contain 4 facet level subscales consisting

of 2 to 3 items each. For the purpose of this study, we are using only the six dimension scales. Participants were asked to read 60 statements and rate how much they agree or disagree with the statement using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). An example of an item is: “I would be quite bored by a visit to an art gallery.” Finally, the 60 items were distributed in a random order. Two attention checks were randomly included between the measure items to emphasize any potentially inaccurate data.

Analytic Procedure

Factor Analyses

Factor analyses were conducted for two aims. The first aim was to evaluate the dimensionality of the six questions on each vignette. This evaluation provided us with direction as to whether we could aggregate the first four items as a reliable composite of HAB and leave the remaining two items as single observed variables measuring separate constructs of level of anger and likelihood to expressing disapproval/grievance. These analyses were conducted for each vignette using exploratory factor analysis with principal axis factoring, oblimin rotation, and inspection of the parallel analysis in combination with the scree plot.

The second aim was to use the composite HAB scores (i.e., the first four items) to evaluate whether the 24 vignettes would form factors similar to the 7 different social contexts. We started with a confirmatory factor analysis and tested our hypothesis of seven factors, and we expected that fewer factors may emerge due to possible lack of differentiation between some social contexts. At that point we followed up with exploratory factor analyses to help us determine the more appropriate number of factors.

Multilevel Modeling

Using a multilevel modeling (MLM) framework (with crossed-random factors), we determined (1) the extent to which participants' responses to a situation were due to their individual characteristics (manifested as cross-vignette stability of their responses), (2) the social psychological effect of particular contextual factors, and (3) the interaction between these two sources (i.e., a person x situation interactions) — with a particular focus on the interaction of gender of the participant by gender of the instigator in the vignettes. We expected that this method would also reveal that participants' responses to any question on any vignette would depend largely on individual differences, but also with respect to the particular context or even the particular vignette. We were able to identify the proportion of variation attributable to each source using this model.

Sample size calculations for this type of design and analytic approach study are complex and would require elaborate simulations and several considerations for the diverse effects in the multilevel models. It is not uncommon in these situations to consider the smallest effect size that we are hypothesizing. In our particular study, one hypothesized effect is an interaction between gender of the participant by gender of the instigator in the vignettes. One issue is that we had a relatively small number of vignettes for a random factor. A size larger than 24 was unfeasible considering the required time to complete the questionnaire. Although beyond the scope of this study, it will be possible to use the data to generate some parameter estimates to conduct a post hoc power analysis as a first step for future studies. It is certainly possible that the study is underpowered to detect a small effect interaction. However, the sample size of participants exceeds requirements to detect small regression coefficients, controlling for other predictors. We

recognize that we have more power to detect predictors of participants than predictors of vignettes (such as power dynamics). The alternative would have been to treat the vignettes as a fixed factor but then we would have traded a way to be able to generalize to the population of vignettes for an increase in power. This tradeoff could be investigated in a future study.

We used both jamovi version 2.5.3.0 and SPSS 28-29 for data inspection, creation of composite scores (SPSS) and to prepare the data files from wide to long format (SPSS) for the MLM analyses. These MLM analyses were performed in the jamovi module GAMLj3 (Version 3.3.1; Gallucci, n.d.).

Results

Sample Characteristics

In addition to general sample characteristics described in the Methods section on p. 20, a few additional demographic statistics are presented below.

The socio-economic status (SES) is presented in *Table 2*. In the SONA sample, most identified as being middle class (39.78%), followed by upper middle class (35.13%). In the Prolific sample, the most commonly chosen SES was middle class (32.34%), closely followed by working class (31.84%), and lower middle class (23.88%).

The distribution of the anticipated program of study is presented in *Table 3* and was a unique demographic for the SONA sample, where most students classified themselves as being within a psychology program (26.16%), followed by business (24.37%), and medical science programs (21.86%).

The distribution of Occupation is presented in *Table 4* and was a unique demographic for the Prolific sample, where most participants indicated they were in a Professional Occupation (34.83%), followed by Other (26.87%).

Table 1

Ethnicity Distribution

Ethnicity/Country of Origin	**% of Entire Sample (n = 479)**	**% of SONA (n = 278)**	**% of Prolific (n = 201)**
Indigenous	0.21	0.36	0.00
White/European	44.05	24.01	71.64
Black/African/Caribbean	3.97	3.58	4.48
Southeast Asian	16.91	23.66	7.46
Arab	1.67	2.87	0.00
South Asian	9.81	14.70	2.99
Latin American	2.30	2.87	1.49
West Asian	4.80	7.17	1.49
North American	13.57	19.35	5.47
Other	2.71	1.08	4.98

Table 2

Socio-Economic Distribution

SES	**% of Entire Sample (*n* = 480)**	**% of SONA (n = 270)**	**% of Prolific (n = 199)**
Lower	4.48	4.30	4.48
Working	16.42	4.66	31.84
Lower Middle	14.07	6.45	23.88
Middle	37.53	39.78	32.34
Upper Middle	23.67	35.13	6.47
Upper	3.84	6.45	0.00

Table 3

Anticipated Program of Study (SONA Sample, n = 279)

Program	**% of Total**
Psychology	26.16
Other Social Science	4.66
Business	24.37
Health Science	10.75
General Science	7.17
Medical Science	21.86
Other	2.51
Uncertain	2.51

Table 4

Occupation Distribution (Prolific Sample, n = 201)

Occupation	**% of Total**
Managers	9.45
Professionals	34.83
Technicians and Associate Professionals	11.44
Clerical Support Workers	2.49
Service and Sales Workers	11.44
Skilled Agricultural, Forestry and Fishery	0.50
Craft and Related Trades Workers	1.00
Plant and Machine Operators and Assemblers	1.49
Elementary Occupations	0.50
Other	26.87

Factor Structure of the HAB Composite Across Vignettes

Conceptually, the first four questions which were written as four indicators of HAB are considered a 'cognitive' component, whereas the fifth question assessing a person's perception that they would be angry in the given hypothetical situation is considered an 'affective' concept. The last question assessing whether the participant would express disapproval to the instigator is considered an assertive type of behaviour. The exploratory factor analysis of the six items was performed for each of the 24 vignettes. For example, the results are shown for one of the vignettes (e.g., Vignette 13) including the correlation matrix in *Table 5* followed by the factor loading matrix (pattern matrix) of the rotated solution in *Table 6*. The results in the example of Vignette 13 were fully replicated in two other vignettes (4 and 11) and mostly replicated (i.e., the Rude item loading on two factors) in three vignettes (9, 14, 16). In the remaining vignettes, although the first four items inter-correlated substantially, the fourth item (Rude) correlated slightly higher with item 5 (Angry) and therefore loaded more highly on a second factor than on the first HAB factor.

Taking into account the results with partial support and the conceptual similarity between the four HAB items, we used the four items in the composite HAB score, recognizing that the Anger item correlates quite strongly with the HAB score. Furthermore, we also calculated the internal consistency reliability (Cronbach alpha) of the four-item composite for the 24 vignettes with values ranging from .79 (vignette 12) to .93 (vignette 13). As the range for these reliability estimates were all quite high, it was determined that it would be reliable to aggregate the four questions into a composite HAB score.

Table 5

Correlation Matrix of the 6 items for Vignette 13

	V13_1	V13_2	V13_3	V13_4	V13_5	V13_6
V13_1	—					
V13_2	0.81	—				
V13_3	0.73	0.79	—			
V13_4	0.80	0.79	0.72	—		
V13_5	0.61	0.59	0.51	0.67	—	
V13_6	0.35	0.37	0.37	0.39	0.44	—

Note. All correlations significant at $p < .001$ (N = 477 to 480)

Table 6

Factor Loading Pattern Matrix for Vignette 13

Variable	Factor 1	Factor 2
V13_1 (Malicious)	.84	.07
V13_2 (Provoke)	.96	-.05
V13_3 (Harm)	.89	-.07
V13_4 (Rude)	.74	.19
V13_5 (Angry)	.01	.94
V13_6 (Disapprove)	.18	.34
Correlation F1 with F2	.68	

Factor Analysis of the Vignettes: Investigating the Social Context Domains

We started with a confirmatory factor analysis of the seven domains each treated as a latent variable with 2 to 5 vignette HAB scores as the indicator variables. One problem with the correlated model was that some of the correlations were above 1 (i.e., 1.01 and 1.05), which is an indication of a non-admissible solution. In addition, most of the other correlations among the latent variables were quite high, ranging from .74 to .98, suggesting a lack of distinctiveness between the vignette HAB. Interestingly, the correlations among the 24 vignette HAB scores were much lower ranging from -.05 to .58.

We then ran an exploratory factor analysis, and a parallel analysis suggested that four factors should be retained. The factor loading pattern matrix is presented in *Table 7*. In order to facilitate interpretation of factors, the full set of vignettes is included in *Table 8* along with descriptive statistics (M, SD, Median) for the total sample in the next section. The first factor in *Table 7* includes vignettes that seem to have stronger confrontations, when considering the full set of 24 vignettes. For example, the vignettes with the highest loadings were vignettes, 14, 11, and 16. These refer to a driver yelling at you to get off the road (v14); a family member relentlessly criticizing your life choices at the dinner table (v11); and a guy groping your female friend on the dance floor at a bar (v16). Factor 2 included vignettes that may cause feelings of frustration. The vignettes with the highest loadings for this factor include vignettes 13, 1, and 23. These vignettes refer to getting pulled over by police for making an unsafe lane change (v13); a new coworker not acknowledging you when seen at a café (v1); and a new person you are dating chooses seeing friends over hanging out with you. Factor 3 included vignettes

which may cause feelings of invalidation. The vignettes with the highest loadings include vignettes 2 and 19. These vignettes refer to scenarios that include your boss not agreeing that your idea would work (v2); and a male peer saying your opinion does not make sense (v19). Finally, Factor 4 included vignettes that may cause feelings of exclusion. The vignettes with the highest loadings include vignettes 10 and 24. These vignettes refer to finding out you were not included in a friend gathering through social media (v10); and your partner lying about studying when they actually went to a concert with friends (v24). Overall, the factor analysis suggested factors that look at the way vignettes may cause the participants to feel rather than the overall contextual theme of the vignette.

The uniqueness values in *Table 7* refer to the proportion of variance in each vignette not explained by the combination of the four factors. In general, many of these values are quite high, with 16 vignettes that have values above .50. Also, only 7 vignettes have one of their loadings above +/- .50. Taken together, these results suggest that the vignettes are fairly unique and only partially explained by the four identified factors.

Table 7

Factor Loadings for Vignette Hostility Composite Scores

Vignettes	Factor				Uniqueness
	1	2	3	4	
1		0.457		0.218	0.558
2			0.693		0.469
3		0.287	0.243	0.222	0.534
4		0.332	0.269		0.564
5	0.436				0.664
6	0.432				0.595
7	0.409	0.262			0.548
8	0.351	0.434			0.450
9	0.383				0.794
10				0.690	0.471
11	0.556			0.200	0.610
12	0.478	0.383			0.415
13		0.474	0.350		0.607
14	0.774				0.449
15		0.431			0.768
16	0.521				0.728
17		0.373			0.686
18				0.342	0.747
19			0.647		0.449
20		0.400	0.353		0.497
21	0.344			0.226	0.532
22		0.380		0.259	0.700
23	-0.211	0.453	0.217	0.288	0.476
24				0.571	0.586

Note. 'Principal axis factoring' extraction method was used in combination with an 'oblimin' rotation. Only loadings of +/- .20 or higher are included.

HAB, Trust, Trait Aggression, and HEXACO Descriptive Statistics

Descriptive Statistics for the Vignette HAB Scores

The 24 vignette descriptions along with the mean, median and SD of the HAB composite scores for the combined samples are presented in *Table 8* separated by social contextual domain. When looking at the composite hostility for the vignettes, the highest average hostility score was for vignettes 16 (M = 5.90, SD = 1.07), followed by vignette 14 (M = 5.61, SD = 1.12). The vignettes with the lowest hostility scores were vignette 23 (M = 2.45, SD = 1.29) and vignette 15 (M = 2.50, SD = 1.35). Although not presented in the table, the skewness values were all under 1.0 with the highest value of -0.902 for vignette 16. The kurtosis values were also low, with the highest value of -0.659 for vignette 12. These values suggest no substantial departure from normality.

Table 8

Descriptive Statistics for Vignette HAB Scores

Vignette	*M*	*Mdn*	*SD*
Employment			
1 You cross path with one of your new coworkers as you walk into a cafe. You are convinced that she recognized you and you say hello to her, but she passes by you without acknowledging you.	3.13	3.00	1.28
2 You are in a Zoom meeting with your work team. Your supervisor has asked the team a question, and you provide what you think is a reasonable answer. Your supervisor responds: “I don’t think that will work; let’s see what the others have to say.”	2.66	2.50	1.29
3 You are having a discussion with a group of colleagues at a new job. Just as you start to speak, one of your colleagues talks over you.	3.36	3.25	1.27
4 You hear about a new project at work. You know that you are the most qualified for the project, but your boss assigns it to the newest employee who has less experience than you.	3.27	3.20	1.42
Interactions with Public/Strangers			
5 You are on a bus sitting in an aisle seat. A man gets on the bus and steps on your foot as he walks past you, gives you a dirty look, and does not apologize.	4.92	5.00	1.27
6 You are boarding a crowded bus. Just as you try to sit down next to a woman, she places her bag in what would have been your seat.	4.19	4.00	1.29
7 You arrive at a leasing office to get more information about the apartments in the complex. The leasing agent notices you, but he continues talking on the phone for a while. Once he gets off the phone, he greets another customer that just walked in instead of you.	4.20	4.25	1.22
8 You are carrying a heavy load of groceries up to a check-out line at the grocery store and just as you are about to enter in line, a woman cuts in front of you. You end up dropping some items on the floor.	4.03	4.00	1.20
9 You are walking downtown and come across a crowd of people demonstrating for a cause that you are opposed to. Three demonstrators walk up to you yelling in your face: “You are with us or against us!” One of them puts their hand on your shoulder to stop you and get your attention.	4.38	4.50	1.43
Family and Friends			
10 You are scrolling on Instagram one day, and a picture of your close friends at a party from the night before pops up, but no one had mentioned the gathering to you.	3.75	3.75	1.43
11 You are having dinner with your family. A family member starts criticizing your life choices and making fun of your achievements. You try to ignore them, but they keep going on and on.	5.60	5.75	1.15

Vignette	*M*	*Mdn*	*SD*
Driving and Police			
12 You have been looking for a parking spot for a while at the mall and finally see one up ahead. You put your signal on and proceed towards the spot, but another driver, who clearly saw you, rushes into the parking spot.	4.47	4.50	1.23
13 You are driving and made a right-hand turn. Right after you turn, you change lanes. A minute later you get pulled over, and the police officer tells you that you made an unsafe lane change as you did not use your signal. You are confident that you had your signal on.	3.01	2.75	1.54
14 You are riding your bike down the street. A car tries to go around you and ends up cutting you off, nearly clipping your front tire and then brakes abruptly at a red traffic light. You are now both stopped at the traffic light, and the driver yells at you to get off the road.	5.61	5.75	1.12
15 You are at a street party following a university football game and the street is packed with people. An ambulance needs to get through to help an injured person. The police are there trying to make space, and in the process one of the officers pulls you aggressively by the arm in order to clear the path.	2.50	2.25	1.35
Bar – Drinking Environments Including Parties			
16 You're dancing at a bar with a group of your friends, and a guy bumps into your female friend from behind and gropes her. Your friend looks a bit shaken and distressed.	5.90	6.13	1.07
17 You have been waiting in line with your friends for over half an hour to get into a bar. You are to be the next ones to get in but two girls who appear to be very intoxicated push their way in front of you.	3.44	3.25	1.23
18 You and a couple of friends are at a university student house party, and you are introduced to several people you don't know. One guy starts talking to you and looks at one of your male friends and back at you and whispers in your ear: "Your friend looks gay!"	4.20	4.25	1.39
Academics			
19 You are having a discussion with a couple of students. You disagree with the other students and express your opinion. One of the male students tells you that you are not making any sense.	3.56	3.50	1.40
20 You have just received your mark on a research assignment, and you think the mark is too low and not reflective of your effort and the quality of your work. You decide to reach out to the professor asking her if she would reconsider the mark. However, she is unwilling to consider this request and replies that the mark is accurate.	2.95	2.75	1.33
21 You are working on a group project with three other classmates. You have done most of the work and assigned tasks to the others, but they have not completed them by the deadline. You confront them about it, but they blame you for being too bossy and not giving them enough time.	4.24	4.25	1.27

Vignette	*M*	*Mdn*	*SD*
Relationships	3.39	3.25	1.52
22 You are at a party with your friends. One of your friends who you are romantically interested in is avoiding you that evening and flirting with someone who you dislike.			
23 You have been on a few dates with a person who you are very attracted to, and you ask them if they would like to go out again next weekend, but they reply that they need some space and want to spend time with their group of friends.	2.45	2.13	1.29
24 You discover that your romantic partner has lied to you, telling you that they were studying for an exam, but you find out that they were going out to a concert with friends.	4.23	4.25	1.35

Note. Responses were using a Likert scale ranging from 1-7. $N = 480$.

Gender and Sample Differences in Vignette HAB Scores

Two by two factorial ANOVA designs were used to investigate gender and sample differences in HAB scores across the 26 vignettes. *Table 9* presents gender by sample means and SDs and indicates which effects were significant. It should be noted that ANOVA model uses Type III (Unique) sums of squares, which means that each effect is adjusted for the other effect (given that the sample sizes are unequal). In order keep Type I error rate at a reasonable level, we are interpreting only results significant at $p < .01$ cautiously and results at $p < .001$ with more confidence.

Table 9 reveals that the SONA sample had significantly higher scores (at $p < .001$) than did the Prolific sample on Vignettes 3, 10, 13, 18, 19, 20, 21, 23. One clear trend is that the SONA sample of university students had higher scores on the three vignettes depicting academic scenarios (i.e., 19, 20, 21). The lower scores in the Prolific sample could be due to not having attended university for some time (or at all) or having attended in the past and using those more distant experiences to form their present-day

perceptions. The Prolific sample had significantly higher scores (at $p < .01$) on Vignettes 6 and 9, both of which involve scenarios with strangers in public.

In terms of gender differences, women had higher scores than men on vignettes 10 and 24 ($p < .001$) and vignettes 19 and 23 at ($p < .01$). A common tread for vignettes 10, 24, and 23 is that they depict being left out or rejected. Men had higher scores on vignette 9 ($p < .01$), depicting a confrontation with a downtown demonstration.

Only one gender by sample interaction for vignette 22 was significant at $p < .05$ and is therefore not interpreted.

Table 9

Vignette HAB Split by Sample and Gender

Vignette	*SONA Sample*		*Prolific Sample*	
	Men (n=104) M (SD)	***Women (n=170) M (SD)***	***Men (n=60) M (SD)***	***Women (n = 138) M (SD)***
1	3.12 (1.30)	3.29 (1.24)	2.92 (1.20)	3.12 (1.33)
2	2.72 (1.28)	2.75 (1.32)	2.50 (1.21)	2.63 (1.32)
3*$^{+++}$	3.38 (1.24)	3.63 (1.28)	2.88 (1.05)	3.26 (1.31)
4	3.30 (1.36)	3.44 (1.31)	3.02 (1.57)	3.20 (1.51)
5+	5.07 (1.25)	5.04 (1.24)	4.85 (1.33)	4.73 (1.30)
6*$^{++}$	3.88 (1.29)	4.16 (1.24)	4.25 (1.09)	4.50 (1.36)
7	4.06 (1.14)	4.25 (1.16)	4.21 (1.42)	4.28 (1.24)
8	4.04 (1.19)	4.16 (1.15)	3.83 (1.19)	4.01 (1.28)
9**$^{++}$	4.48 (1.35)	4.07 (1.42)	4.81 (1.43)	4.49 (1.46)
10***$^{+++}$	3.47 (1.40)	4.32 (1.29)	2.86 (1.32)	3.66 (1.40)
11^{+}	5.36 (1.35)	5.58 (1.12)	5.64 (1.06)	5.79 (1.05)
12	4.42 (1.38)	4.59 (1.15)	4.36 (1.15)	4.48 (1.24)
13^{+++}	3.40 (1.57)	3.13 (1.50)	2.77 (1.59)	2.69 (1.46)
14	5.48 (1.27)	5.58 (1.13)	5.74 (1.02)	5.69 (1.05)
15	2.32 (1.23)	2.59 (1.39)	2.63 (1.39)	2.49 (1.40)
16+	5.80 (1.15)	5.80 (1.07)	6.14 (0.91)	5.98 (1.07)
17*	3.33 (1.26)	3.52 (1.22)	3.26 (1.09)	3.56 (1.27)
18^{+++}	4.29 (1.46)	4.46 (1.30)	3.77 (1.37)	4.02 (1.41)
19**$^{+++}$	3.48 (1.44)	3.97 (1.34)	3.05 (1.23)	3.37 (1.41)
20^{+++}	3.25 (1.32)	3.12 (1.31)	2.72 (1.33)	2.64 (1.31)
21^{+++}	4.51 (1.31)	4.53 (1.14)	3.63 (1.14)	4.00 (1.30)
22$^{\#}$	3.22 (1.50)	3.63 (1.50)	3.45 (1.42)	3.24 (1.56)
23**$^{+++}$	2.42 (1.33)	2.76 (1.35)	2.00 (1.05)	2.34 (1.22)
24***$^{+}$	4.12 (1.33)	4.49 (1.26)	3.64 (1.28)	4.34 (1.41)
Composite *$^{+}$	3.87 (0.73)	4.04 (0.73)	3.70 (0.74)	3.86 (0.77)

Note. Gender *$p < .05$, **$p < .01$, ***$p < .001$; Sample $^{+}p < .05$, $^{++}p < .01$, $^{+++}p < .001$; Gender by Sample interaction $^{\#}p < .05$

Descriptive Statistics and Intercorrelations of Variables in MLM Models

Prior to presenting the multilevel models in the next section, this section includes descriptive statistics for the main variables in *Table 10* and their intercorrelations in *Table 11*. It is important to note that in this section, the statistics are all aggregated between individuals and — in a sense — disregard the variation in the vignettes. These measures include the composite HAB score aggregated across vignettes as well as the individual items. These are all capturing individual differences in dispositions to report high or low levels of HAB, anger, and expression of disapproval regardless of the particular vignette. Other variables include the Trust scale, the Aggression Questionnaire subscales (for the SONA sample only), and the six HEXACO personality trait scales (for the Prolific sample only). For all the variables in *Table 10*, inspection of the distributions revealed no substantial departure from normality with the highest skewness value of 0.849 for Physical Aggression, and the highest kurtosis value of -0.825 for trust.

The correlations in *Table 11* indicate that all follow-up items are highly correlated with each other, including the HAB composite (0.429 to 0.933), as was seen with the factor analysis. Trust was also significantly negatively correlated with all follow-up questions and HAB composite, except for expressing disapproval (-0.134 to -0.216), indicating that when trust would increase, responses in the first five follow up questions (maliciousness, intentionality, provocation, rudeness, and anger) and HAB responses would decrease. Looking at the HEXACO, honesty-humility correlated negatively with all follow-up questions and the HAB composite (-0.202 to -0.274); therefore, when honesty-humility scores increased, the follow-up and HAB responses decreased. Emotionality was only significantly correlated with rudeness (0.222). Extraversion,

openness to experience, and conscientiousness had no significant correlations, however agreeableness was correlated negatively with all follow-up items and HAB composite (-0.232 to -0.351) and positively correlated with trust (0.245). Physical aggression was significantly correlated with all outcome variables (i.e., follow-up question and HAB composite), including trust (trust = -0.183; 0.198 to 0.401), indicating that when trust increased, physical aggression scores decreased, but when all HAB follow-up responses and mean scores increased, so did physical aggression scores. Verbal aggression was significantly correlated with rudeness, anger, and expressing disapproval (0.235 to 0.362). Buss-Perry Anger was significant with all variables except for perceived intent to cause harm (-0.251 with trust; 0.181 to 0.291). Finally, Buss-Perry Hostility was significantly correlated with all variables: negative with trust (-0.329) and positive with all other outcome variables (0.212 to 0.369).

Table 10

Descriptive Statistics of Variables Used in MLM Models

Measures	***M (SD)***	***Mdn***
Vignette Items		
HAB Composite	3.89 (0.75)	3.86
Maliciousness	4.05 (0.89)	4.04
Provocation	3.72 (0.86)	3.68
Harm	3.00 (0.92)	2.94
Rude	4.81 (0.71)	4.80
Angry	4.58 (0.86)	4.58
Disapproval	3.87 (0.90)	3.88
General Trust Scale	2.93 (0.96)	3.00
HEXACO		
Honesty-Humility	3.48 (0.67)	3.54
Emotionality	3.40 (0.64)	3.42
Extraversion	2.92 (0.73)	2.92
Agreeableness	3.19 (0.66)	3.25
Conscientiousness	3.65 (0.60)	3.67
Openness to Experience	3.37 (0.75)	3.50
Aggression Questionnaire		
Physical Aggression	2.76 (1.18)	2.56
Verbal Aggression	3.61 (1.29)	3.60
Hostility	3.62 (1.24)	3.63
Anger	2.98 (1.21)	2.71

Note. Sample size for combined samples = 480 (Vignettes and Trust scale); sample size for Prolific sample = 201 (HEXACO) and SONA sample = 279 (Aggression Questionnaire). Vignette follow-up questions were rated on a scale from 1 (none) to 7 (extremely). The General Trust Scale items were rated on a scale from 1 (strongly disagree) to 5 (strongly agree). The HEXACO items were rated on a scale from 1 (strongly disagree) to 5 (strongly agree). The Aggression Questionnaire's items were rated on a scale from 1 (extremely uncharacteristic) to 7 (extremely characteristic).

Table 11

Correlation Matrix

	HAB Score	Malic	Provoke	Harm	Rude	Angry	Disap	Trust
HAB Score	--							
Malic	0.932	--						
Provoke	0.933	0.830	--					
Harm	0.864	0.709	0.795	--				
Rude	0.804	0.760	0.658	0.500	--			
Angry	0.756	0.730	0.663	0.527	0.794	--		
Disap	0.556	0.492	0.537	0.429	0.525	0.634	--	
Trust	-0.203	-0.216	-0.216	-0.134	-0.153	-0.179	-0.023	--
HH	-0.257	-0.221	-0.236	-0.259	-0.202	-0.259	-0.274	0.177
EM	0.137	0.135	0.083	0.078	0.222	0.199	-0.075	-0.051
EX	-0.071	-0.059	-0.017	-0.096	-0.085	-0.099	0.131	0.157
AG	-0.292	-0.254	-0.267	-0.292	-0.232	-0.351	-0.305	0.245
CO	-0.051	-0.030	-0.054	-0.104	0.022	-0.015	-0.029	-0.061
OP	-0.026	-0.031	-0.064	0.004	-0.000	-0.106	-0.048	-0.010
BP_PAgg	0.238	0.200	0.225	0.212	0.198	0.174	0.401	-0.183
BP_VAgg	0.166	0.138	0.134	0.095	0.235	0.246	0.362	-0.132
BP_Anger	0.209	0.181	0.207	0.125	0.235	0.291	0.255	-0.251
BP_Hos	0.328	0.291	0.317	0.212	0.354	0.369	0.277	-0.329

Note. HH: Honesty-Humility, EM: Emotionality, Ex: Extraversion, AG: Agreeableness, CO: Conscientiousness, OP: Openness, BP_PAgg: Physical Aggression, BP_Vagg: Verbal Aggression, BP_Anger: Anger, BP_Hos: Hostility. Sample size for combined samples = 480 (Vignettes and Trust scale); sample size for Prolific sample = 201 (HEXACO) and SONA sample = 279 (Aggression Questionnaire). All correlations in the rows starting at the HAB score and ending at Trust are significant at $p < .001$, except correlation between Trust and Express Disapproval *ns*. Correlations in the HEXACO rows (HH to OP) equal or greater than +/- .20 significant at $p < .005$. Correlations in Aggression Questionnaire rows above +/- .17 significant at $p < .005$

Multilevel Models

A number of models were run in three hierarchical steps starting with an Intercept only model, followed by a model with all predictors but no interaction, and a third model with an interaction between gender of participant by gender of instigator in the vignettes (male, female, or not specified). These models were all run with the Restricted Maximum Likelihood (REML) estimator to obtain the best estimate of the random effect variance estimates. No random slopes were included in any of the models to limit complexity. The SONA and Prolific samples were combined into one total sample for models on which data was provided by both samples. In those models we included Sample as a binary predictor. Categorical predictors were dummy coded and left uncentered, while continuous predictors were grand mean centered.

Six tables were created to outline the results for the multilevel models. Three tables include models based on the total sample, where each single table reports the results for the three outcome variables: hostility, anger, and disapproval. The final three tables report results of models for the three outcome variables, run separately on the SONA sample that completed the Buss-Perry Aggression Questionnaire and the Prolific sample that completed the HEXACO personality questionnaire. Although we identified effects that were significant at a nominal alpha of .05, we interpreted only those that were significant at .01 or smaller to maintain the Type I error rate at a smaller level.

HAB Composite with Full Sample

The results of the models with the HAB composite score as the outcome variable are presented in *Table 12*. The intercept only model indicates that the intraclass correlation (ICC) for participants was 0.21. This value indicates that 21% of the variance

in HAB scores is attributable to participant differences. The ICC for vignettes was 0.34, indicating that 34% of the variation in HAB composite responses can be accounted for by the vignettes themselves. Taken together these results indicate that there is more heterogeneity of responses across the vignettes than across participants. The remaining sources of variance occur within the participant by vignette level. The intercept had a value of $\beta = 3.89$ representing the expected or predicted value of the HAB composite responses when all the predictors are 0. Although the intercept is often statistically significant, we disregarded this test because it simply indicates that the value is different from 0. In the intercept only model, there are no fixed effects predictors and therefore the 3.89 value is the grand mean of HAB scores across all participants' responses to each vignette.

In the second model in *Table 12*, there are no predictors at the within participant by vignette level. We will see examples of these predictors in later models (i.e., HAB composite for one vignette predicting the anger response for that vignette). The Between participant predictor that was significant at $p < .01$ was trust ($\beta = -0.15$, $p < .001$), in line with the hypothesis for the third objective. The regression coefficient indicates that for a one unit increase in trust, the HAB composite score is expected or predicted to decrease by 0.15 units, taking into account the other predictors in the model. The between vignette predictor that was significant was power dynamic (perpetrator power), $\beta = 1.17$, $p < .01$. This was a binary predictor with the instigator having a status Equal = 1 and Superior = 0 to the participant, therefore indicating that HAB composite scores had means that were 1.17 point higher when the instigator was equal than superior, adjusting for other predictors in the model. The Conditional R^2 value of 0.554 is the total amount of variance

in the outcome explained by the combination of the random factors (i.e., participants and vignettes) and fixed predictors, whereas the Marginal R^2 of .160 represents the proportion of variance explained by the fixed predictors only (not unlike an R^2 in regular multiple linear regression).

The third model in *Table 12* was a test of an interaction between gender of the participants by gender of the instigator. Given the small number of non cis participants, we only included cis women and cis men in these analyses. However, the vignette gender categories required two contrasts (male vs. female) and (non-specified vs. female). The two interaction terms in the model were not statistically significant indicating the male and female participants did not differ significantly in how they viewed male vs. female instigators or not specified vs. female instigators.

Table 12

Outcome Variable: HAB Composite

Model Parameters	**Intercept only**	**Without interaction**	**With interaction**
Intercept	3.893 (.198)	2.702(.496)	2.709(.497)
Within participant by vignette predictors			
Participant Gender by Vignette Gender:			
[Female - Male] x [Male – Female]			-0.021(.067)
[Female - Male] x [Not sp. - Female]			0.026(.055)
Between participant predictors			
Sample (1 = SONA; 0 = Prolific)		0.105(.070)	0.105(.070)
Gender *[Female - Male]*		0.153(.071)*	0.142(.083)
Trust		-0.151(.036)***	-0.151(.036)***
Between vignette predictors			
Perpetrator gender *[Male - Female]*		0.649(.498)	0.663(.500)
Perpetrator gender *[Not specified - Female]*		0.468(.411)	0.451(.412)
Perpetrator power [*Equal - Superior*]		1.168(.403)**	1.168(.403)**
Perpetrator known *[yes - no]*		-0.580(.337)	-.0580(.337)
Variance components			
Residual	1.210	1.213	1.213
Participants intercepts (random effect)	0.551 icc = .206	0.483	0.483
Vignette intercepts (random effect)	0.917 icc = .342	0.590	0.590
Conditional R^2	0.541	0.554	0.554
Marginal R^2	-	0.160	0.160
Model summary			
Log likelihood	-18088.077	-17785.772	-17789.370
Number of estimated parameters	4	11	13
Number of cases (observations/partic/vig)	11516/480/24	11324/472/24	11324/472/24

Note. *$p < .05$, **$p < .01$, ***$p < .001$. Regression coefficients with standard errors in parentheses are unstandardized. REML refers to Restricted Maximum Likelihood Estimation.

Anger with Full Sample

The results of the models with the vignette Anger response score as the outcome variable are presented in *Table 13*. We considered treating this outcome variable as an ordinal variable with a categorical approach to MLM; however, we opted to treat it as continuous given that it had a fairly normal distribution, and that in another similar project, we did run models with both approaches and found little differences in the results. The intercept only model indicates that the ICC for participants was 0.22. This value indicates that 22% of the variance in Anger scores is attributable to participant differences. The ICC for the vignettes was 0.27, indicating that 27% of the variation in Anger scores can be accounted for by the vignettes themselves. Together, these results indicate that there is slightly more diversification of responses across the vignettes than across participants. The remaining sources of variance occur within participant by vignette level. The intercept had a value of $\beta = 4.58$, representing the predicted value of the Anger responses when all predictors are 0, and it is also the grand mean Anger score across all participants' responses to each vignette.

In the second model in *Table 13*, the one predictor at the within participant by vignette level, the HAB composite has a significant coefficient, $\beta = 0.76$, $p < .001$. This indicates that with one unit increase in the Hostile composite score, there is an expected increase in Anger scores by 0.76 units within participants, taking into account the other predictors in the model. This represents a very strong effect but it should be recognized that it is due to the strong association between participants responses to HAB composite items and the Anger item on the vignettes. No between participant predictors or between vignette predictors were significant. The Conditional R^2 value of 0.689 is the total

amount of variance in the outcome explained by the combination of the random factors and fixed predictors. The Marginal R^2 of 0.512 represents the proportion of variance explained by the fixed predictors only.

The third model in *Table 13* was the interaction test between gender of the participants by gender of the instigator. The two interaction terms were not statistically significant, indicating that male and female participants did not differ significantly in how they viewed male vs. female instigators or not specified vs. female instigators with Anger as the outcome variable.

Table 13

Outcome Variable: Vignette Anger Response

Model Parameters	**Intercept only**	**Without interaction**	**With interaction**
Intercept	4.580 (.190)	4.660(.302)	4.650(.303)
Within participant by vignette predictors			
Hostility composite		0.762(.008)***	0.762(.008)***
Participant Gender by Vignette Gender:			
[Female - Male] x [Male – Female]			-0.026(.057)
[Female - Male] x [Not sp. - Female]			-0.018(.047)
Between participant predictors			
Sample (1 = SONA; 0 = Prolific)		-0.011(.055)	-0.011(.055)
Gender *[Female - Male]*		0.060(.055)	0.076(.066)
Trust		-0.040(.028)	-0.040(.028)
Between vignette predictors			
Perpetrator gender *[Male - Female]*		-0.151(.301)	-0.134(.304)
Perpetrator gender *[Not specified - Female]*		0.032(.248)	0.044(.250)
Perpetrator power [*Equal - Superior*]		-0.135(.244)	-0.135(.244)
Perpetrator known *[yes – no]*		0.013(.204)	0.013(.204)
Variance components			
Residual	1.582	0.880	0.880
Participants intercepts (random effect)	0.666 icc = .217	0.284	0.284
Vignette intercepts (random effect)	0.828 icc = .269	0.215	0.215
Conditional R^2	0.486	0.689	0.689
Marginal R^2	-	0.512	0.512
Model summary			
Log likelihood	-19599.204	-15900.471	-15904.658
Number of estimated parameters	4	12	14
Number of cases (observations/partic/vig)	11500/480/24	11308/472/24	11308/472/24

Note. *$p < .05$, **$p < .01$, ***$p < .001$. Regression coefficients with standard errors in parentheses are unstandardized. REML refers to Restricted Maximum Likelihood Estimation.

Expressing Disapproval with Full Sample

The results for the models of the Expressing Disapproval score as the outcome variable are presented in *Table 14*. This outcome variable was another one that we decided to treat as continuous given its fairly normal distribution. The intercept only model indicates that the ICC for participants was 0.17, implying that 17% of the variance in Expressing Disapproval scores is attributable to participant differences. The ICC for vignettes was 0.32, indicating that 32% of the variation in Expressing Disapproval scores can be accounted for by the vignettes themselves. Taken together, these results suggest that there is almost twice as much heterogeneity of responses across vignettes than across participants. This also suggests that the situational context has a larger impact than individual differences for this outcome. The remaining sources of variance occur within participant by vignette level. The intercept of $\beta = 3.87$ is the grand mean of the Expressing Disapproval scores across all participants' responses to each vignette.

In the second model in *Table 14*, the two predictors at the within participant by vignette level were significant: HAB composite ($\beta = 0.32$, $p < .001$) and Anger ($\beta = 0.47$, $p < .001$). Again, it is not surprising that these two predictors were significant; however, given their overlap, we did not expect that both would be significant in the regression model. Two Between participants predictors were significant at the $p < .01$: sample ($\beta = -0.27$, $p < .001$) and gender ($\beta = -0.42$, $p < .001$). Sample was a binary predictor with a SONA = 1 and Prolific = 0, therefore indicating that Expressing Disapproval scores had means that were 0.27 point lower for the SONA sample than for the Prolific sample, adjusting for other predictors in the model. Gender was also a binary predictor with female = 1 and male = 0, therefore indicating that Expressing Disapproval scores had

means that were 0.42 point lower for women than men, adjusting for other predictors in the model. The Conditional R^2 value of 0.639 is the total amount of variance in the outcome explains by the combination of random factors and fixed predictors. The Marginal R^2 of 0.469 represents the proportion of variance explained by the fixed predictors only.

The third model in *Table 14* was the interaction test between gender of the participants and gender of the instigator. The two interaction terms in the model (male vs female and non-specified vs female) were significant, but only at $p < .05$: male vs. female; $\beta = 0.167$, non-specified vs female; $\beta = 0.137$. An interpretation of this interaction is provided below recognizing that is only significant at $p < .05$. The first value suggest that women are more influenced than men in their responses to Expressing Disapproval by the contrast of whether the instigator was a man vs. a woman, with higher scores when the instigator is a man. The second value has a similar interpretation with the only difference that the instigator man is replaced by unknown gender.

Table 14

Outcome Variable: Vignette Expressing Disapproval response

Model Parameters	**Intercept only**	**Without interaction**	**With interaction**
Intercept	3.870 (0.245)	3.735(0.373)	3.810(.375)
Within participant by vignette			
Hostile composite		0.320(.014)***	0.320(.014)***
Anger		0.472(.012)***	0.473(.012)***
Participant Gender by Vignette Gender:			
[Female - Male] x [Male - Female]			0.167(.074)*
[Female - Male] x [Not sp. - Female]			0.137(.061)*
Between participant predictors			
Sample (1 = SONA; 0 = Prolific)		-0.271(.063)***	-0.272(.063)***
Gender *[Female - Male]*		-0.416 (.063)***	-0.531(.078)***
Trust		0.070(.032)*	0.070(.032)*
Between vignette predictors			
Perpetrator gender *[Male - Female]*		0.377(.373)	0.268(.376)
Perpetrator gender *[Not specified -*		0.298(.308)	0.209(.310)
Perpetrator power *[Equal - Superior]*		0.539(.302)	0.539(.302)
Perpetrator known *[yes – no]*		-0.235(.253)	-0.235(.253)
Variance components			
Residual	2.216	1.461	1.461
Participants intercepts (random effect)	0.720 icc = .166	0.360	0.360
Vignette intercepts (random effect)	1.403 icc = .323	0.329	0.329
Conditional R^2	0.489	0.639	0.639
Marginal R^2	-	0.469	0.469
Model summary			
Log likelihood	-21512.798	-18712.205	-18712.842
Number of estimated parameters	4	13	15
Number of cases (observations/partic/vig)	11517/480/24	11308/472/24	11308/472/24

Note. *$p < .05$, **$p < .01$, ***$p < .001$. Regression coefficients with standard errors in parentheses are unstandardized. REML refers to Restricted Maximum Likelihood Estimation.

Hostile HAB Composite Split by Sample

Models were run for each outcome variable split by the sample. All predictors were included within one model (i.e., we did not start with intercept only model and did not include an interaction test model). By splitting the samples, we were able to incorporate other predictors, such as the Buss-Perry (BP) Aggression Questionnaire (SONA sample) and HEXACO (Prolific sample) variables. The BP predictors include BP Anger, BP Hostility, BP verbal aggression, and BP physical aggression. The HEXACO predictors include Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience. These models are presented in Table 15.

With this model of the HAB as outcome variable, there are no predictors at the within participant by vignette level. For the SONA sample, the between participant predictors that were significant at $p < .01$ were gender ($\beta = 0.257$, $p < .01$) and BP Hostility ($\beta = 0.161$, $p < .001$). Gender was a binary predictor with female = 1 and male = 0, therefore indicating that HAB composite scores had means that were 0.257 units higher for female SONA participants than male SONA participants, taking into account other predictors in the model. Further, with every one unit increase in BP Hostility scores, there is a predicted 0.161 unit increase in HAB composite scores. The between vignette predictor that was significant at $p < .01$ was perpetrator power ($\beta = 1.109$, $p < .01$). Once again, this is a binary predictor (Equal = 1 and Superior = 0), therefore indicating that HAB composite scores had means that were 1.109 point higher when the perpetrator was equal in power, adjusting for other predictors in the model for the SONA sample. The Conditional R^2 value was 0.538 and is the variance in outcome explained by

both the random factors and fixed predictors. The Marginal R^2 of 0.172 is the proportion of variance explained by the fixed predictors only.

For the Prolific sample, the between participant predictors were significant at $p <$ .01 were Honesty-Humility (β = -0.220, $p < .01$) and agreeableness (β = -0.235, $p < .01$). The regression coefficients indicate that with every one unit increase in the Honesty-Humility there is a 0.220 unit decrease in HAB composite scores, and for every one unit increase in agreeableness scores, there is a 0.235 point decrease in HAB composite scores, taking into account the other predictors in the model. The latter of the two is in line with the hypothesis for the third objective. No between vignette predictors were significant at $p < .01$. The Conditional R^2 value of variance for the random factor and fixed predictors combined was 0.596, and the Marginal R^2 value for the variance of the fixed predictors was 0.194.

Table 15

Outcome Variable: Hostile HAB composite SONA and Prolific samples

Model Parameters	**SONA sample**	**Prolific sample**
Intercept	2.719(.472)	2.761(.549)
Between participant predictors		
Gender *[Female - Male]*	0.257(.092)**	0.111(.128)
Trust	-0.066(.048)	-0.088(.057)
Anger BP	-0.051(.049)	
Hostility BP	0.161(.040)***	
Verbal Aggression BP	0.025(.039)	
Physical Aggression BP	0.113(.047)*	
Honesty Humility		-0.220(.084)**
Emotionality		0.068(.093)
Extraversion		-0.001(.080)
Agreeableness		-0.235(.085)**
Conscientiousness		-0.004(.092)
Openness to Experience		0.013(.070)
Between vignette predictors		
Perpetrator gender *[Male - Female]*	0.730(.473)	0.538(.547)
Perpetrator gender *[Not specified - Female]*	0.470(.390)	0.467(.451)
Perpetrator power *[Equal - Superior]*	1.109(.382)**	1.256(.442)*
Perpetrator known *[yes - no]*	-0.446(.320)	-0.765(.371)
Variance components		
Residual	1.198	1.182
Participants intercepts (random effect)	0.417	0.470
Vignette intercepts (random effect)	0.531	0.710
Conditional R^2	0.538	0.596
Marginal R^2	0.172	0.194
Model summary		
Log likelihood	-10257.052	-7397.089
Number of estimated parameters	14	16
Number of cases (observations/partic/vig)	6550/273/24	4726/197/24

Note. *p<.05, **p < .01, ***p < .001. Regression coefficients with standard errors in parentheses are unstandardized. REML refers to Restricted Maximum Likelihood Estimation.

Anger Split by Sample

The results of the models with Anger scores as the outcome variable split by sample is presented in *Table 16*. For the SONA sample, there was one significant within participant by vignette predictor, the HAB composite ($\beta = 0.722$, $p < .001$). This regression coefficient suggests that with one unit increase in HAB composite scores, there is a 0.722 unit increase in Anger scores for the SONA sample within participants, across the vignettes. No between participant or between vignette predictors were significant at $p < .01$ for the SONA sample. Although we would have expected stronger regression coefficients for the Anger BP and Hostility BP it is likely that those were adjusted downward due to their overlap with HAB composite. The Conditional R^2 value was 0.674 and is the variance in outcome explained by both the random factors and fixed predictors. The marginal R^2 of 0.480 is the proportion of variance explained by the fixed predictors only.

For the Prolific sample, there was also one significant within participant by vignette predictor, the HAB composite ($\beta = 0.799$, $p < .001$). This predicts that with a one unit increase in HAB composite scores, there would be an increase of 0.799 unit for Anger scores for the Prolific sample within participants, across vignette responses. One between participant predictor was significant at $p < .01$, agreeableness ($\beta = -0.168$, $p < .01$), suggesting that with every one unit increase in agreeableness, there would be a 0.168 unit decrease in Anger scores, taking into account the other predictors in the model. No between vignette predictors were significant. The Conditional R^2 value of variance for the random factor and fixed predictors combined was 0.729, and the Marginal R^2 value for the variance of the fixed predictors was 0.593.

Table 16

Outcome Variable: Vignette Anger Response for SONA and Prolific samples

Model Parameters	SONA sample	Prolific sample
Intercept	4.501(.328)	4.896(.285)
Within participant by vignette predictors		
Hostile composite	0.722(.011)***	0.799(.012)***
Between participant predictors		
Gender *[Female - Male]*	0.181(.076)*	-0.166(.091)
Trust	-0.005(.040)	0.032(.041)
Anger BP	0.026(.040)	
Hostility BP	0.067(.033)*	
Verbal Aggression BP	0.052(.033)	
Physical Aggression BP	0.038(.039)	
Honesty Humility		-0.064(.060)
Emotionality		0.141(.066)*
Extraversion		-0.000(.057)
Agreeableness		-0.168(.061)**
Conscientiousness		0.095(.066)
Openness to Experience		-0.081(.050)
Between vignette predictors		
Perpetrator gender *[Male - Female]*	-0.105(.327)	-0.199(.279)
Perpetrator gender *[Not specified - Female]*	0.081(.270)	-0.025(.230)
Perpetrator power *[Equal - Superior]*	-0.129(.265)	-0.134(.226)
Perpetrator known *[yes - no]*	0.095(.222)	-0.093(.189)
Variance components		
Residual	0.899	0.820
Participants intercepts (random effect)	0.283	0.230
Vignette intercepts (random effect)	0.253	0.182
Conditional R^2	0.674	0.729
Marginal R^2	0.480	0.593
Model summary		
Log likelihood	-9293.233	-6487.216
Number of estimated parameters	15	17
Number of cases (observations/partic/vig)	6542/273/24	4718/197/24

Note. *$p < .05$, **$p < .01$, ***$p < .001$. Regression coefficients with standard errors in parentheses are unstandardized. REML refers to Restricted Maximum Likelihood Estimation.

Expressing Disapproval Split by Sample

The results of the models with Expressing Disapproval as the outcome variable split by sample is presented in *Table 17*. For the SONA sample, both within participant by vignette predictors were significant: HAB composite ($\beta = 0.351$, $p < .001$) and Anger ($\beta = 0.437$, $p < .001$). At the between participant level, gender ($\beta = -0.217$, $p < .01$), BP verbal aggression ($\beta = 0.111$, $p < .01$), and BP physical aggression ($\beta = 0.144$, $p < .001$) were all significant. As gender was a binary predictor, the coefficient indicates that Expressing disapproval scores had means that were 0.217 point higher for male SONA participants compared with Female SONA participants. The BP verbal aggression coefficient suggests that with every one unit increase in verbal aggression scores, it is predicted that expressing disapproval will increase by 0.111 unit. Further, the BP physical aggression coefficient predicts that with every one unit increase in physical aggression, there will be a 0.144 unit increase in Expressing Disapproval, adjusting for the other predictors in the model. No between vignette predictors were significant for the SONA sample. The Conditional R^2 value was 0.625 and is the variance in outcome explained by both the random factors and fixed predictors. The marginal R^2 of 0.453 is the proportion of variance explained by the fixed predictors only.

For the Prolific sample, both within participant by vignette predictors were significant: HAB composite ($\beta = 0.269$, $p < .001$) and Anger ($\beta = 0.510$, $p < .001$). The HAB regression coefficient suggests that with every one unit increase in HAB composite scores, there is a 0.351 unit increase in Expressing Disapproval scores. Further, with every one unit increase in Anger scores, Expressing Disapproval is predicted to increase by 0.437 unit, taking into account the other predictors in the model. Three predictors at

the between participant level were significant at $p < .01$: gender ($\beta = -0.497$, $p < .001$), extraversion ($\beta = 0.272$, $p < 001$), and agreeableness ($\beta = -0.244$, $p < .01$). As gender was a binary predictor, the coefficient indicates that Expressing disapproval scores had means that were 0.497 point higher for male Prolific participants compared with Female Prolific participants. The extraversion coefficient suggests that with every one unit increase in extraversion, there is a 0.272 unit increase in Expressing disapproval. Further, the agreeableness coefficient predicts that with every one unit increase in agreeableness, there will be a 0.244 unit decrease in Expressing Disapproval, adjusting for the other predictors in the model. No between vignette predictors were significant at $p < .01$. The Conditional R^2 value of variance for the random factor and fixed predictors combined was 0.675, and the Marginal R^2 value for the variance of the fixed predictors was 0.527.

Table 17

Outcome Variable: Vignette Expressing Disapproval SONA and Prolific samples

Model Parameters	**SONA sample**	**Prolific sample**
Intercept	3.274(.412)	3.829(.345)
Within participant by vignette predictors		
Hostile composite	0.351(.018)***	0.269(.021)***
Anger	0.437(.016)***	0.510(.018)***
Between participant predictors		
Gender *[Female - Male]*	-0.217(.081)**	-0.497(0.111)***
Trust	0.059(.042)	0.116(.050)*
Anger BP	-0.057(.043)	
Hostility BP	-0.035(.035)	
Verbal Aggression BP	0.111(.035)**	
Physical Aggression BP	0.144(.041)***	
Honesty Humility		-0.095(.073)
Emotionality		-0.117(.081)
Extraversion		0.272(.070)***
Agreeableness		-0.244(.074)**
Conscientiousness		0.036(.080)
Openness to Experience		0.019(.061)
Between vignette predictors		
Perpetrator gender *[Male - Female]*	0.478(.413)	0.246(.337)
Perpetrator gender *[Not specified - Female]*	0.407(.340)	0.163(.278)
Perpetrator power *[Equal - Superior]*	0.502(.334)	0.611(.273)*
Perpetrator known *[yes - no]*	-0.226(.279)	-0.249(.229)
Variance components		
Residual	1.539	1.324
Participants intercepts (random effect)	0.301	0.338
Vignette intercepts (random effect)	0.401	0.265
Conditional R^2	0.625	0.675
Marginal R^2	0.453	0.527
Model summary		
Log likelihood	-10995.696	-7609.236
Number of estimated parameters	16	18
Number of cases (observations/partic/vig)	6542/273/24	4718/197/24

Note. *$p < .05$, **$p < .01$, ***$p < .001$. Regression coefficients with standard errors in parentheses are unstandardized. REML refers to Restricted Maximum Likelihood Estimation.

Discussion

The overarching purpose of this study was to improve on previous self-report measures of the Hostile Attribution Bias (HAB) in adults by refining the representation of hypothetical scenarios across many domains of daily life interpersonal conflicts. The measure itself is intended mainly for more in-depth research on the role of personality and social contextual properties in HAB.

The **first specific purpose** was to evaluate the factor structure of the six response items to the vignettes. We intended to build a composite score of HAB that would include four items depicting four variations on the theme of hostile intentionality: (1) maliciousness of the act, (2) perception that the instigator meant to provoke, and (3) to harm, and (4) the perceived rudeness of the behaviour. The first three questions were used from previous research, as ambiguity, maliciousness, and provocation are central to the definition of the HAB (Epps & Kendall, 1995; Wilkowski & Robinson, 2010; Gagnon & Rochat, 2017). The factor analyses showed strong evidence of the clustering of the first three items. The fourth item referring to the perceived rudeness of the behaviour correlated strongly with those three variables, but in the majority of vignettes, it also correlated strongly with the anger item. In other words, a person who perceived the behaviour as rude was highly likely to report that they would feel angry in the given situation. We opted to keep the rudeness item on the HAB factor based on the clear definitional association between perception of rudeness with the other three items. Rudeness is certainly different from anger conceptually, even if the two have a strong association.

Previous research has connected both anger and frustration with rudeness (Park et al., 2014), so it is not surprising that rudeness seemed to correlate highest with the outcome of anger. Further, rudeness is an interesting variable to explore as it can increase negative affect (including anger) and impact one's ability to complete tasks when witnessed (Gilam et al., 2020), all of which could connect with the HAB. Anger itself is also highly related to the HAB as well as our first three follow-up questions (maliciousness, provocation, and intentionality). However, 'rudeness' was considered more of a cognitive component, compared with affect, and therefore was able to be added with the first three follow-up questions to make up the hostile composite.

As part of the first purpose to build in seven categories of situational contexts in the HAB measure, we tested the extent that these seven categories would be revealed in a confirmatory factor analysis. The seven-factor model was not supported. Instead, a follow-up exploratory factor analysis revealed four factors that were not specific to any of the seven categories but instead depicted other properties of the vignettes. including a more direct confrontation (factor 1), frustration (factor 2), invalidation (factor 3), and exclusion (factor 4).

The **second purpose** of this study was to investigate the extent to which participants HAB responses on a given vignette would be influenced by individual differences and vignette differences as well as the interaction between the two. When looking at the intercept only models for the outcome variables, namely the HAB composite, the ICC for the vignettes was substantially larger (21% and 34% respectively), meaning more variation between the vignettes than between individuals. When the ICC for participants is small, there is more agreement in the perceptions of the

hypothetical situations. It is not uncommon in our experiences with MLM studies of repeated observations nested within individuals to see ICCs as high at .40 on measures where the individual differences are quite large. This typically happens in diary studies in which the same variable is assessed across occasions and where ICCs show how stability across occasions. In our case we were exposing participants to very different situations, and therefore the cross-situational stability is a bit lower.

Although we focused on the HAB outcome variable, we also included MLM models with anger as the outcome variable, and then expressing disapproval as the outcome. Our rationale for these additional models was based on the literature review suggesting that HAB can influence Anger responses which in turn can influence aggressive responses. In our study we used an assertive response instead (i.e., expressing disapproval) because verbal and aggressive responses are less frequent. The ICCs for participants and vignettes in the Anger responses were much closer, with .22 and .27 respectively. A larger difference was found between ICCs for expressing approval, where, once again, participants account for less variance (17%) than vignettes (32%). Overall, situational or contextual factors account for more variability in our set of vignettes on the three outcome measures.

In addition to obtaining a partition of the variance between Participants and Vignettes in the MLM analyses, it is also possible to take a closer look at the characteristics of the individual vignettes. For example, in the results we provided the 24 vignettes along with overall means as SDs on the HAB responses. The vignettes with the overall highest hostile composite (or HAB) scores include vignette 16 (at a bar where your friend is bumped into and groped), vignette 14 (you’re riding your bike and get cut

off by a car), and vignette 11 (a family member criticizes your achievements at the dinner table). Overall, these vignettes include scenarios that are direct and intentional, so it is not surprising that they had the highest HAB scores. Further, these three vignettes also had high factors loadings on the first factor in the exploratory analysis of vignettes, once again demonstrating they were more directly confrontational. On the other side, the vignettes with the lowest HAB scores were vignette 23 (a new person you're dating rejects your offer to hangout to see friends instead), vignette 15 (police pulls you out of the way at a street party), and vignette 20 (your professor is unwilling to remark an assignment). Interestingly, these three vignettes all had the highest loadings in factor 2; potential feelings of frustration. It may be interesting to look into this connection in future research to better understand the underpinnings of the situational contexts that influence perceptions of hostility/intentionality.

A participant gender by vignette gender interaction was added to the MLMs. Only weak evidence (i.e., $p < .05$) of this interaction was found, and only for the outcome variable of expressing disapproval, where women were less likely to express a disapproval if the target gender was female. Interestingly, this follows the same pattern as previous research into trust where women had a type of intergroup bias where they viewed women's faces as being more inherently trustworthy (Mattarozzi et al., 2015). However, this trend was not observed with men or with faces expressing certain facial expressions (Mattarozzi et al., 2015). This once again invites the importance of studying these two variables together, as trust and the HAB seem to have an important relationship. Statistical interactions involving individual differences often have small effects, which make them difficult to support without very large sample sizes. In our

study we had an adequate sample size of participants, but the sample size of 24 vignettes might be too low. One future direction with our sample will be to run a Monte Carlo study using the parameter estimates from the presented models and determine the post hoc power estimates for all the parameters of interest. This knowledge will help us make clear recommendations for sample size of stimuli recognizing that it is often not feasible to have large sets of stimuli. A second future analysis will be to re-analyze the models treating the vignettes as a fixed factor, which provides us with the trade-off of more power to investigate the predictors between vignettes and within person by vignettes, but weaker generalization to a larger population of vignettes.

The **third purpose** of the study was to investigate the relationships between the HAB composite score and personality variables consisting of dispositional trust, trait aggression, HEXACO six factor personality measures, and contextual variables such as power differential, gender of the instigator and whether the person was a stranger or not.

Our hypothesis regarding dispositional trust was supported within the MLMs and the preliminary correlation section, where we found a negative relationship between the hostile composite and dispositional trust. This supports the idea that those who tend to be wearier or trust the world less in their everyday life are more likely to interpret ambiguous situations as being hostile. This extends on previous research that has found a connection between trust and aggression (Rotenberg et al., 2021; Mattarozzi et al., 2015), as there is also a connection at the cognitive step, not just the behavioral.

When looking at agreeableness more generally, our prediction that it would also have a negative relationship with the hostile composite was supported. This hypothesis conforms with previous research linking agreeableness with aggression more generally,

specifically physical aggression (Barlett & Anderson, 2012). Further, this also conforms with previous research that connects trust and agreeableness (Crowe et al., 2018). No specific predictions were made for other common personality traits; however, it was found that honesty-humility was significant (negative) with the hostile composite. Those who have higher scores on Honesty-Humility scales are characterized as being less entitled, more humble and sincere, with low manipulation tendencies (Lee & Ashton, 2009). Given these traits, it is understandable that these individuals may be less likely to attribute malice to others during unclear situations. The other traits found to be significant were with the other single item factors (i.e., anger and expressing disapproval) and included emotionality and extraversion, respectively.

Trait aggression facets of hostility and physical aggression were related to the HAB composite in the MLMs. Interestingly, Anger was the only subscale of the Buss-Perry questionnaire not found to be significant with any outcome variables. However, Anger from the vignette follow-up questions was significant. The modest overlap between those two predictors (r = .29) in Table 11, may have rendered one of these non-significant in the models.

With respect to any gender differences between vignettes, only one vignette had a gender difference where men had a higher HAB score: vignette 9 (a protestor or demonstrator gets in your face). The other vignettes with the highest gender differences were vignettes 10 (on Instagram you find out that friends did not invite you to a gathering) and vignette 24 (your partner told you they were studying when they were out with friends), both where women had significantly higher HAB ratings. Based on these few vignettes, men had higher HAB scores when the circumstances related to a stranger

being confrontational, whereas women had higher HAB scores when those closest to them excluded or lied to them. Gender was found to be significant with expressing disapproval, suggesting men were more likely to report that they would express their disapproval within the scenarios. No other contextual factors that were explored were found to be significant.

Although we did not control for age in the MLM analyses, we did control for Sample which is highly correlated with age as shown in the descriptive statistics with an older mean age for the Prolific sample. Looking at the vignettes individually, the SONA sample had significantly higher mean HAB scores (controlling for gender) on the following vignettes: vignette 3 (a work colleague speaks over you), vignette 10 (on Instagram you find out that friends did not invite you to a gathering), vignette 13 (an officer pulls you over for an unsafe lane change), vignette 18 (at a university party, someone refers to your friend as gay), vignette 19 (while having a discussion with classmates, your male classmate says you don't make sense when expressing your opinion), vignette 20 (your professor is unwilling to remark your assignment), vignette 21 (you complete most of a group assignment, and when you press others for their part, they blame you for being overbearing), and vignette 23 (a new person you're dating rejects your offer to hangout to see friends instead). Most of these vignettes have to do with either a school/university age setting, or new life experiences (e.g., dating someone new), all of which may represent more sensitive issues for younger participants. These participants may be able to put themselves in these vignettes to a different degree than the Prolific sample, as they may be experiencing these situations in the here and now.

No formal hypotheses were proposed regarding the situational variables that were explored (i.e., power dynamics, gender of instigator, stranger status). However, for power dynamics, a strong effect was found in the models such that vignettes that depicted an instigator who has more power than the participant generated lower HAB composite scores than those where there was no imbalance. Previous research with power imbalances found that less trust is facilitated when a power imbalance is present (du Plessis et al., 2023), so due to the relationship seen between the HAB and trust, power imbalances could potentially extend to have a relationship with the HAB, but that was not support in the current study. However, previous research has found that those with less power seem to respond more based on situational factors compared with their dispositions (Guinote et al., 2012). In the sample of 24 vignettes, 5 of those vignettes depicted a powerful person. All five vignettes have low scores: zoom meeting with boss (#2), another interaction with boss (#4), the two police scenarios (13 and 15 which are quite mild in provocation), and the student trying to get a higher mark (20). The police scenarios were also quite interesting because it occurs within the backdrop of unfavorable opinions of the police being commonly broadcast on social media (e.g., Oglesby-Neal et al., 2019). We can say that these five vignettes were mild in provocation, and perhaps these people are not seen as malicious but just doing their job.

Limitations

While the study had some interesting and promising results, it is important to take into consideration some of the limitations. First, although we had a fairly substantial sample of 24 vignettes, this number is probably on the low side for treating it as a random factor and then expecting to find significant effects for predictors between vignettes. We

did find one of these effects (i.e., power differential) but it was a very large effect and partly a result of the fact that we constructed vignettes with the powerful instigator in conjunction with vignettes depicting relatively low provocation. One way to increase the number of vignettes while maintaining a feasible required completion time would be to increase the participant sample size and have participants complete a random sample rather than the total set of vignettes.

Second, one of the samples used participants from a single Canadian university, where the majority of students that participated were women. Statistics show that the majority of students enrolled within universities in Canada are women (Wilson et al., 2019). As this is not reflective of the general population, it may have impacted the results and the generalizability to the general public. The Prolific sample also saw more women participants, which is not uncommon for psychological research (Fanny, 2023). Given that the faculty of psychology is female-dominated, it could be that women are more likely to participate in psychological research, even outside of being a university student. Regardless, we did statistically control for the gender imbalance in the analyses and therefore this limitation applies only to generalizability and larger standard errors for the male sample.

The Buss-Perry Aggression Questionnaire may also pose some limitations. Previous research has found that, although the four-factor model is generally a good fit, a couple items within the hostility domain could be removed to improve factor loadings (Harris, 1995). Further, the study revealed that social desirability was related to the measure (Harris, 1997). Taking into account these limitations, the Buss-Perry Aggression Questionnaire is still considered a reasonable measure. Given the nature of our study, it is

possible that social desirability affected the results of the vignettes as well. Hostility and associated anger may not be a desired trait or state, and often has very negative connotations, even if it seems like a reasonable response to high provoking scenarios. Therefore, it is possible that some of the responses are underrated.

Future Directions and Implications

Vignettes are a unique and interesting tool used within research that allow the researchers to manipulate the context of a situation without having to expose an individual to the scenario itself. Written vignettes pose limitations of realism, and therefore, future research should focus on expanding the use of vignettes with technology. One study used headphones, added another level of realism or being able to immerse oneself within a scenario without directly being involved (Lien van der Schans et al., 2019). It can be difficult to accurately gauge one's authentic reaction when not exposed to the stimuli of interest, i.e., how you think you would respond may not be the way you actually respond when exposed. Therefore, building on Lien van der Schans et al., (2019) it would be interesting for research to create scenarios based in virtual reality to immerse people within ethical scenarios, so participants have a better gauge of their responses.

Finally, future research should continue exploring the relationship between dispositional trust and the HAB. Now that a potential connection has been supported by the current study, it would be interesting for researchers to further understand the factors that influence this complex relationship. For instance, previous research has found a complex relationship between power dynamics and trust (Guinote et al., 2012), so future

research may benefit from exploring potential mediators and moderators between trust and the HAB.

Various vignette research designs are used frequently in the social and health sciences (Atzmüller, & Steiner, 2010; Baguley et al. in press; Ouwehand et al. 2006; Wallander, 2009). Our focus on the design opens the door to an alternative or complementary way of investigating person by situation interactions. One way to see our approach is in terms of taking a wide lens or opting for "bandwidth" by using many vignettes and introducing more natural variability in dimensions such as intentionality-ambiguity and power differential. In addition to these design-modeling framework contributions, this research provides a clear direction for helping break the barriers of how to address conflict in society, taking into account the interaction between individual differences and the strength of particular social-contextual factors.

Further, we have added to the literature on the HAB, where we have found new support with the bias's connection with trait levels of trust. This builds on previous research that looked at related variables, but not directly with the HAB. It has been stated that less is known about the acquisition of the HAB (Smeijers et al., 2019), so looking at the bias's relationship with other variables, such as trust, allows for a more encompassing view, which could potentially help chip away at understanding the acquisition of the bias.

Finally, as research has established a strong connection between different types of aggression and violence with the HAB, understanding more about the bias, including both individual differences and the influence of situational contexts, will allow researchers to understand what to focus on for interventions. For instance, is there a way to increase dispositional trust, or a way to reframe certain types of situations, specifically

ones that may increase feelings of frustration, in order to decrease the likelihood of reactive aggression? Understanding connections between personality variables and contextual factors pave the way for answering these types of questions and helping to decrease societal issues, including reactive aggression.

Conclusion

In summary, the current study provides insight into the relationship between dispositional trust and HAB, along with other personality variables such as agreeableness, honesty-humility, emotionality, extraversion, openness to experience, conscientiousness, and trait aggression. Overall, the current research highlights the complex relationships between individual differences and social-contextual factors that influence the HAB, and therefore, the affective and behavioural reactions that accompany the cognitive bias. For the associated negative behaviours (i.e., reactive aggression) to be reduced in the future, research will need to continue to investigate these complicated relationships to get a better understanding of what impacts the HAB.

References

Archer, R. H. (1961). A preliminary study of hostility: The hostility ratio. *Psychiatric Quarterly, 35*, 562-574. https://doi.org/10.1007/BF01573624

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*, 340-345.

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology*, *6*, 128-138.

Aunger, R. (2020). Toward a model of situations and their context. *Review of General Psychology, 24*(3), 268–283. https://doi.org/10.1177/1089268020931767

Bailey, C. A., & Ostrov, J. M. (2008). Differentiating forms and functions of aggression in emerging adults: Associations with hostile attribution biases and normative beliefs. *Journal of Youth and Adolescence*, *37*(6), 713-722.

Baguley, T., Dunham, G., & Steer, D. (in press). Statistical modelling of vignette data in psychology. *British Journal of Psychology*.

Barlett, C. P., & Anderson, C. A. (2012). Direct and indirect relations between the Big 5 personality traits and aggressive and violent behavior. *Personality and Individual Differences, 52*(8), 870-875. https://doi.org/10.1016/j.paid.2012.01.029

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*, 122-142.

Bowen, K. N., Roberts, J. J., & Kocian, E. J. (2016). Decision making of inmates: Testing social information processing concepts using vignettes. *Applied Psychology in Criminal Justice, 12*(1), 1-17.

https://www.lib.uwo.ca/cgibin/ezpauthn.cgi?url=http://search.proquest.com/scholarly-journals/decision-making-inmates-testing-social/docview/1861353663/se-2

Brehm, J., & Rahn, W. (1997). Individual-level evidence for the causes and consequences of social capital. *American Journal of Political Science, 41*(3), 999-1023.

Brion, S., Mo, R., & Lount, R. B. (2019). Dynamic influences of power on trust: Changes in power affect trust in others. *Journal of Trust Research*, *9*(1), 6–27. https://doi.org/10.1080/21515581.2018.1552591

Bronchain, J. (2023). The Elemental Psychopathy Assessment: A personality-based measure of psychopathy. Jonason, Peter Karl [Ed]. Shining light on the dark side of personality: Measurement properties and theoretical advances. Gottingen, Germany: Hogrefe, Germany; pp. 151-158. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc21&NEWS=N&AN=2022-84375-015.

Bushman, B. J. (2016). Violent media and hostile appraisals: A meta-analytic review. *Aggressive Behavior*, *42*(6), 605–613. https://doi.org/10.1002/ab.21655

Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology, 63*(3), 452–459. https://doi.org/10.1037/0022-3514.63.3.452

Butler, M., & Maruna, S. (2009). The impact of disrespect on prisoners' aggression: outcomes of experimentally inducing violence-supportive cognitions. *Psychology, Crime & Law*, *15*(2–3), 235–250. https://doi.org/10.1080/10683160802190970

Chen, P., Coccaro, E. F., & Jacobson, K.C. (2014). Hostile Attribution Bias, Negative Emotional Responding, and Aggression in Adults: Moderating Effects of Gender

and Impulsivity. *Aggressive Behavior, 38*(1), 47-63. https://doi.org/10.1002/ab.21407

Choe, D. E., Lane, J. D., Grabell, A. S., Olson, S. L., & Eccles, J. S. (2013). Developmental Precursors of Young School-Age Children's Hostile Attribution Bias. *Developmental Psychology, 49*(12), 2245–2256. https://doi.org/10.1037/a0032293

Coccaro, E. F., Noblett, K. L., & Mccloskey, M. S. (2009) Attributional and emotional responses to socially ambiguous cues: validation of a new assessment of social/emotional information processing in healthy adults and impulsive patients. J. *Psychiatr Res, 43*, 915-925. 10.1016/j.jpsychires.2009.01.012

Combs, D. R., Penn, D. L., Wicher, M., & Waldheter, E. (2007). The Ambiguous Intentions Hostility Questionnaire (AIHQ): a new measure for evaluating hostile social-cognitive biases in paranoia. *Cognitive neuropsychiatry*, *12*(2), 128–143. https://doi.org/10.1080/13546800600787854

Couch, L. (1994). *The development of the Trust Inventory*. Unpublished mater's thesis, University of Tennessee, Knoxville, TN.

Couch, L. L., Adams, J. M., & Jones, W. H. (1996). The Assessment of Trust Orientation. *Journal of Personality Assessment, 67*, 305-323.

Crowe, M. L., Lynam, D. R., & Miller, J. D. (2018). Uncovering the structure of agreeableness from self-report measures. *Journal of Personality, 86*(5), 771-787. https://doi.org/10.1111/jopy.12358

D'Costa, M. (2021). Perpetrator Blame Attribution in Heterosexual Intimate Partner Violence: The Role of Gender and Perceived Injury. *Electronic Thesis*

and Dissertation Repository. 7791. https://ir.lib.uwo.ca/etd/7791

De Castro, B. O, Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2003). Hostile Attribution of Intent and Aggressive Behavior: A Meta-Analysis. *Child Development, 73*(3), 916-934. https://doi.org/10.1111/1467-8624.00447

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development*, 162-170.

Dodge K. A. (2006). Translational science in action: hostile attributional style and the development of aggressive behavior problems. *Development and psychopathology*, *18*(3), 791–814. https://doi.org/10.1017/s0954579406060391

Dodge, K. A., & Crick, N. R. (1990). Social Information Processing Bases of Aggressive Behavior in Children. *Personality and Social Psychology Bulletin, 16*(1), 8-22.

Dodge, K. A., Lansford, J. E., Burks, V. S., Bates, J. E., Pettit, G. S., Fontaine, R., & Price, J. M. (2003). Peer Rejection and Social Information-Processing Factors in the Development of Aggressive Behavior Problems in Children. *Child Development, 74*(2), 374-393. https://doi.org/10.1111/1467-8624.7402004

du Plessis, C., Nguyen, M. H. B., Foulk, T. A., & Schaerer, M. (2023). Relative power and interpersonal trust. *Journal of Personality and Social Psychology, 124*(3), 567-592. https://doi.org/10.1037/pspi0000401

Dunning, D., Fetchenhauer, D., & Schlösser, T. (2019). Why people trust: Solved puzzles and open mysteries. *Current Directions in Psychological Science, 28*(4), 366-371. https://doi.org/10.1177/0963721419838255

Epps, J., & Kendall, P. C. (1995). Hostile attributional bias in adults. *Cognitive Therapy and Research, 19*(2), 159-178. https://doi.org/10.1007/BF02229692

Ermisch, J., Gambetta, D., Laurie, H., Siedler, T., & Uhrig, S. C. N. (2009). Measuring people's trust. *Journal of the Royal Statistical Society, 172*(4), 749-769.

Fang, X., Zhang, K., Chen, J., Chen, M., Wang, Y., & Zhong, J. (2023). The Effects of Covert Narcissism on Chinese College Students Cyberbullying: The Mediation of Hostile Attribution Bias and the Moderation of Self-Control. *Psychology Research and Behavior Management*, *16*, 2353–2366. https://doi.org/10.2147/PRBM.S416902

Fanny. (2023, March 31). *Why do more women than men participate in clinical trials?*. Trialy. https://www.trialy.se/en/why-do-more-women-than-men-participate-in-clinicaltrials/#:~:text=In%20our%20work%20with%20recruiting,women%20than%20men%20that%20apply.

Ferraccio, B. J. (2018). *Implicit Racial Attitude, Hostile Attribution Bias, and the Desire to Enter Law Enforcement: Exploring Possible Relationships* (Order No. 10746389). Available from ProQuest Dissertations & Theses Global; ProQuest Dissertations & Theses Global Closed Collection. (2029172768). https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/dissertations-theses/implicit-racial-attitude-hostile-attribution-bias/docview/2029172768/se-2

Fields, S. K. (2013). *Narcissism and Intimate Partner Violence: An Establishment of the Link and Investigation of Multiple Potential Mediators* (Order No. 1524020). Available from ProQuest Dissertations & Theses Global. (1442774558). https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/dissertations

-theses/narcissism-intimate-partner-violence/docview/1442774558/se-2

Fiske, D. W. (1994). Two Cheers for the Big Five!. *Psychological Inquiry, 5*(2), 123-124. DOI: 10.1207/s15327965pli0502_5

Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Bessette, M., & Jolicoeur, P. (2016). Neural mechanisms underlying attribution of hostile intention in nonaggressive individuals: An ERP study. *International Journal of Psychophysiology*, *110*, 153–162. https://doi.org/10.1016/j.ijpsycho.2016.08.007

Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Brochu, A. F., Bessette, M., & Jolicoeur, P. (2016). An ERP study on hostile attribution bias in aggressive and nonaggressive individuals. *Aggressive Behavior, 43*(3), 217-229. https://doi.org/10.1002/ab.21676

Gagnon, J., & Rochat, L. (2017). Relationships between hostile attribution bias, negative urgency, and reactive aggression. *Journal of Individual Differences, 38*(4), 211-219. https://doi.org/10.1027/1614-0001/a000238

Gallucci, M. (n.d.) General analyses for linear models (Version 3.3.1) [Computer software]. Retrieved from https://gamlj.github.io/ and https://blog.jamovi.org/2023/10/14/gamlj.html

Gilam, G., Horing, B., Sivan, R., Weinman, N., & Mackey, S. C. (2020). The Decline in Task Performance After Witnessing Rudeness Is Moderated by Emotional Empathy-A Pilot Study. *Frontiers in psychology*, *11*, 1584. https://doi.org/10.3389/fpsyg.2020.01584

Godleski, S. A., & Murray-Close, D. (2023). Assessment of hostile intent attributions across escalating conflict stories. *Aggressive Behavior*, *49*(3), 249–260.

https://doi.org/10.1002/ab.22066

Golembiewski, R. T., & McConkie, M. (1975). The centrality of interpersonal trust in group processes. In C. L. Cooper (Ed.), *Theories of group processes* (pp. 131-185). New York: John Wiley

Guinote, A., Weick, M., & Cai, A. (2012). Does Power Magnify the Expression of Dispositions? *Psychological Science*, *23*(5), 475–482. https://doi.org/10.1177/0956797611428472

Hales, A. H., Kassner, M. P., Williams, K. D., & Graziano, W. G. (2016). Disagreeableness as a cause and consequence of ostracism. *Personality and Social Psychology Bulletin, 42*(6), 782-797. https://www.lib.uwo.ca/cgibin/ezpauthn.cgi?url=http://search.proquest.com/scholarly-journals/disagreeableness-as-cause-consequence-ostracism/docview/1789765878/se-2

Hansen-Brown, A. A., & Freis, S. D. (2021). Assuming the worst: Hostile attribution bias in vulnerable narcissists. *Self and Identity, 20*(2), 152–164. https://doi.org/10.1080/15298868.2019.1609574

Harris, J. (1995). Confirmatory Factor Analysis of the Aggression Questionnaire. *Behav. Res. Ther., 33*(8), 991-993.

Harris, J. (1997). A Further Evaluation of the Aggression Questionnaire: Issues of Validity and Reliability. *Behav. Res. Ther., 35*(11), 1047-1053.

Hiemstra, W., Orobio De Castro, B., & Thomaes, S. (2019). Reducing Aggressive Children's Hostile Attributions: A Cognitive Bias Modification Procedure. *Cognitive Therapy and Research, 43*(2), 387–398.

https://doi.org/10.1007/s10608-018-9958-x

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1975). Category Accessibility and Impression Formation. *Journal of Experimental Social Psychology, 13*, 141-154.

Holtzworth-Munroe, A., & Hutchinson, G. (1993). Attributing negative intent to wife behavior: the attributions of maritally violent versus nonviolent men.*Journal of Abnormal Psychology*, *102*(2), 206.

Johansson, A., Rötkönen, N., & Jern, P. (2021). Is the association between childhood \ maltreatment and aggressive behavior mediated by hostile attribution bias in women? A discordant twin and sibling study. *Aggressive Behavior, 47*(1), 28–37. https://doi.org/10.1002/ab.21928

Jonason, P. K., Webster, G. D., Li, N. P., Schmitt, D., & Crysel, L. C. (2012). The rise of the an-tihero in popular culture: A life history theory of the Dark Triad. *Review of General Psychology, 16*, 192–199. http://dx.doi.org/10.1037/a0027914.

Kastner, R. M. (2012). *The taste of aggression: A model for psychopathy and reactive aggression* (Order No. 1534145). Available from ProQuest Dissertations & Theses Global. (1314985498). https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/dissertations-theses/taste-aggression-model-psychopathy-reactive/docview/1314985498/se-2

Kim, S. H., Smith, R. H., & Brigham, N. L. (1998). Effects of Power Imbalance and the Presence of Third Parties on Reactions to Harm: Upward and Downward Revenge. *Personality & Social Psychology Bulletin*, *24*(4), 353–361. https://doi.org/10.1177/0146167298244002

Lakey, C. E., Kernis, M. H., Heppner, W. L., & Davis, P. J. (2005). *Development of the*

Hostile Attribution Bias Scale. Unpublished, University of Georgia.

Laue, C., Griffey, M., Ping-I, L., Wallace, K., Menno van, d. S., Horn, P., Pedapati, E., & Barzman, D. (2018). Eye Gaze Patterns Associated with Aggressive Tendencies in Adolescence. *Psychiatric Quarterly, 89*(3), 747-756. https://doi.org/10.1007/s11126-018-9573-8

Lee, K., & Ashton, M. C. (2014). The Dark Triad, the Big Five, and the HEXACO model. *Personality and Individual Differences*, *67*, 2–5. https://doi.org/10.1016/j.paid.2014.01.048

Lee, K., Ashton, M. C., Wiltshire, J., Bourdage, J. S., Visser, B. A., & Gallucci, A. (2013). Sex, power, and money: Prediction from the Dark Triad and honesty–humility. *European Journal of Personality, 27*, 169–184.

Lee, K., & Ashton, M. C. (2005). Psychopathy, machiavellianism, and narcissism in the Five-Factor model and the HEXACO model of personality structure. *Personality and Individual Differences, 38*, 1571–1582.

Li, M., He, Q., Zhao, J., Xu, Z., & Yang, H. (2022). The effects of childhood maltreatment on cyberbullying in college students: The roles of cognitive processes. *Acta Psychologica*, *226*(Complete). https://doi.org/10.1016/j.actpsy.2022.103588

Lien van der Schans, K., Karremans, J. C., & Holland, R. W. (2019). Mindful social inferences: Decentering decreases hostile attributions. *European Journal of Social Psychology, 50*(5), 1073-1087. https://doi.org/10.1002/ejsp.2657

Lin, P., Hsieh, C., Juan, C., Hossain, M. M., Erickson, C. A., Lee, Y., & Su, M. (2016). Predicting aggressive tendencies by visual attention bias associated with hostile

emotions. *PLoS ONE, 11*(2), 8. https://doi.org/10.1371/journal.pone.0149487

Mamuza, J. M., & Simourd, D. J. (1997). *The Hostile Interpretations Questionnaire*. Unpublished Manual.

Mattarozzi, K., Todorov, A., Marzocchi, M., Vicari, A., & Russo, P. M. (2015). Effects of gender and personality on first impression. *PLoS ONE, 10*(9), 13. https://doi.org/10.1371/journal.pone.0135529

Matthews, B. A., & Norris, F. H. (2006). When Is Believing "Seeing"? Hostile Attribution Bias as a Function of Self-Reported Aggression. *Journal of Applied Social Psychology, 32*(1), 1-31. https://doi.org/10.1111/j.1559-1816.2002.tb01418.x

McFall, R. M., & Dodge, K. A. (1982). Self-management and interpersonal skills learning. In P. Karoly & F. H. Kanfer (Eds.), *Self-management and behavior change: From theory to practice* (pp. 353-392). Elmsford, NY: Pergamon Press.

Menon, T., Sheldon, O. J., & Galinsky, A. D. (2014). Barriers to transforming hostile relations: Why friendly gestures can backfire. *Negotiation and Conflict Management Research, 7*(1), 17-37. https://doi.org/10.1111/ncmr.12023

Milburn, M. A., Niwa, M., & Patterson, M. D. (2013). Authoritarianism, Anger, and Hostile Attribution Bias: A Test of Affect Displacement. *Political Psychology, 35*(2), 225-243. https://doi.org/10.1111/pops.12061

Mooijman, M. (2023). Power dynamics and the reciprocation of trust and distrust.*Journal of Personality and Social Psychology,* https://doi.org/10.1037/pspi0000424

Naef, M., & Schupp, J. (2009). Measuring trust: Experiments and surveys in contrast and combination. *IZA Discussion Paper, 4087*, 1-44.

Nasby, W., Hayden, B., & DePaulo, B. M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology*, *89* (3), 459.

Nelson, J. A., & Perry, N. B. (2015). Emotional reactivity, self-control and children's hostile attributions over middle childhood. *Cognition and Emotion, 29*(4), 592-603. https://doi.org/10.1080/02699931.2014.924906

Oglesby-Neal, A., Tiry, E., & Kim, K. (2019, February). *Public perceptions of police on social media*. Justice Policy Center. https://www.urban.org/sites/default/files/publication/99789/public_perceptions_of_police_on_social_media_0.pdf

Orford, J. (1986). The rules of interpersonal complementarity: Does hostility beget hostility and dominance, submission? *Psychological Review, 93*(3), 365-377. https://doi.org/10.1037/0033-295X.93.3.365

Ouwehand, C., DeRidder, D. T. D., & Bensing, J. M. (2006). Situational aspects are more important in shaping proactive coping behaviour than individual characteristics: A vignette study among adults preparing for aging. *Psychology and Health*, *21*, 809-825.

Park, A., Ickes, W., & L. Robinson, R. (2014). More f#!%ing rudeness: reliable personality predictors of verbal rudeness and other ugly confrontational behaviors. *Journal of Aggression, Conflict and Peace Research*, *6*(1), 26–43. https://doi.org/10.1108/JACPR-04-2013-0009

Rotenberg, K. J., Manley, E., & Walker, K. M. (2021). The relation between young adults' trust beliefs in others and interpersonal hostility. *Aggressive*

Behavior, https://doi.org/10.1002/ab.21969

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35*, 651-665.

Schönenberg, M., & Jusyte, A. (2014). Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *European Archives of Psychiatry and Clinical Neuroscience, 264*(1), 61-69. https://doi.org/10.1007/s00406-013-0440-1

Shambare, R. (2016). What is Disposition to Trust. *IGI Global*. Retrieved May 30, 2023 from https://www.igi-global.com/dictionary/disposition-to-trust/35290

Simourd, D. J., & Mamuza, J. M. (2000). The Hostile Interpretations Questionnaire: Psychometric Properties and Construct Validity. *Criminal Justice and Behavior, 27*(5), 645-663. https://doi.org/10.1177/0093854800027005006

Smeijers, D., Bulten, E., & Brazil, I. A. (2019). The Computation of hostile biases (CHB) model: Grounding hostility biases in a unified cognitive framework. *Clinical Psychology Review, 73*(6). 10.1016/j.cpr.2019.101775

Spector, P. E., & Zhou, Z. E. (2014). The Moderating Role of Gender in Relationships of Stressors and Personality with Counterproductive Work Behavior. *Journal of Business and Psychology, 29*(4), 669-681. https://doi.org/10.1007/s10869-013-9307-8

Stirrat, M., & Perrett, D. I. (2010). Valid Facial Cues to Cooperation and Trust: Male Facial Width and Trustworthiness. *Psychological Science, 21*(3),

349. https://www.lib.uwo.ca/cgibin/ezpauthn.cgi?url=http://search.proquest.com/scholarly-journals/valid-facial-cues-cooperation-trust-male-width/docview/871575983/se-2

Tett, R. P., Toich, M. J., & Ozkum, S. B. (2021). Trait Activation Theory: A Review of the Literature and Applications to Five Lines of Personality Dynamics Research. *Annual Reviews, 8*, 199-233. https://doi.org/10.1146/annurev-orgpsych-012420-062228

Thierry, S. M., Twele, A. C., & Mondloch, C. J. (2021). Mandatory first impressions: Happy expressions increase trustworthiness ratings of subsequent neutral images. *Perception, 50*(2), 103-115. https://doi.org/10.1177/0301006620987205

Tremblay, P. F., & Belchevski, M. (2004). Did the instigator intend to provoke? A key moderator in the relation between trait aggression and aggressive behaviour. *Aggressive Behavior*, *30*, 409-424.

Tuente, S. K., Bogaerts, S., & Veling, W. (2019). Hostile attribution bias and aggression in adults – a systematic review. *Aggression and Violent Behavior, 26*, 66-81. https://doi.org/10.1016/j.avb.2019.01.009

Wallander, L. (2009). 25 years of factorial survey in sociology: A review. *Social Science in Research*, 38, 505-520.

Walters, G. D. (2022). Crime and social cognition: a meta-analytic review of the developmental roots of adult criminal thinking. *Journal of Experimental Criminology, 18*, 183-207. https://doi.org/10.1007/s11292-020-09435-w

Wang, Y., Cao, S., Dong, Y., & Xia, L. X. (2019). Hostile attribution bias and angry rumination: A longitudinal study of undergraduate students. *PloS one*, *14*(5),

e0217759. https://doi.org/10.1371/journal.pone.0217759

White, A. E., Kenrick, D. T., Li, Y. J., Mortensen, C. R., Neuberg, S. L., & Cohen, A. B. (2012). When nasty breeds nice: Threats of violence amplify agreeableness at national, individual, and situational levels. *Journal of Personality and Social Psychology, 103*(4), 622-634. https://doi.org/10.1037/a0029140

Wilkowski, B. M., & Robinson, M. D. (2010). The anatomy of anger: An integrative cognitive model of trait anger and reactive aggression. *Journal of Personality, 78*(1), 9-38. https://doi.org/10.1111/j.1467-6494.2009.00607.x

Wilson, C. A., Babcock, S. E., & Saklofske, D. H. (2019). Sinking or swimming in an academic pool: A study of resiliency and student success in first-year undergraduates. *Canadian Journal of Higher Education*, *49*(1), 60–84. https://doi.org/10.7202/1060824ar

Wrightsman, L. S. (1974). *Assumptions about human nature: A social-psychological approach*. Monterey, CA: Brooks Cole

Yamagishi, T. & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion, 18*, 129-166.

Yang, Y., Zhao, M., Dong, Y., & Xia, L. (2024). Longitudinal Associations Between Interpersonal Distrust and Social Aggression During College: Disentangling the Within-Person Process from Stable Between-Person Differences. *Journal of Youth and Adolescence, 53*, 849-862. https://doi.org/10.1007/s10964-023-01874-8

Yu, M., Saleem, M., & Gonzalez, C. (2014). Developing trust: First impressions and experience. *Journal of Economic Psychology, 43*, 16-29. https://doi.org/10.1016/j.joep.2014.04.004

Zajenkowska, A., Bodecka-Zych, M., Gehrer, N., Krejtz, K., Lawrence, C., Schoenenberg, M., & Jusyte, A. (2023). Gender differences in sensitivity to provocation and hostile attribution bias toward ambiguous facial cues in violent offenders and community-based adults. *Motivation and Emotion, 47*(1), 115–124. https://doi.org/10.1007/s11031-022-09972-z

Zajenkowska, A., Prusik, M., Jasielska, D., & Szulawski, M. (2021). Hostile attribution bias among offenders and non-offenders: Making social information processing more adequate. *Journal of Community & Applied Social Psychology, 31*(2), 241-256. https://doi.org/10.1002/casp.2493

Zmerli, S., & Newton, K. (2008). Social trust and attitudes toward democracy. *Public Opinion* Quarterly, 72(4), 706-724.

Appendix A

Demographic Questionnaire

Question 1: Age
What is your age?

- [Open-ended answer]

Question 2: Gender Identity
What gender do you identify as? 'Cisgender' means that your assigned sex at birth (e.g., female) matches your gender identity (e.g., woman). 'Transgender' means that your assigned sex at birth (e.g., male) does not match your gender identity (e.g., woman).

- Woman (cisgender)
- Man (cisgender)
- Transgender Man
- Transgender Woman
- Non-Binary
- Prefer not to say
- Self-Identify: Open-ended option

Question 3: Socioeconomic Status
How would you describe your own socio-economic status? I consider myself to be…

- Lower class
- Working class
- Lower middle class
- Middle class
- Upper middle class
- Upper class
- Prefer not to answer

Question 4: Ethnicity/Country of Origin
How would you describe your ethnicity? Ethnicity refers to a shared cultural heritage that distinguishes one group of people from another including ancestry, a sense of history, language, religion, foods, and clothing (e.g., Japanese, Eastern European, Nigerian, Greek, Canadian). You may type in more than one ethnicity.

- [Open-ended answer]

Question 5 SONA: Program
How would you describe the program(s) of study that you plan to pursue during your undergraduate studies (e.g., Psychology, Computer Science, Nursing, Music, Biochemistry, English)? If you are not certain, indicate uncertain.

- [Open-ended Answer]

Question 5 Prolific: Occupation

How would you describe your job/occupation (e.g., elementary school teacher, secretary, truck driver, etc...)? If you are not certain, indicate uncertain.

- [Open-ended Answer]

Appendix B

Vignettes

The participants were given the following instructions before reading the vignettes, "In this section you will be presented with 24 short written scenarios depicting social interactions that people may encounter in their lives. You are asked to imagine yourself in those situations and respond to the questions asking about your thoughts, feelings, and actions."

Vignettes 1-4: Work

Vignette #1:

"You cross paths with one of your new coworkers as you walk into a cafe. You are convinced that she recognized you and you say hello to her, but she passes by you without acknowledging you."

Vignette #2:

"You are in a Zoom meeting with your work team. Your supervisor has asked the team a question, and you provide what you think is a reasonable answer. Your supervisor responds: "I don't think that will work; let's see what the others have to say."

Vignette #3:

"You are having a discussion with a group of colleagues at a new job. Just as you start to speak, one of your colleagues talks over you."

Vignette #4:

"You hear about a new project at work. You know that you are the most qualified for the project, but your boss assigns it to the newest employee who has less experience than you."

Vignettes 5-9: Interactions with Public/Strangers

Vignette #5:

"You are on a bus sitting in an aisle seat. A man gets on the bus and steps on your foot as he walks past you, gives you a dirty look, and does not apologize."

Vignette #6:

"You are boarding a crowded bus. Just as you try to sit down next to a woman, she places her bag in what would have been your seat."

Vignette #7:

“You arrive at a leasing office to get more information about the apartments in the complex. The leasing agent notices you, but he continues talking on the phone for a while. Once he gets off the phone, he greets another customer that just walked in instead of you.”

Vignette #8:

“You are carrying a heavy load of groceries up to a check-out line at the grocery store and just as you are about to enter in line, a woman cuts in front of you. You end up dropping some items on the floor.”

Vignette #9:

“You are walking downtown and come across a crowd of people demonstrating for a cause that you are opposed to. Three demonstrators walk up to you yelling in your face: “You are with us or against us!” One of them puts their hand on your shoulder to stop you and get your attention.”

Vignettes 10-11: Family and Friends

Vignette #10:

“You are scrolling on Instagram one day, and a picture of your close friends at a party from the night before pops up, but no one had mentioned the gathering to you.”

Vignette #11:

“You are having dinner with your family. A family member starts criticizing your life choices and making fun of your achievements. You try to ignore them, but they keep going on and on.”

Vignettes 12-15: Driving and Police

Vignette #12:

“You have been looking for a parking spot for a while at the mall and finally see one up ahead. You put your signal on and proceed towards the spot, but another driver, who clearly saw you, rushes into the parking spot.”

Vignette #13:

“You are driving and made a right-hand turn. Right after you turn, you change lanes. A minute later you get pulled over, and the police officer tells you that you made an unsafe lane change as you did not use your signal. You are confident that you had your signal on.”

Vignette #14:

“You are riding your bike down the street. A car tries to go around you and ends up cutting you off, nearly clipping your front tire and then brakes abruptly at a red traffic

light. You are now both stopped at the traffic light, and the driver yells at you to get off the road."

Vignette #15:

"You are at a street party following a university football game and the street is packed with people. An ambulance needs to get through to help an injured person. The police are there trying to make space, and in the process one of the officers pulls you aggressively by the arm in order to clear the path."

Vignettes 16-18: Drinking Environments Including Parties

Vignette #16:

"You're dancing at a bar with a group of your friends, and a guy bumps into your female friend from behind and gropes her. Your friend looks a bit shaken and distressed."

Vignette #17:

"You have been waiting in line with your friends for over half an hour to get into a bar. You are to be the next ones to get in but two girls who appear to be very intoxicated push their way in front of you."

Vignette #18:

"You and a couple of friends are at a university student house party, and you are introduced to several people you don't know. One guy starts talking to you and looks at one of your male friends and back at you and whispers in your ear: "Your friend looks gay!""

Vignettes 19-21: Academics

Vignette #19:

"You are having a discussion with a couple of students. You disagree with the other students and express your opinion. One of the male students tells you that you are not making any sense."

Vignette #20:

"You have just received your mark on a research assignment, and you think the mark is too low and not reflective of your effort and the quality of your work. You decide to reach out to the professor asking her if she would reconsider the mark. However, she is unwilling to consider this request and replies that the mark is accurate."

Vignette #21:

"You are working on a group project with three other classmates. You have done most of the work and assigned tasks to the others, but they have not completed them by the

deadline. You confront them about it, but they blame you for being too bossy and not giving them enough time."

Vignettes 22-24: Intimate Relationships

Vignette #22:

"You are at a party with your friends. One of your friends who you are romantically interested in is avoiding you that evening and flirting with someone who you dislike."

Vignette #23:

"You have been on a few dates with a person who you are very attracted to, and you ask them if they would like to go out again next weekend, but they reply that they need some space and want to spend time with their group of friends."

Vignette #24:

"You discover that your romantic partner has lied to you, telling you that they were studying for an exam, but you find out that they were going out to a concert with friends."

Appendix C

Vignette Follow-up Questions

Includes the 6 follow-up questions given after each vignette. Please note, Question 1 was altered per vignette to include context for the participant. For example, it would reference the target individual of interest in the vignettes.

Question 1:
How malicious or mean to you perceive *this person* to have been? (1 = not malicious at all, 7 = extremely malicious)

Question 2:
How likely is it that this person meant to provoke you? (1 = not likely at all, 7 = extremely likely)

Question 3:
How likely is it that this person meant to harm you? (1 = not likely at all, 7 = extremely likely)

Question 4:
How rude to you perceive this person to have been? (1 = not at all rude, 7 = extremely rude)

Question 5:
How angry would you feel in this situation? (1 = not at all angry, 7 = extremely angry)

Question 6:
How likely is it that you would express disapproval in words or gestures to the person? (1 = not at all likely, 7 = extremely likely)

Appendix D

Recruitment Letters

Online SONA Recruitment Advertisement

Study Name: The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Brief Abstract: Using short descriptions of everyday social interactions that may include provocation, participants will rate their perceptions of how they would feel and react in those situations. They will also short measures of dispositional trust, anger and aggression.

Detailed Description

You are invited to participate in an online study investigating descriptions of hypothetical social interactions in which a person may provoke you. The aim of the study is to identify the types of social interactions that people find more provoking or frustrating and to investigate why some people perceive more or less provocation than others, and how these differences may relate to personality dispositions. Participants will read through 24 short vignettes and will be asked follow-up questions regarding their perceptions of how they might feel and react in the described situations. The online survey also contains short measures of dispositional trust, anger and aggression and a few demographic questions.

If you have questions about this study, please contact either:

Mackenzie Smith, Master's Thesis Student, Department of Psychology, Western University, by email at [redacted]

Paul Tremblay, PhD., Principal Investigator, Department of Psychology, Western University
Email: [redacted]
Phone: [redacted]

Eligibility Requirements
The requirements are to be 17 or older and an ability to understand English fluently (in order to understand the descriptions of the vignettes) and enrollment in a psychology course that uses the Psychology Research Participant pool at Western University.

Duration (Minutes)
30

Credits

0.5

Online Prolific Recruitment Advertisement

Study Name: The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Brief Abstract: Using short descriptions of everyday social interactions that may include provocation, participants will rate their perceptions of how they would feel and react in those situations. They will also complete a short measure of dispositional trust and a general personality questionnaire.

Detailed Description

You are invited to participate in an online study investigating descriptions of hypothetical social interactions in which a person may provoke you. The aim of the study is to identify the types of social interactions that people find more provoking or frustrating and to investigate why some people perceive more or less provocation than others, and how these differences may relate to personality dispositions. Participants will read through 24 short vignettes and will be asked follow-up questions regarding their perceptions of how they might feel and react in the described situations. Please note that some of the situations deal with scenarios that may be more common for younger adults (college/university age) and others may deal with behaviors more common in workplace scenarios. However, you are still invited to imagine yourself in these scenarios and answer any and all questions that you are comfortable with, regardless of age. The online survey also contains short measures of dispositional trust a personality questionnaire, and a few demographic questions.

If you have questions about this study, please contact either:

Mackenzie Smith, Master's Thesis Student, Department of Psychology, Western University, by email at [redacted].

Paul Tremblay, PhD., Principal Investigator, Department of Psychology, Western University
Email: [redacted]
Phone: [redacted]

Eligibility Requirements
The requirements are an age 18 or older; residing in Canada, the United States of America, the United Kingdom, or Australia; and an ability to understand English fluently (in order to understand the description of the vignettes).

Duration (Minutes)
30

Compensation

£4.50 (approximately $5.63 USD), a rate of £9.00 ($11.25 USD) per hour

Appendix E

Letters of Information and Consent

SONA

Letter of Information & Consent

Project Title: The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Principal Investigator: Dr. Paul F. Tremblay Department of Psychology, [redacted]

Additional Researchers: Mackenzie Smith, [redacted]

You are invited to participate in an online study investigating the influence of individual differences in personality dispositions on perceptions of hostility in the actions of others in various social situations. This project is conducted by Mackenzie Smith, as part of her master's thesis, and her collaborator and supervisor Dr. Paul Tremblay in the department of Psychology. The purpose of this letter is to provide you with information to make an informed decision regarding participation in this research.

Study Information: The aim of the study is to identify the types of social interactions that people find more provoking or frustrating and to investigate why some people perceive more or less provocation than others, and how these differences may relate to personality dispositions. If you consent to participate, you will be asked to read through 24 short vignettes and answer on rating scales follow-up questions regarding your perceptions of how you might feel and react in the described situations. The online survey also contains short measures of dispositional trust, anger and aggression and a few demographic questions. The survey takes approximately 30 minutes.

Potential Risks and Resource Information: The risks of participating in this study are considered minimal. Some may feel discomfort when thinking about some of the potentially provoking interactions, but you are not compelled to respond to any question that you find unpleasant. If you feel distressed while thinking about potentially provoking

interactions, some available on-campus services are listed here: Student Development Services is available at 519-661-3031 or Student Health Services at 519-661-3030. If you feel you need academic support, the Student Success Centre is available at 519-661-3559, and Peer Support Network is available at 519-661-3574.

Benefits to Participation: You will not benefit directly from this research. However, your participation in this study will provide valuable information regarding the relationships between different personality variables and their influence on reactions to potentially provoking scenarios, as well as the influence that situational contexts have on the responses (e.g., in a work/school setting, with friends or family, etc..).

Compensation: Participants enrolled in the introductory psychology course will be rewarded with a 0.5 research credit toward that course. For students in other non-introductory psychology courses, you will be compensated as indicated on your relevant course syllabus.

Your Rights as a Participant: Your participation in this research is completely voluntary, and you have the right to withdraw at any stage, without needing to provide a reason. You may choose not to respond to any question you find distressing or uncomfortable. You can exit the questionnaire at any given moment without any penalty. Participating in this study does not compromise any of your legal rights. If you want to discontinue the study, you may do so at any point by closing the survey window. Any data collected before your withdrawal will be excluded from our analysis. However, due to the anonymity of the data collected, once your responses to the survey have been submitted, it would be impossible for the researchers to remove your data.

Confidentiality: All information that we obtain from you is confidential. Your responses to our questionnaires will be collected anonymously through a third party, secure online survey platform called Qualtrics. Qualtrics uses encryption technology and restricted access authorizations to protect the privacy and security of all data collected and retained, including personal information. In addition, Western's Qualtrics server is in Ireland,

where privacy standards are maintained under the European Union's General Data Protection Regulation, which is consistent with Canada's privacy legislation. Please refer to Qualtrics' Privacy Policy (https://www.qualtrics.com/privacystatement/) for more details about Qualtrics' information management practices. The data will then be exported from Qualtrics and securely stored on Western University's server. The collected data will be stored electronically in password-protected, encrypted files for 7 years, per Western University guidelines. While we do our best to protect your information, there is no guarantee that we will be able to do so.

Usually, it is only the research staff that will have access to the data. However, representatives of Western University's Non-Medical Research Ethics Board may require access to your study-related records to monitor the conduct of the research. In addition, in the interest of promoting research transparency and facilitating independent scrutiny of our data, anonymized data from the study, excluding all demographics except for age and gender, and containing no information that could identify you, may be uploaded onto the lab's Open Science Framework (OSF; https://osf.io) site so that data may be inspected and analyzed by other researchers.

Contacts for Further Information: If you have any questions about your rights as a research participant or the conduct of this study, you may contact The Office of Human Research Ethics (519) 661-3036 or toll-free at 1-844-720-9816, email: ethics@uwo.ca. You may also choose to direct any questions about this research or to address any concerns about your participation to Dr. Paul Tremblay at The University of Western Ontario, in London Ontario by email at: [redacted]

This letter is yours to keep for future reference. You can download it here.

Consent: Before beginning the survey online, you will be asked to indicate your acknowledgement of having read this letter of information and your consent to participate by clicking yes or no, below this letter. By clicking 'yes' below, you indicate that you have read the letter of information, and voluntarily consent to participate in this study.

Consent to participate in the study: Do you consent to participate in this survey?

- Yes I consent to participate in the study

No I do not consent to participate in the study

Prolific

Letter of Information & Consent

Project Title: The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Principal Investigator: Dr. Paul Tremblay, Department of Psychology, [redacted]

Additional Researcher: Mackenzie Smith, [redacted]

You are invited to participate in an online study investigating the influence of individual differences in personality dispositions on perceptions of hostility in the actions of others in various social situations. This project is conducted by Mackenzie Smith, as part of her master's thesis, and her collaborator and supervisor Dr. Paul Tremblay in the department of Psychology. The purpose of this letter is to provide you with information to make an informed decision regarding participation in this research.

Study Information: The aim of the study is to identify the types of social interactions that people find more provoking or frustrating and to investigate why some people perceive more or less provocation than others, and how these differences may relate to personality dispositions. If you consent to participate, you will be asked to read through 24 short vignettes and answer on rating scales follow-up questions regarding your perceptions of how you might feel and react in the described situations. The online survey also contains short measures of dispositional trust, anger and aggression and a few demographic questions. The survey takes approximately 30 minutes.

Potential Risks and Resource Information: The risks of participating in this study are considered minimal. Some may feel discomfort when thinking about some of the potentially provoking interactions, but you are not compelled to respond to any question that you find unpleasant.

Available Resources:

United States:

- You can call the suicide and crisis lifeline anytime by dialing 988
- Text HOME to 741741 to connect with a volunteer Crisis Counselor 24/7 (https://www.crisistextline.org/)

Canada:

- You can visit suicide.ca or you can contact Talk Suicide Canada 24/7 by phone at 1-833-456-4566 or by text at 45645, or for residents of Quebec, call 1-866-277-3553.
- For phone counselling available 24/7 call 1-866-585-0445

United Kingdom:

- You can Contact Samaritans any time on 116 123. They offer a listening service (https://www.samaritans.org/how-we-can-help/contact-samaritan/talk-us-phone/).
- Text SHOUT at any time to 85258 (https://giveusashout.org/get-help/how-shout-works/)

Australia:

- Call lifeline Australia 24/7 for crisis support at 13 11 14 (https://www.lifeline.org.au/131114)
- Text (SMS) 0477 13 11 14 or receive online chat support at https://www.lifeline.org.au/crisis-chat/

Benefits to Participation: You will not benefit directly from this research. However, your participation in this study will provide valuable information regarding the relationships between different personality variables and their influence on reactions to potentially provoking scenarios, as well as the influence that situational contexts have on the responses (e.g., in a work/school setting, with friends or family, etc..).

Compensation: Participants will be rewarded with £4.50 (approximately $5.63 USD) for their participation, a rate of £9.00 ($11.25 USD) per hour, which will be added to your Prolific profile upon providing consent at the beginning of the study.

To receive compensation for your participation in the study you must submit the survey; however, you are not compelled to respond to any items in the survey. Due to the anonymous nature of your data, once your survey responses have been submitted, the researchers will be unable to withdraw your data.

Your Rights as a Participant: Your participation in this study is voluntary. You may decide not to participate in this study. Even if you consent to participate, you have the right to not answer individual questions or to withdraw from the study at any time.

Confidentiality: All information that we obtain from you is confidential. Your responses to our questionnaires will be collected anonymously through a third party, secure online survey platform called Qualtrics. Qualtrics uses encryption technology and restricted access authorizations to protect the privacy and security of all data collected and retained, including personal information. In addition, Western's Qualtrics server is in Ireland. Please refer to Qualtrics' Privacy Policy (https://www.qualtrics.com/privacystatement/) for more details about Qualtrics' information management practices. The data will then be exported from Qualtrics and securely stored on Western University's server.

Usually, it is only the research staff that will have access to the data. However, representatives of Western University's Non-Medical Research Ethics Board may require access to your study-related records to monitor the conduct of the research. In addition, in the interest of promoting research transparency and facilitating independent scrutiny of our data, anonymized data from the study, excluding all demographics except for age and gender, and containing no information that could identify you, may be uploaded onto the lab's Open Science Framework (OSF; https://osf.io) site so that data may be inspected and analyzed by other researchers.

Contacts for Further Information: If you have any questions about your rights as a research participant or the conduct of this study, you may contact The Office of Human Research Ethics (519) 661-3036 or toll-free at 1-844-720-9816, email: ethics@uwo.ca. You may also choose to direct any questions about this research or to address any concerns about your participation to Dr. Paul Tremblay at Western University, in London, Ontario by email at: [redacted]

This letter is yours to keep for future reference. You can download it into a new window here.

Consent. Before beginning the survey online, you will be asked to indicate your acknowledgement of having read this letter of information and your consent to participate by clicking yes or no, below this letter. By clicking 'yes' below, you indicate that you have read the letter of information, and voluntarily consent to participate in this study.

Do you consent to participate in this survey?

○ Yes I consent
○ No I do not consent

Appendix F

Debrief Letters

SONA

Debrief

Project Title: The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Principal Investigator: Dr. Paul Tremblay, Department of Psychology at Western University, [redacted]
Co-investigator: Mackenzie Smith, Masters Thesis Student, [redacted]

Thank you for your participation in our study. The objective of this study is to investigate individual differences in responses to potentially provoking situations and how these individual differences may relate to dispositional measures of trust, anger, hostility and aggression. We also want to investigate differences in reactions to situations across different social contexts (e.g., at work, in school/university or in interactions with strangers). The investigation will help us better understand the influence of the hostile attribution bias (HAB; defined as a tendency to see more hostility, malice, or intent to harm than might exist in the actions of others) on anger and confrontation.

In social interpersonal interactions involving some level of provocation, the intent of the instigator is not always clear to the receiver or target of that provocation. The interpretation of ambiguous interpersonal situations as hostile is not uncommon and can lead to confrontation by the target and potentially an exchange of aggressive reactions and behaviours. Research has confirmed individual differences in the HAB in adults (Bowen et al., 2016) as well as in children (Nelson & Perry, 2015). The HAB has also been shown to lead to aggressive responses (Bowen et al., 2016). Ongoing research also investigates whether high levels of the HAB may also be associated with narcissism

(Fields, 2013), dominant-submissive behaviors (Orford, 1986), and general personality dimensions such as agreeableness (Barlett & Anderson, 2012). As part of our study we will investigate whether high HAB is associated with low levels of trust and whether HAB differs or remains fairly stable in various domains of the social context and the role of the gender and power dynamics of the people involved in the social interaction.

References.

Barlett, C. P., & Anderson, C. A. (2012). Direct and indirect relations between the Big 5 personality traits and aggressive and violent behavior. *Personality and Individual Differences, 52*(8), 870-875. https://doi.org/10.1016/j.paid.2012.01.029

Bowen, K. N., Roberts, J. J., & Kocian, E. J. (2016). Decision making of inmates: Testing social information processing concepts using vignettes. *Applied Psychology in Criminal Justice, 12*(1), 1-17. https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/scholarly-journals/decision-making-inmates-testing-social/docview/1861353663/se-2

Fields, S. K. (2013). *Narcissism and Intimate Partner Violence: An Establishment of the Link and Investigation of Multiple Potential Mediators* (Order No. 1524020). Available from ProQuest Dissertations & Theses Global. (1442774558). https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/dissertations-theses/narcissism-intimate-partner-violence/docview/1442774558/se-2

Nelson, J. A., & Perry, N. B. (2015). Emotional reactivity, self-control and children's hostile attributions over middle childhood. *Cognition and Emotion, 29*(4), 592-603. https://doi.org/10.1080/02699931.2014.924906

Orford, J. (1986). The rules of interpersonal complementarity: Does hostility beget hostility and dominance, submission? *Psychological Review, 93*(3), 365-377. https://doi.org/10.1037/0033-295X.93.3.365

If you have any questions about this research study, please contact Dr. Paul Tremblay, Department of Psychology, Western University, [redacted], email: [redacted]

If you have any questions about your rights as a research participant or the conduct of this study, you may contact The Office of Human Research Ethics (519) 661-3036, email: ethics@uwo.ca. Or contact the toll-free long-distance phone number for the Office of Human Resource Ethics at 1-844-720-9816.

If you or someone you know is experiencing distress, please contact your local mental health services.

Available Campus Resources:

- Health and Wellness, Western University (519)661-2111
- Mental Health Support Counselling, Western University (519) 661-3030
- For resources specific to gender based violence, please visit https://www.uwo.ca/health//student_support/survivor_support/index.html
- For available resources specific to gender identity and sexual orientation, please visit https://www.uwo.ca/health//psych/2SLGBTQIA+.html

Prolific

Debrief

Project Title: The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Principal Investigator: Dr. Paul Tremblay, Department of Psychology at Western University, [redacted]
Co-investigator: Mackenzie Smith, Master's Thesis Student, [redacted]

Thank you for your participation in our study. The objective of this study is to investigate individual differences in responses to potentially provoking situations and how these individual differences may relate to dispositional measures of trust, anger, hostility and aggression. We also want to investigate differences in reactions to situations across different social contexts (e.g., at work, in school/university or in interactions with strangers). The investigation will help us better understand the influence of the hostile attribution bias (HAB; defined as a tendency to see more hostility, malice, or intent to harm than might exist in the actions of others) on anger and confrontation.

In social interpersonal interactions involving some level of provocation, the intent of the instigator is not always clear to the receiver or target of that provocation. The interpretation of ambiguous interpersonal situations as hostile is not uncommon and can lead to confrontation by the target and potentially an exchange of aggressive reactions and behaviours. Research has confirmed individual differences in the HAB in adults (Bowen et al., 2016) as well as in children (Nelson & Perry, 2015). The HAB has also been shown to lead to aggressive responses (Bowen et al., 2016). Ongoing research also investigates whether high levels of the HAB may also be associated with narcissism (Fields, 2013), dominant-submissive behaviors (Orford, 1986), and general personality dimensions such as agreeableness (Barlett & Anderson, 2012). As part of our study we will investigate whether high HAB is associated with low levels of trust and whether

HAB differs or remains fairly stable in various domains of the social context and the role of the gender and power dynamics of the people involved in the social interaction.

References.

Barlett, C. P., & Anderson, C. A. (2012). Direct and indirect relations between the Big 5 personality traits and aggressive and violent behavior. *Personality and Individual Differences, 52*(8), 870-875. https://doi.org/10.1016/j.paid.2012.01.029

Bowen, K. N., Roberts, J. J., & Kocian, E. J. (2016). Decision making of inmates: Testing social information processing concepts using vignettes. *Applied Psychology in Criminal Justice, 12*(1), 1-17. https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/scholarly-journals/decision-making-inmates-testing-social/docview/1861353663/se-2

Fields, S. K. (2013). *Narcissism and Intimate Partner Violence: An Establishment of the Link and Investigation of Multiple Potential Mediators* (Order No. 1524020). Available from ProQuest Dissertations & Theses Global. (1442774558). https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/dissertations-theses/narcissism-intimate-partner-violence/docview/1442774558/se-2

Nelson, J. A., & Perry, N. B. (2015). Emotional reactivity, self-control and children's hostile attributions over middle childhood. *Cognition and Emotion, 29*(4), 592-603. https://doi.org/10.1080/02699931.2014.924906

Orford, J. (1986). The rules of interpersonal complementarity: Does hostility beget hostility and dominance, submission? *Psychological Review, 93*(3), 365-377. https://doi.org/10.1037/0033-295X.93.3.365

If you have any questions about this research study, please contact Dr. Paul Tremblay, Department of Psychology, Western University, [redacted], email: [redacted]

If you have any questions about your rights as a research participant or the conduct of this study, you may contact The Office of Human Research Ethics, Western University, at (519) 661-3036, email: ethics@uwo.ca. Or contact the toll-free long-distance phone number for the Office of Human Resource Ethics at 1-844-720-9816.

If you or someone you know is experiencing distress, please contact your local mental health services.

Other Available Resources:

United States:

- You can call the suicide and crisis lifeline anytime by dialing 988
- Text HOME to 741741 to connect with a volunteer Crisis Counselor 24/7 (https://www.crisistextline.org/)

Canada:

- You can visit suicide.ca or you can contact Talk Suicide Canada 24/7 by phone at 1-833-456-4566 or by text at 45645, or for residents of Quebec, call 1-866-277-3553.
- For phone counselling available 24/7 call 1-866-585-0445

United Kingdom:

- You can Contact Samaritans any time on 116 123. They offer a listening service (https://www.samaritans.org/how-we-can-help/contact-samaritan/talk-us-phone/).
- Text SHOUT at any time to 85258 (https://giveusashout.org/get-help/how-shout-works/)

Australia:

- Call lifeline Australia 24/7 for crisis support at 13 11 14 (https://www.lifeline.org.au/131114)

Text (SMS) 0477 13 11 14 or receive online chat support at

https://www.lifeline.org.au/crisis-chat/

Appendix G

Ethics Approval

Western Research

Date: 28 November 2023

To: Professor Paul F. Tremblay

Project ID: 123449

Study Title: The Influence of Dispositions and Everyday Social Factors on the Hostile Attribution Bias

Short Title: Dispositions and Social Factors on the Hostile Attribution Bias

Application Type: NMREB Initial Application

Review Type: Delegated

Full Board Reporting Date: 12/Jan/2024

Date Approval Issued: 28/Nov/2023 14:57

REB Approval Expiry Date: 28/Nov/2024

Dear Professor Paul F. Tremblay

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the above mentioned study, as of the date noted above. NMREB approval for this study remains valid until the expiry date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

This research study is to be conducted by the investigator noted above. **All other required institutional approvals and mandated training must also be obtained prior to the conduct of the study.**

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Prolific_Survey	Online Survey		
SONA_Survey	Online Survey		
Recuitment Letter_SONA_Updated	Recruitment Materials		
Debrief Letter_Prolific_Updated	Debriefing document	08/Nov/2023	2
Debrief Letter_SONA_Updated	Debriefing document	09/Nov/2023	2
Recruitment Letter_Prolific_Updated	Recruitment Materials	30/Oct/2023	2
Letter of Information_SONA_Updated	Implied Consent/Assent	08/Nov/2023	2
Letter of Information_Prolific_Updated	Implied Consent/Assent	08/Nov/2023	2

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Dr. Trevor Bieber, Research Ethics Officer on behalf of Dr. Isha DeCoito, NMREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Curriculum Vitae

Name:	Mackenzie Smith
Post-secondary Education and Degrees:	Western University London, Ontario, Canada 2018-2022 B.A.
Honours and Awards:	Dean's Honor Roll Award 2020, 2021, 2022
Related Work Experience	Teaching Assistant Western University 2022-2024 Community Policing Assistant Ontario Provincial Police 2023 Undergraduate Student Research Internship Western University 2022

Publications:

Undergraduate Student Research Internship - Poster

Smith, M. C. (2022). Connecting the hostile attribution bias with habit formation. Western Repository.

Undergraduate thesis –

Smith, M. C. (2022). The relationship between violent crime and alcohol consumption at the country level. Western Repository.