Fall 1-6-2021

A Study of How We Perceive People

Muriel Kosaka
CUNY Hunter College

A Study of How We Perceive People

by

Muriel Kosaka

Submitted in partial fulfillment
of the requirements for the degree of
Master of Arts General Psychology, Hunter College
The City University of New York

2020

12/14/2020	Professor Jason Young
Date	Thesis Sponsor
12/14/2020	Professor Rebecca Huselid
Date	Second Reader

Abstract

In a now classic study, Srull and Wyer (1979) found that by priming participants with hostility related stimuli in a sentence unscrambling task caused subsequent judgment of a person (Donald) behaving ambiguously to be perceived as more hostile. As part of a larger, multi-lab collaborative study conducted by Randy McCarthy, the purpose of this study is to replicate the findings found in Srull and Wyer (1979) by focusing on two conditions that demonstrated a clear priming effect, the use of 30 sentences in the sentence unscrambling task and the immediate testing condition. Participants first completed a sentence unscrambling task, followed by the reading of a brief vignette about a man named Donald behaving ambiguously hostile and then rated him on a set of personality traits. As an indicator of hostility, we examined ratings of three traits (*hostile, unfriendly,* and *dislikeable*) to create an average hostility score for participants of both groups. Participants in the experimental group who descrambled mostly hostile sentences rated Donald as 0.18 points more hostile than did those who descrambled neutral sentences. Analysis found these results not to be significant. Results indicated a significant interaction for the trait rating of *dislikeable,* females in the control group rated the trait more negatively compared to males in the control group. To provide additional context for the results in our sample, we re-analyzed data from three external samples in this multi-site project. Results from all samples suggest that the currently used methods do not produce a hostile priming effect as detected in the original Srull and Wyer (1979) study.

A Study of How We Perceive People

The essential goal of studies in social psychology is to examine the subtle and unanticipated impact of individuals' social environments on their behaviors and thoughts (Molden, 2014). Priming is a phenomenon in which exposure to a stimulus influences how individuals respond to a subsequent stimulus. In social psychology, priming is defined in terms of how actions/events influence the activation of stored knowledge (Higgins, 1996), primarily studying how exposing participants to different types of information activates social representations such as stereotypes or traits, that then may affect subsequent perceptions or actions. Different types of priming include semantic priming, which occurs when the response to a target (i.e. violin) is preceded by a semantically-related prime (i.e. trumpet) because both the target and prime are from the same category and share similar features (Sperber et al., 1979). Repetition priming involves one's implicit memory in which exposure to a stimulus enables the subsequent processing of the same or related stimulus (Tulving & Schacter, 1990). Assimilative priming occurs when judgements are consistent with the primed category (i.e. believing an individual to be hostile after being exposed to hostile-related stimuli), such as in Srull and Wyer (1979). Such studies show that mere exposure to socially relevant stimuli can prime actions, perceptions, and emotions, often outside of a person's awareness (Takarada & Nozaki, 2018; Bargh, Chaiken, Govender, & Pratto, 1992; Bargh, Chaiken, Raymon, & Hymes, 1996). Research on priming has become an important area of study because it involves a diverse set of phenomena and processes whose limits are still being explored.

The importance of priming originates from the consequences it can have for an individual's behaviors, thoughts, and how s/he interacts with others. It can affect a person's perception towards others and the way they interpret their behavior. For instance, hostile priming

due to watching an aggressive movie or playing an aggressive video game may trigger a high probability that the participant will later perceive someone else's ambiguous behaviors as aggressive. In such a situation, their social perception has been changed due to increased contact with aggressive traits without the awareness of priming exposure. The effect of priming social categories is the same as the effect of priming individual behaviors (Anderson & Bushman, 2001). For example, when participants were primed with the elderly category they were observed to have more conservative attitudes and participants primed with the aggressive behaviors via violent video games, were observed to act more aggressively. Researchers have proposed that even more complex social traits can be primed without the awareness of the perceivers (Kawakami, Dovidio, & Dijksterhuis, 2003), such that participants may assimilate their attitudes to those of the primed category even when the shift is in a direction that is not personally or socially desirable, demonstrating a conversion rather than a compliance process in their attitudes. This phenomenon is known as the behavior-perception link. Since behavior and perception overlap, priming can cause individuals to behave in accordance with the primed concept (Bargh, Chen, & Burrows, 1996). Priming behaviors can also involve activating mental representations of goals (Dijksterhuis & Aarts, 2010). For example, in a study conducted by Bargh et al. (2001), they found that behavioral goals such as performing well on a task can become activated without any consciously made choice in the form of performance on a word search puzzle and the goal to cooperate on a task can become activated in the form of a resource-dilemma task (Komorita & Parks, 1995). Bargh, Chen, and Burrows (1996) were able to prime rude behavior using rudeness-related stimuli in a sentence unscrambling task, where participants were given a list of 30 items consisting of five-word phrases (i.e. "they her bother see usually") and were asked to construct a four-word grammatically correct sentence ("they usually bother her"). The sentence

unscrambling task was developed by Costin (1969) as an implicit means to prime participants with hostility. Participants primed with hostility via this task were more likely to interrupt a conversation between an experimenter and a confederate reliably faster than those exposed to politeness-related stimuli.

The impact of priming on behavior is prevalent in social-cultural literature these days. According to Jacobs, Grainger, and Ferrand (2015), priming may affect almost any type of individual behavior, altering it in a given context without the subject's knowledge. The primary research on priming (Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979) focused on how passive activation of certain traits in a given context can have a profound impact on how the subject perceived issues in other unrelated scenarios. An experimental study by Ven (2019) revealed that exposing individuals to words related to "kindness," as part of a class language activity, led them to later view the target person as more kind, as compared to the perception formed by the control group. Ven (2019) argues that the priming impact should be seen as a single concept or category that help seize the behavioral input in competition with each other, feeding requisite ideals, with this construal of input serving as the ground for later, deliberately made decisions.

Over the past decade, there has been a tremendous improvement in knowledge of different psychological concepts and processes that can be influenced through priming or put in motion unconsciously. Mattler (2006) identifies different social norms that can guide or direct personal behavior in various contexts, such as goals to enhance high performance, how to handle an opponent, or being just. In recent years, a nature more complex has been discovered, i.e., incognizant activation of deep cultural ideologies, which has brought to light in-depth information about behavioral traits that help us in understanding and often defining ambiguous

social behavior. For example, in a study by David and Ozaki (2010), they found through semantic priming they were able to activate the existence of colonial mentality- consistent cultural knowledge schemas in Filipino Americans. Colonial mentality refers to the notion that pleasantness or desirability are associated with any cultural values, objects, and behaviors that are American or Western (David and Ozaki, 2010). David and Ozaki (2010) conducted three studies to examine colonial mentality's covertness and automaticity, using a word fragment completion task, lexical decision priming task, and implicit association task. They found that colonial mentality had been deeply internalized, such that Filipino-related stimuli are automatically associated with ideas of unpleasantness, while American-related stimuli are automatically associated with ideas of pleasantness.

Factors that influence person perception include the social categorization of the target person who is being evaluated. Otten and Stapel (2006) suggest that the social categorization of a target person as either an in-group or out-group member affects the interpretation of their ambiguously aggressive behavior. Using a brief vignette which described the behaviors of an individual as ambiguously aggressive (Srull & Wyer, 1979), Otten and Stapel (2006) conducted three experiments among Dutch participants and found that judgements of an outgroup target (Moroccan, Surinamese, German) committing an ambiguously aggressive action were rated as more aggressive compared to the judgements made about an ingroup target (Dutch) and this effect depended on the extent to which the target group is associated with aggressiveness, if the target is a member of a stereotypically aggressive outgroup (i.e. Moroccans) then they were rated as more aggressive than members of a stereotypically non-aggressive outgroup (i.e. the Surinamese). They also found that by subtly priming an association between in-group and aggressive behavior, subsequent evaluations of the in-group target were influenced negatively.

Person exemplar priming, in which participants are primed with a specific person as a comparison standard to the target person, can also influence person perception. For example, Herr (1986) found that by priming participants with a person exemplar of hostility, such as Hitler, participants were more likely to subsequently rate Donald as kind. Participants primed with a person exemplar of kindness, such as Shirley Temple, were more likely to rate Donald as hostile (Herr, 1986). This study found that person exemplar priming leads to contrast effects in judgements of unambiguous targets. Participants primed with person exemplars of moderate categories (moderately hostile or moderately kind) evaluated a target person behaving ambiguously consistent with those categories. Participants primed with person exemplars of extreme categories (extremely hostile or extremely kind) evaluated the same target person in the opposite direction from the activated category. These results extended the findings of Herr, Sherman, and Fazio (1983), where subjects were primed with exemplars of different levels of animal size prior to making judgements about ambiguously sized animals and found that those that were primed with exemplars of extreme categories before making judgements of an ambiguously sized animal resulted in contrast effects.

Previous research has shown that different personality traits or motivations can be primed. In a widely known research study, Srull and Wyer (1979) found that exposing participants to stimuli related to hostility led to more hostile interpretations of the actions of an individual described in a brief vignette. In a study by Chartrand and Bargh (1996), it was found that by using words such as “strive” and “achieve,” researchers were able to prime motivation by activating the goal of achievement in participants. It is also possible to prime cooperation or competition on the construal of and decisions made in a Prisoner’s Dilemma game (Kay & Ross, 2003). Kay and Ross (2003) primed participants using the sentence unscrambling task with

words associated with competition or cooperation, making thoughts about competitiveness or cooperation more cognitively available. Afterwards, they were asked to judge an ambiguous situation (a Prisoner's Dilemma game). Those who were primed with competitiveness were more likely to defect, to rebel and those who were primed with cooperativeness were more likely to cooperate (remain silent) – participants were more prone to behave in a manner coherent with their previous judgement.

It is possible to prime individuals using different types of stimuli. Higgins, Rholes, and Jones (1977) used a color priming task to show impression formation and category activation. Participants were first shown 10 slides containing different words (i.e. "sky," "tree," "yellow") on different colored backgrounds and were instructed to quickly name the background color after presentation of the slide. Before each slide, participants were presented with a memory word drawn from a list that included 6 object-nouns (e.g., "corner" or "furniture") and 4 personality trait terms (two of which were positive and two of which were negative— e.g., "adventurous," "self-confident," "reckless," or "conceited") that they had to repeat immediately after naming the background color of the slide. Afterwards, participants were asked to read a vignette about an individual named Donald and were then given two questionnaires, one that asked participants to characterize Donald's personality and another that asked factual questions about the Donald vignette. Subjects were then asked to return between 10 and 14 days later to fill out a questionnaire and to probe for suspicion. Results from this study found that participants who memorized positive words formed positive impressions of Donald, while participants who memorized negative words formed negative impressions of Donald.

Bargh, Chen, and Burrows (1996) found that participants (41 non-African-American undergraduate students) primed with photographs of African American faces behaved in a more

hostile manner compared to those primed with Caucasian faces. Participants were asked to sit at a computer and complete a visual task. Before each trial, a photo of either a young African American or Caucasian face appeared subliminally. On the 130th trial, an error message appeared on the screen informing the participant that their data were not saved and that they would have to redo the visual task. Participants' facial reactions to this message were recorded via a hidden video camera in the lab room. After it was revealed to participants that the computer did actually save their data, they were asked to complete the Racial Ambivalence Scale (Katz & Hass, 1988) and the Modern Racism Scale (McConahay, 1986). Results indicated that participants who were primed with photographs of African American faces showed a more hostile expression than those primed with Caucasian photographs. The possibility that the results found were influenced by attitudes toward African Americans was examined by calculating the correlation between average hostility ratings and participants' level of racism measured by the two racism scales. Results indicated that participants that had low racist attitudes toward African Americans were just as likely to behave in a hostile manner as participants that had high racist attitude, irrespective of their priming condition.

As mentioned above, Srull and Wyer (1979) conducted a noteworthy study showing evidence of a hostile priming effect in participants. Two experiments were conducted in their study, one priming hostility and the other priming kindness. As a pilot test for each experiment, researchers had 43 participants not part of the main experiment rate a vast amount of individual behaviors along a scale from 0 ("not at all hostile") to 10 ("extremely hostile"). After analyses of the data, five behaviors were selected that represented high hostility, five that represented low hostility, and ten ambiguous behaviors which were split into two groups of five behaviors and then used to construct a vignette describing a hypothetical target person. This provided two

stimulus replications of target information. Additionally, all twenty behaviors were used as test items in the main experiment where participants had to rate how hostile/kind they considered each behavior to be.

The procedure for the main part of the two experiments was similar, except where noted. In each experiment, participants first completed a sentence unscrambling task, where they descrambled a total of either 30 phrases or 60 phrases, and each phrase consisted of four words. The participant's task was to underline three of the words that would make a complete sentence as quickly as possible. In the hostile (Experiment 1) or kindness (Experiment 2) experimental condition, 80% of the descrambled sentences contained words related to the priming concept, and the remaining descrambled sentences were not related to the priming concept. Examples of the priming sentences included "l*eg break his arm*" (Experiment 1) and "*the hug boy kiss*" (Experiment 2); an example of the neutral sentences that were used included "*her found knew I*" (Srull & Wyer, 1979). In both experiments contained a neutral condition, where only 20% of the sentences contained words related to the priming concept. After completing the sentence unscrambling task, participants were either told to continue onto the next task, return in one-hour or return twenty-four hours later. When participants returned, they were asked to read one of two different versions of a brief vignette about a man named Donald who behaved ambiguously hostile (Experiment 1) or kind (Experiment 2) and rate him on 12 traits along a scale from 0 ("not at all hostile/kind") to 10 ("extremely hostile/kind"). Participants were then asked to rate the hostility (Experiment 1) or kindness (Experiment 2) represented by the 20 behaviors that were selected from the pilot study described earlier, on a scale from 0 ("not at all hostile/kind") to 10 ("extremely hostile/kind"). Lastly, participants were then asked to rate the likelihood of trait co-occurrence, which is simply the likelihood that two traits can co-occur, such as "*If a*

person is hostile/kind, how likely is it that he is ____(mean/friendly)" and rate this likelihood on a scale from 0 ("not at all") to 10 ("extremely").

Deviating from Experiment 1, in Experiment 2, after completion of the three rating tasks, participants were asked to fill out a post-experimental questionnaire, indicating which of the four experimental tasks (sentence unscrambling task, rating of Donald, rating the kindness of 20 individual behaviors, estimation of the co-occurrence of kindness with 11 other traits) were most likely related to each other, if they suspected a relationship of the priming task and impression formation task. This served to study the extent to which participants were aware of the objectives of the experiment.

Preliminary analysis of the 12 traits, found that six of these traits (hostile, unfriendly, dislikable, kind, considerate, and thoughtful) implied a high or a low degree of hostility/kindness, whereas the other six traits were unrelated to either hostility or kindness. Researchers averaged the scores of the six traits that demonstrated a high/low degree of hostility/kindness (after appropriate reverse coding) to provide a single composite hostility rating. Results from the study indicated that priming participants with hostility/kindness-related stimuli caused subsequent information to be encoded using the same trait concept. The likelihood of this effect to occur increased with the number of times the trait was activated previously; participants who unscrambled 60 sentences compared to 30 sentences displayed a stronger, but not significant effect. Srull and Wyer (1979) found that, once participants were primed with hostility or kindness, their subsequent judgements on the same trait concept would increase. It was also found that the effect of the prime decreased over time, the accessibility of the trait category (hostility and kindness) and its effect on the interpretation of subsequent information decreased over time (no time delay, 1 hour, and 24 hours). Furthermore, the accessibility of the

trait category kindness decreased more rapidly over time than did the accessibility of hostility. Lastly, Srull and Wyer (1979) found that participants were not aware of the relatedness of the priming and experimental tasks; only 1 out of 96 participants believed the sentence unscrambling task was related to any of the subsequent experimental tasks (rating of an individual presented in the vignette, ratings of individual behaviors, and ratings of trait co-occurrence).

Components of Study	**Srull and Wyer (1979)**	**Srull and Wyer (1980)**
Participants	96 Total participants (8 in each cell)	96 Total participants (8 in each cell)
Pilot Study	To select behavioral descriptions to be used in the main study, subjects who did not participate in the main experiment rated a list of individual behaviors along a scale. Based on average scores, 5 were selected that conveyed high hostility, 5 were selected that conveyed low hostility, and 10 were selected that were considered "ambiguous" according to criteria set by the researchers.	Followed same process as Srull and Wyer (1979)
Sentence Unscrambling Task	Four-word phrases Total: 30 or 60 items (Experimental group: 80% trait-related phrases and 20% neutral phrases) (Control group: 20% trait-related phrases and 80% neutral phrases)	Four-word phrases Total: 50 items (Experimental group: 70% trait-related phrases and 30% neutral phrases) (Control group: 30%

		trait-related phrases and 70% neutral phrases)
Vignette	Donald Vignette	Donald Vignette
Judgement Tasks	After reading the vignette, participants rated Donald along 12 separate trait dimensions. Then participants rated the hostility (Experiment 1) or kindness (Experiment 2) conveyed by each of the 20 behaviors selected from the pilot study.	Followed same process as Srull and Wyer (1979) – therefore different trait dimensions compared to original study.
Presentation of delay	All experienced delay after the sentence unscrambling task.	Half experienced delay between sentence unscrambling and presentation of stimulus information. Other half experienced delay after reading the vignette and before presentation of rating tasks.
Length of delay	None, 1 hour, or 24 hours	None, 24 hours, or 1 week

Subsequent studies (Srull & Wyer, 1980; Bargh & Pietromonaco, 1982; Carver, Ganellen, Froming & Chambers, 1983) have been conducted to extend the results of the original

Srull and Wyer (1979) study. In Srull and Wyer (1980), three experiments were conducted to determine if the extent of the priming effects vary with different time intervals between the priming task and the vignette, and between the vignette and the rating task (Experiment 1 and Experiment 2), while Experiment 3 sought to determine if priming effects are due to the accessibility of categories at the time of encoding or in this case at the time of the presentation of the vignette.

Experiments 1 and 2 were similar to the original study, where Experiment 1 primed hostility, Experiment 2 primed kindness, but had some differences in design and materials. In Experiments 1 and 2, the vignettes were the same as those used in the original study, but as done in the original study, a pilot study was conducted for Experiments 1 and 2 (1980) to determine 20 behavioral descriptions of hostility (Experiment 1) or kindness (Experiment 2) to be used as test items. Srull and Wyer (1980) removed the length manipulation (30 phrases or 60 phrases) in the sentence unscrambling task used in the original study, and, as a substitute, used 50 items with either 70% or 30% hostility-related priming items (Experiment 1) or kindness-related priming items (Experiment 2) in sentences rather than the 80%/20% split used in the original study. In the original study, the researchers placed a time delay between the sentence unscrambling task and the presentation of stimulus information (i.e., the reading of the vignette and rating tasks). Srull and Wyer (1980) examined the same delay in half of their participants; the other half experienced a delay between reading the vignette and the rating tasks. This was done to further examine the manner in which the magnitude of priming effects varies with various time intervals. The length of the time delay was also longer in the 1980 study; participants were either presented with the stimulus information or the rating tasks after either no delay, 24 hours, or 1 week.

Experiment 3 was similar in design to Experiment 1, examining the priming of hostility, except that only two delay intervals (no delay vs. 24 hours) were examined. Also, the vignette was presented before the sentence unscrambling task; this was done to examine if a priming effect is evident when a trait category is activated after information is first interpreted or encoded into memory by a subject. By having participants read the vignette first, encoding into memory information about the person in the vignette, and then priming a trait concept, followed by the rating task, Srull and Wyer (1980) were determining if priming effects still occur after initial encoding and if these effects are due to recoding information during the rating task. Participants first read the Donald vignette, then either immediately afterwards or following a 24-hour delay, were presented the priming stimuli, and then asked to rate Donald along the same trait dimensions as those used in Experiment 1.

The three experiments were conducted to determine if the extent of the priming effects vary with different time intervals between the priming task and the vignette, and between the vignette and the rating task (Experiment 1 and Experiment 2), while Experiment 3 sought to determine if priming effects are due to the accessibility of categories at the time of encoding or, at the time of the presentation of the vignette.

First, participants were presented with the sentence unscrambling task, in which 70% of the phrases connoted hostility (Experiment 1) or kindness (Experiment 2) in the experimental condition and 30% of the phrases were not related to the priming concept. The control group was presented with sentences in which 30% of the phrases connoted hostility (Experiment 1) or kindness (Experiment 2) and 70% of phrases were not related to the priming concept. Half of the participants then experienced a time delay (no delay, 24 hours, or 1 week) and were then asked to read the Donald vignette, rate him along 12 trait dimensions, and rate the hostility

(Experiment 1) or kindness (Experiment 2) represented by each of the 20 behavioral descriptions that were chosen from the pilot study. The other half of participants immediately read the Donald vignette after the sentence unscrambling task and then experienced a corresponding delay (no delay, 24 hours, or 1 week) before being presented with the rating tasks.

In Experiments 1 and 2, Srull and Wyer (1980) found that, the more a participant was primed (number of trait-related items), the more likely the judgement ratings of Donald matched the target prime (hostility or kindness), but when the delay was placed between the priming and the rating of Donald, the effect declined over time, a result similar to the original study. Conversely, there was no decline in ratings across time when the delay was placed between reading the vignette and the rating of Donald. Lastly, their third experiment found that, in order to see an effect, the priming needs to be placed before reading the vignette.

Another clear extension of the Srull and Wyer (1979) study was conducted by Bargh and Pietromonaco (1982), using the Donald vignette and the trait measures from the original study. Bargh and Pietromonaco (1982) were interested in determining whether categories can be activated outside awareness and still influence participants' judgements of Donald. Instead of using the sentence unscrambling task, researchers presented one of three lists of 100 words parafoveally for 100 ms, containing either 0, 20, or 80 hostile words, taken from the original Srull and Wyer (1979) study. These words were flashed on a CRT display at one of four locations around a fixation point, and participants were asked to indicate on which side of the screen the flash occurred. Researchers found that participants were unaware of the words in the flashes. Participants were then asked to read the Donald vignette and rate Donald on 12 traits. Results indicated that participants' impressions of Donald were directly related to the amount of

hostile words to which they were exposed, supporting the idea that trait categories can be primed outside of conscious awareness and influence ratings.

In a study conducted by Carver, Ganellen, Froming, and Chambers (1983), they extended the original study by using the Donald vignette, trait ratings, and the sentence unscrambling task in two separate experiments. In Experiment 1, researchers were looking to see if watching hostile behavior on a videotape would influence trait ratings of Donald. In Experiment 2, researchers were looking to see if unscrambling hostile sentences had an effect on subsequent behavior, teaching a task to another person (confederate) and administering electric shocks when they answered incorrectly (an experiment similar to Milgram's obedience to authority experiment). Carver et al. (1983) found a significant effect in both experiments. Watching hostile behavior caused participants to rate Donald as a more hostile person than did the control group, and unscrambling hostile sentences caused participants to administer electric shocks of greater intensity to another person (i.e., the confederate).

Further studies have found the sentence unscrambling task to be effective in priming hostility/kindness (Srull & Wyer, 1980; Carver et al., 1983). The sentence unscrambling task from the original study was used to prime hostility/kindness in a subsequent study by Bargh, Chen, and Burrows (1996). They looked to extend the priming effect to rudeness/politeness. Researchers used negative priming stimuli such as *bold, bother, annoyingly,* and *disturb* in their sentence unscrambling task and found that those participants who saw these words were more likely to interrupt the experimenter more frequently and quickly than those primed with politeness-related stimuli such as *courteous, honor, behaved,* and *cordially* (Bargh, Chen, & Burrows, 1996).

In another study, Higgins, Bargh, and Lombardi (1985) extended the findings of the original study by using the sentence unscrambling task to extend the priming effect from hostility/kindness into other domains to increase generalizability (independent/aloof, adventurous/reckless, and persistent/stubborn). Higgins et al. (1985) modified the sentence unscrambling task such that words were presented on the screen, and the participant was asked to say the sentence aloud. Examples of this task included a group of four words, "*you unconventional are quiet*" which appeared on the screen for three seconds. Participants would use three of the words to create a grammatically correct sentence, such as "*you are unconventional*" or "*you are quiet*," and stated it out loud. After the sentence unscrambling task, participants were asked to count backwards by three for either 15-seconds or 120-seconds and then to complete other tasks. Higgins et al. (1985) found that the frequency effect (the more frequently a construct is primed, the higher is its level of activation) decreased as the delay between priming and stimulus presentation increased, and that recently primed constructs predominate in the short term.

Recent studies have shown that priming has an effect on trait-related cognition (involving misrecognition of trait-related information leading to faulty impressions of the material presented) rather than on person perception. Buchanan (2015) conducted a study using Facebook advertisements to prime aggression and their effect on rating a person's Facebook profile as aggressive, levels of aggressive mood (measured by POMS), and a word recognition task. Participants were randomly assigned to be in the aggressive condition or neutral condition and were asked to study a simulated Facebook page for a person named "Stephen Green," with four advertisements, which either displayed aggressive advertisements or neutral advertisements. Participants then completed a mood rating questionnaire (POMS subscales), a word recognition

task (to measure aggressive cognition), and a rating task, where they rated how aggressive they believed “Stephen Green” was from his Facebook profile, similar to Srull and Wyer (1979). The results of their study found that aggressive cognitions were more accessible for those in the aggressive condition than in the neutral condition but found no effects of priming on person perception or mood. Discrepancies such as these are why we are interested in studying the effects of priming on person perception.

Registered Replication Report (RRR, McCarthy et al., 2018)

In 2018, McCarthy et al. (2018) were interested in replicating the results of Experiment 1 of the original Srull and Wyer (1979) study due to failures to replicate priming effects in past studies (Cheung et al., 2016; Doyen, Klein, Pichon, & Cleeremans, 2012). Using 26 independent replications (N = 7,373), McCarthy et al. (2018) found that the methods used based on the Srull and Wyer (1979) study did not consistently produce a significant priming effect.

While the procedure was similar to Experiment 1 of the original study, McCarthy et al. (2018) modified some of the materials, although they were made in consultation with one of the original researchers, Wyer. The reason for the use of different stimuli in this replication was that the stimuli for the original sentence unscrambling task could not be found. Wyer assisted in creating stimuli that were consistent with the original and had it piloted. In the original study, one of the experimental tasks included rating a list of behaviors individually. In the RRR, Wyer and researchers modified the pronouns to be gender neutral. Some wording was changed as well (“slamming down a handset” = “slamming down a phone”), to account for current young adults being unaware with the actions described in the original study. Lastly, the name Donald from the vignette was changed to Ronald to avoid any association with current President Donald Trump.

Other differences between the Registered Replication Report and the original study included procedural modifications. Researchers focused their study on comparing two conditions from the original study that demonstrated a strong and clear priming effect: the immediate-testing condition (no time delay between the completion of the priming task and the presentation of the target information) using 30 items in the sentence unscrambling task. For the hostile priming condition 24/30 sentences involved hostile words, and for the neutral condition 6/30 sentences involved hostile words, similar to the original Srull and Wyer (1979) study. Rather than inform participants that they would be participating in two unrelated studies, the RRR had participants complete the sentence unscrambling task and the rating tasks all together in a large classroom setting. The RRR was completed in a classroom setting was because another RRR (Verschuere et al., 2018) was collecting data at the same time and required participants to be run in a classroom setting.

The reasoning behind conducting the current research study is based on criticisms of the 2018 RRR. The RRR deviated from the original Srull and Wyer (1979) study and these deviations raised some uncertainty about whether or not the RRR demonstrated a clear hostile priming effect. Most of the deviations are considered to have had a trivial effect on the results, such as the change of the list of behaviors to be gender neutral, and the words used for the sentence unscrambling task. The most notable deviation is the setting in which the study occurred. The RRR had participants complete the sentence unscrambling task and the rating tasks altogether in a large classroom setting, whereas in the original study, participants were tested in a distraction-free lab setting in groups of 4-8 (Srull & Wyer, 1979) which may have had an effect on results.

Current Study

In our current study, we were interested in conducting a close replication of the original Srull and Wyer (1979) study that addressed criticisms of the 2018 RRR due to its deviations from the original study. Similar to the 2018 RRR, in the current study participants were presented with the brief vignette and rating task immediately following the sentence unscrambling task (immediate testing condition) and the use of 30 items in the sentence unscrambling task, two conditions that demonstrated a clear priming effect in earlier studies.

We were interested in testing the existence of the hostile priming effect by improving conditions that would be the most favorable in demonstrating a priming effect including participants physical environment during a study, which has been found to have an effect on automatic cognition and behavior. Having participants sit in an isolated setting may provide a more favorable context to study priming effects where fight-related action semantics in response to primes are greater than when participants are in an open field setting (Cesario, Plaks, Hagiwara, Navarrete, & Higgins, 2010). Cesario et al. (2010) explained that nonhuman animals engage in defensive behavior such that when there is not an opportunity for escape this leads to fight-related semantics and prevents distancing behavior, whereas in opportunities where escape is available, this leads to flight-related semantics and allows for distancing behavior. Cesario et al. (2010) were interested in applying this methodology to priming participants with photos of African American males versus Caucasian males in an enclosed booth versus an open field to examine if they would engage in fight/flight-related semantics. Results from their study found that participants primed in an enclosed booth showed increased fight-related semantics compared to those primed in an open field, showing that primes have a stronger effect in isolated settings.

In the current study, the sentence unscrambling task from the 2018 RRR was also used. However, there were slight departures in materials from the original 1979 study. Firstly, for the

hostile priming condition, 24/30 sentences involved hostile words, and for the neutral condition 0/30 sentences involved hostile words, which is a slight departure from the original 1979 study but was done to increase the chances of a priming effect. Second, rather than using two vignettes which served as stimulus replications in the original Srull and Wyer (1979) study, the current study used one--the Ronald vignette. As conducted in Experiment 2 of the original Srull and Wyer (1979) study, participants were also asked if they were aware of the influence of the prime. Awareness of the prime can have an effect on results as demonstrated by Loersch and Payne (2012), where participants were more likely to show evidence of priming effects when they were led to misattribute prime-related content to their own thoughts. Lastly, positive controls, which are additional experimental conditions that are run, in which the correct result is very well known and used to assess test validity providing some assurance that the experiment was conducted properly, were not included in the original Srull and Wyer (1979) study and the RRR. In the current study, positive controls were created by McCarthy such as responses to “Reading books is a hobby of mine” and “How many books have you read for pleasure in the last year?”, and gender differences in self-reported height will serve as positive controls. Positive controls help control extraneous variables that may have an effect on results, such as whether the participant was paying attention during the study. Lastly, the traits used for the trait rating task were not the same as those used in the 2018 RRR/original study. In the 2018 RRR and original study, the traits *hostile, unfriendly, dislikeable, kind, considerate,* and *thoughtful* were used to form an index of the extent to which Donald/Ronald was perceived as hostile. However, in the current study the traits *hostile, smart, angry, honest, unfriendly, outgoing,* and *dislikable* were used.

By integrating the above features into the present study, we were testing to see if the current study would obtain the same results as the original Srull and Wyer (1979) study, that

participants who were primed with hostility were more likely to interpret the ambiguous behaviors of an individual as hostile, compared with participants in a comparison condition involving a neutral prime. The current study is part of a multi-site collaborative study, where each individual lab conducts a close and conceptual replication of the original Srull and Wyer (1979) study, where the conceptual replication involved researchers developing and using stimuli unique their participant pool. In the current paper, I will be focusing on the close replication of the original study.

Method

Participants

Sixty-six undergraduate students from an urban college in New York City participated in this study. Students participated in this research study as part of a requirement for a class. Nineteen participants were excluded from the study for meeting one or more exclusion criteria including not completing the sentence unscrambling task, not providing ratings for each of the traits, or failing either one of the two attention checks, leaving the total number of eligible participants at forty-seven**.** Participants included 36 women aged 18 to 52 years old (M = 20.52, SD = 6.97), 10 men aged 18 to 28 years old (M = 23.5, SD = 12.36), and one individual who did not disclose their sex. In the hostile condition, there were 23 participants (9 male and 14 female) aged 18-54 (M = 20.75, SD = 7.92) and in the control condition, there were 23 participants (22 female and one male) aged 18-52 (M = 21.43, SD = 8.62).

Materials

The questionnaire was administered via the college's Qualtrics System website at a computer in the college's PsychologyLab. Online consent was used, which contained information about the purpose of the study, procedures, potential risks/benefits, participants'

rights, and contact information of the Principal Investigator and the Research Compliance Administrator.

The sentence unscrambling task was adapted from the RRR (2018) study. The stimuli for the sentence unscrambling task from the Srull and Wyer (1979) study were unavailable, therefore new stimuli were created and pretested (see https://osf.io/32pkz/ for details on the pretesting) and subsequently used in the RRR (2018) study. The sentence unscrambling task used in the present study consisted of 30 trials with each trial containing a 4-word phrase in a scrambled order (i.e. "*milk pour the spill*") and participants were asked to create a 3-word sentence ("*spill the milk*" or "*pour the milk*"). Participants in the hostile group were asked to unscramble 24 (out of 30) sentences that described aggressive behaviors (i.e. "*shoot I'll you hurt*"). Participants in the control group were asked to unscramble sentences that described non-aggressive behaviors (i.e. "*your pen key use*").

Adapted from the original Srull and Wyer (1979) study, the Donald vignette (see Appendix A) was also used in the current study. As done in the RRR (2018), the name Donald was changed to Ronald due to any unwanted association with President Donald Trump. The rating task that followed the vignette was adapted from Srull and Wyer (1979). This task consisted of seven traits (*hostile, smart, angry, honest, unfriendly, outgoing,* and *dislikeable*), which participants were asked to use to rate Ronald using a 0 to 10 scale (0 = *not at all* to 10 = *extremely*). The ratings for traits *hostile, unfriendly,* and *dislikable* were averaged together to create an average hostility rating for each participant during analyses. This is a deviation from the original Srull and Wyer (1979) study, where six traits (*hostile, unfriendly, dislikable, kind, considerate,* and *thoughtful*) were averaged together (the latter three traits were reverse-scored) to create an average hostility rating.

Randy McCarthy (2018) generated the stimuli used for the attention check and positive control. Positive controls are additional experimental conditions that are run, in which the correct result is very well known and used to assess test validity providing some assurance that the experiment was conducted properly. Participants were presented with a paragraph that instructed participants to answer "*Completely Disagree*" to the first two statements, "*Watching TV is a hobby of mine"* and "*Playing video games is a hobby of mine*" and to answer honestly to the next two statements, "*Reading books is a hobby of mine*" and "*How many books have you read for pleasure in the past year?*" The responses to the first two items served as an attention check while a positive correlation between the latter two items served as a positive control, in which responses to these items are known. Therefore, we should expect to find that those participants that agree that reading books is a hobby of theirs, should have read more books for pleasure this year compared to those who disagree with the statement.

The suspicion probe (see Appendix B) was modeled after the debriefing example in Bargh and Chartrand (2000), which served to probe in a systematic way for any suspicions the participant has about the expected effect of the prime on their subsequent performance on a task in the experiment. Three questions were adapted from the debriefing example: "*What do you think the purpose of the study was?*", "*Do you think any of the tasks in this study were related?*", and "*To what extent do you feel like the 'scrambled sentence task' influenced your ratings of Ronald?*" The first question was open-ended, the second question was a forced choice of either yes/no, and the third question used a 0 to 7 rating scale (0 = *Did not influence at all*, 7 = *Influenced a lot*).

Procedure

As part of the psychology research requirement, students decided what study to participate in once they logged into the SONA website used to manage research assignments. Participants then chose a timeslot to come to the research lab, where they were tested individually at a computer via the college's Qualtrics system website. After reading and signing the consent form via the Qualtrics system, participants were randomly assigned to either the hostile priming condition or neutral condition of [either the close replication or conceptual replication group].

Participants were first asked to complete a 30-trial sentence descrambling task; participants in the hostile priming condition were asked to descramble 24/30 sentences that contained hostile words and 6/30 sentences contained neutral words. For example, participants in this condition were shown a 4-word set "*face his kick leg*," and were asked to descramble this set into a 3-word phrase that made a complete sentence such as, "*kick his face*" or "*kick his leg.*" Participants in the neutral priming condition were asked to descramble zero sentences that formed hostile sentences. For example, participants in this condition were shown a 4-word set "*story poem a write*" and were asked to descramble this set into a 3-word phrase, "*write a poem*" or "*write a story*." Participants were then asked to read the Ronald vignette (see Appendix A) and were then asked to rate Ronald using a 0 to 10 scale for traits *hostile, smart, angry, honest, unfriendly, outgoing,* and *dislikeable* (0 = *not at all* to 10 = *extremely*).

Participants then viewed a screen that asked them to report the extent to which they agreed with the statements, "playing video games is a hobby of mine," "watching TV is a hobby of mine," and "reading books is a hobby of mine" on a 1 to 7 scale (1 = *completely disagree* to 7 = *completely agree*). They were then asked to self-report how many books in the past year they have read for pleasure. At the top of the screen read instructions that informed participants

how to respond to the three statements, and one question, where the responses to the first two statements served as an attention check and a positive correlation between the last two items served as a positive control. Participants were then asked to fill out demographic information including their age, gender, and their height. Gender and height were used as an additional positive control since it has been found that on average, men are taller than women (Gustafsson & Lindenfors, 2006), therefore we should expect to see this same trend in self-reported measures of height and gender. Lastly, after completing the study, participants were asked questions to probe for suspicion.

Results

To demonstrate evidence of a hostile priming effect as was shown in Srull and Wyer (1979), first, average hostility ratings were computed from the three traits (*hostile, unfriendly,* and *dislikable*) for the hostile group and the control group. Then, using these average hostility ratings, hypothesis tests were conducted to determine if the difference in average hostility ratings were significantly different between experimental conditions and if gender had influenced these ratings. Lastly, positive controls were evaluated to assess for test validity.

In the hostile group, the average hostility rating was 8.43 (SD = 1.14) and in the control group, the average hostility rating was 8.26 (SD = 2.21). In the hostile group there were 14 female (M = 8.10, SD = 2.46) and 9 male (M = 8.52, SD = 1.86) participants. In the control group there were 22 female (M = 8.62, SD = 0.96) participants, one male (M = 5.67) participant, and one person who did not identify their sex (M = 7).

An independent samples t-test was conducted between the experimental group and control group with the average hostility score being the dependent variable to determine if the sentence unscrambling task led to a hostile priming effect. Results indicated that there was not a

significant difference between the experimental group (M = 8.26, SD = 2.21) and control group (M = 8.43, SD = 1.14) scores of average hostility, $t(32.56) = .33$, $p > .05$, as shown in Figure 1. In other words, the sentence unscrambling task did not cause significantly higher hostility ratings in the experimental group versus the control group, therefore showing no evidence of a hostile priming effect as demonstrated in Srull and Wyer (1979).

Insert Figure 1 here

A two-way analysis of variance (ANOVA) was conducted to determine if participants' gender (male/female) and experimental condition (experimental/control) had a significant effect on average hostility scores, however no significant main effects or interactions were found. This led us to examine if ratings for the specific traits *hostile, unfriendly, and dislikable* differed individually, based upon gender and experimental condition since these three traits were used to calculate average hostility scores for each participant. Two-way ANOVA's were conducted and found that the traits *hostile* and *unfriendly* did not yield any significant main effects or interactions. However, a significant main effect and interaction were found for the trait *dislikable*. Results indicated a main effect for gender was significant, $F(1,42) = 4.21$, $p < .05$. As shown in Figure 2, the average rating for the trait *dislikable* among male participants (M = 8.90, SD = 2.03) differed significantly from the average rating for the trait *dislikable* among female participants (M = 9.28, SD = 1.85). Specifically, female participants rated Ronald *dislikable* more negatively compared to male participants.

Insert Figure 2 here

There was a significant interaction between the gender of the participant and the experimental condition the participant was in, on the rating of the trait *dislikable*, F (1,42) = 8.06, $p < .05$. Trait ratings for *dislikable* depended on whether the participant was male or female and

whether the participant was in the experimental or control group. To interpret this interaction and determine where there were significant differences, we examined the cell means as shown in Table 1. The male participants in the control condition had a lower average trait rating for *dislikable* (M = 5.00; 95% CI = 1.48, 8.52) than did female participants in the control condition (M = 9.73; 95% CI = 8.98, 10.48). Because there was no overlap in CI's between male and female participants in the control condition, we can conclude that this difference is significant. This shows evidence that participants had significantly rated Ronald more negatively for the trait *dislikable* if they were female than if they were male in the control group, as shown in Figure 3.

Insert Figure 3 here

To assess for test validity, two sets of positive controls were used in the current study, which assessed test validity where the correct result is very well known. A Pearson product-moment correlation coefficient was computed to assess the relationship between the items "Reading books is a hobby of mine" and "How many books have you read for pleasure in the past year?" As expected, we found a significant positive correlation between the two responses, r = .57, n = 47, p < .05. Participants that were more likely to agree that reading books is a hobby were more likely to have read more books compared to participants who disagreed with the statement. These results indicate that the primary result we found of no presence of a hostile priming effect is not due to participants reading instructions carefully blindly answering rating Ronald on individual traits. For the second positive control, an independent samples t-test was performed to test whether there were significant differences in self-reported height by gender. Results indicated that there was no significant difference between male (M = 67.9, SD = 3.51) and female (M = 65.44, SD = 8.46) self-reported height (in inches), $t(44) = .89$, $p > .05$. We

expected to find a significant effect in both positive controls, however failing to find a significant difference between self-reported heights of male and female participants indicates that there may have been an error in test administration, participants not paying attention to the questionnaire. These results were not expected due to past research supporting that men are generally taller than women (Gustafsson & Lindenfors, 2006), but our findings could be due to the sample size where we had a total of 36 females and 10 males in the study.

To provide additional context for the results in our sample, we re-analyzed data from three other samples in this multi-site project. These samples were also part of the multi-site study being conducted by Randy McCarthy so therefore they all have the same variables. Samples were selected from the following locations: Athens, Ohio, Redmond, Washington, and Richmond, Virginia. Ratings for the traits *unfriendly, hostile,* and *dislikable* were averaged into a new variable for each analysis.

Ohio Sample

First, we evaluated the results of an urban university at Ohio. Participants included 59 undergraduate students, however 11 were excluded due to failing one or two of the attention checks, leaving the total number of participants at 48. Participants included 13 men aged 18 to 22 years old (M = 19.46, SD = 1.13) and 35 women aged 18 to 23 years old (M = 19.18, SD = 1.09). The experimental group included 9 males and 17 females, while the control group included 4 males and 18 females. An independent samples t-test was conducted between the experimental and control group with the average hostility score being the dependent variable. Results indicated that there was not a significant difference between the experimental group (M = 8.38, SD = 2.01) and the control group (M = 8.89, SD = 1.62) scores of average hostility $t(46) = .96$, $p > .05$, as shown in Figure 4. The sentence unscrambling task did not cause a significant difference in hostility ratings of the target person Ronald in individuals primed with

hostile sentences versus neutral sentences, results identical to those obtained in the Hunter College sample. Interestingly, the control group in both samples had slightly higher average hostility ratings.

Insert Figure 4 here

A two-way ANOVA was conducted as well to determine if gender (male/female) and/or experimental condition (experimental/control) had a significant effect on hostility ratings. Male participants in the hostile group (M = 8.44, SD = 1.67) and control group (M = 8.83, SD = 1.29) did not have significantly different hostility ratings from female participants in the hostile group (M = 8.35, SD = 2.21) and control group (M = 8.91, SD = 1.72). Additional two-way ANOVA's were conducted analyzing if gender and/or experimental condition had an effect on the traits *hostile, unfriendly,* and *dislikable* as done in analyses for the Hunter sample. No significant effects were found among the three traits in the Ohio sample, whereas the trait *dislikable* had a significant effect in the Hunter sample.

As was done in the current study, two positive controls were analyzed in the Ohio sample. The first positive control assessed the relationship between items "Reading books is a hobby of mine" and "How many books have you read for pleasure in the past year?" Using a Pearson product correlation coefficient, results indicated a significant effect, r = .58, n = 48, p < .05. Participants that were more likely to agree with the statement read more books than those that did not agree with the statement. These results are consistent with the findings found in the current study. Secondly, gender differences in self-reported height was analyzed among the Ohio sample using an independent samples t-test. Results indicated that there was a significant difference between male (M = 73.85, SD = 4.96) and female (M = 64.63, SD = 4.09) self-reported height (in inches), $t(46) = 6.55$, $p < .05$. This finding is consistent with past research

(Gustafsson & Lindenfors, 2006), but was not also found in the Hunter sample.

Washington Sample

The next sample evaluated was the results of a sample from Washington. Participants included 67 undergraduate students, however 17 were excluded for failing to pass one or both of the attention check items, leaving the total number of participants at 50. Participants included 14 men aged 18 to 29 years old (M = 20.57, SD = 2.77), 34 women aged 18 to 32 years old (M = 21.21, SD = 3.49), and two participants who chose not to indicate their sex or age. The experimental group included 7 males and 26 females, and the control group included 11 males and 21 females. An independent samples t-test was used to determine if the sentence unscrambling task primed participants' judgements of an ambiguously behaving individual. No significant effect was found. Participants primed with hostile-related stimuli (M = 8.44, SD = 1.25) did not rate the individual to be more hostile than participants primed with neutral-related stimuli (M = 7.78, SD = 1.88), $t(48) = -1.45$, $p > .05$, as shown in Figure 5. This result is identical to those found in the Hunter sample. In comparison to the Hunter sample, the control group had a slightly lower average hostility rating for Ronald compared to the experimental group.

Insert Figure 5 here

A two-way ANOVA was conducted to determine if gender and/or group had a significant effect on hostility ratings. Male participants in the hostile group (M = 8.27, SD = 1.30) and control group (M = 7.70, SD = 1.18) did not have significantly different hostility ratings from female participants in the hostile group (M = 8.54, SD = 1.29) and control group (M = 7.67, SD = 2.17). Results indicated no significant main effects or interaction effects; in the Washington sample, participants' gender and experimental condition had no effect on hostility ratings of an

ambiguously behaving individual. This result is identical to the results from the Hunter sample. Additional two-way ANOVAs were conducted analyzing if gender and/or experimental condition had an effect on the traits *hostile, unfriendly,* and *dislikable* as was evaluated in the current study. No significant effects were found among the three traits in the Washington sample, whereas the trait *dislikable* had a significant effect in the current study.

Lastly, two sets of positive controls were analyzed in the Washington Sample. The first positive control assessed the relationship between the items "Reading books is a hobby of mine" and "How many books have you read for pleasure in the past year?" Using a Pearson product correlation coefficient, results indicated a significant effect, $r = .55$, $n = 50$, $p < .05$. Participants that were more likely to agree with the statement read more books than did those that did not agree with the statement. These results are consistent with the findings of the current study, which helps ensure the validity of the questionnaire given. Secondly, gender differences in self-reported height was analyzed among the Washington sample using an independent samples *t*-test. Results indicated that there was a significant difference between male ($M = 60.14$, $SD = 13.25$) and female ($M = 70.29$, $SD = 3.68$) self-reported height (in inches), $t(13.83) = 2.82$, $p < .05$. This finding is consistent with past research (Gustafsson & Lindenfors, 2006). Finding the expected results in both positive controls indicates we can interpret the result of failing to detect a hostile priming effect in the sample, not due to auxiliary hypotheses such as errors in questionnaire administration, data collection, participants paying attention, etc.

Virginia Sample

The final external sample examined were the results of a sample from Virginia. Participants included 64 undergraduate students. However, 17 participants were excluded from

the analyses for failing one or both of the attention check items, leaving the total number of participants at 47. Participants included seven men aged 18 to 29 years old (M = 21, SD = 3.92) and 40 women aged 18 to 24 years old (M = 18.97, SD = 1.19). The experimental group included four men and 18 women, while the control group included three men and 22 women. An independent samples t-test was conducted to determine if the average hostile ratings were significantly different between the experimental and control group. Results indicated no significant difference between the experimental group (M = 8.64, SD = 2.09) and the control group (M = 8.17, SD = 2.12), $t(45) = -.75$, $p > .05$, as shown in Figure 6. The result from the Virginia sample is identical to the results found in the Hunter sample. In comparison to the Hunter study, the control group in the Virginia sample also had a slightly lower average hostility rating for Ronald compared to the experimental group. Collectively, all five samples examined did not have a significant effect; participants primed with hostile-related stimuli were not more likely to rate an individual behaving ambiguously as more hostile compared to participants primed with neutral-related stimuli.

Insert Figure 6 here

A two-way ANOVA was conducted to determine if gender and/or experimental condition had a significant effect on hostility ratings. Male participants in the hostile group (M = 7.50, SD = 0.79) and control group (M = 7.00, SD = 2.85) did not have significantly different hostility ratings from female participants in the hostile group (M = 8.89, SD = 2.22) and control group (M = 8.33, SD = 2.04). Results indicated no significant main or interaction effects; in the Virginia sample, participants' gender and group assignment had no effect on hostility ratings of an ambiguously behaving individual. Two-way ANOVA's were conducted and found that the traits *hostile* and *dislikable* did not yield any significant main effects or interactions. However, a

significant main effect was found for the trait *unfriendly*, $F(1,43) = 4.45$, $p < .05$. As shown in Figure 7, the average rating for Ronald for the trait *unfriendly* male participants ($M = 6.43$, $SD = 3.10$) differed significantly from female participants ($M = 8.80$, $SD = 2.57$). Specifically, female participants rated Ronald more severely for the trait *unfriendly* compared to male participants. While we did not find this same effect on the rating of Ronald for the trait *unfriendly* in the results for the Hunter sample, we did find the same effect of gender on the rating of Ronald, specifically that female participants rated Ronald for the trait *dislikable* more negatively compared to male participants. In both samples we did not find a main effect of group and/or gender on the trait *hostile.*

Insert Figure 7 here

Lastly, two sets of positive controls were analyzed in the Virginia sample. The first positive control item asked participants to respond honestly to items "Reading books is a hobby of mine" and "How many books have you read for pleasure in the past year?" Using a Pearson product correlation coefficient, results indicated a significant effect, $r = .54$, $n = 47$, $p < .05$. Participants that were more likely to agree with the statement read more books than those that did not agree with the statement. These results are consistent with the findings of the current study, which indicate that the result we found of no presence of a hostile priming effect is likely not due to participants reading instructions carefully blindly answering rating Ronald on individual traits. The second positive control item asked participants for their gender and height (in inches); an independent samples *t*-test was used to analyze gender differences in self-reported height among the Virginia sample. Results did not find a significant difference in self-reported height between male ($M = 56.43$, $SD = 25.16$) and female ($M = 58.48$, $SD = 15.86$) participants, $t(45) = -0.29$, $p > .05$. This finding is inconsistent with past research, where men

are typically taller than females (Gustafsson & Lindenfors, 2006). We expected to find a significant effect in both positive controls, however failing to find a significant difference between self-reported heights of male and female participants indicates that there may have been an error in test administration, participants not paying attention to the questionnaire. These possibilities can also be a possible explanation as to why a significant difference in hostility ratings were not found.

Discussion

The current study found that undergraduate students from an urban college in New York City when presented with hostile-related stimuli in the form of a sentence unscrambling task, did not show any effects on trait ratings of an individual named Ronald behaving ambiguously hostile. Participants who completed the sentence unscrambling task with 80% hostile primes rated Ronald to be 0.17 points more hostile than did participants who completed the sentence unscrambling task with 0% hostile primes (i.e., all neutral sentences), a difference that was not statistically significant. These results are not consistent with what was found in the original study conducted by Srull and Wyer (1979). Using conditions believed to demonstrate a clear hostile priming effect, the immediate-testing condition (no time delay between the completion of the priming task and the presentation of the target information) using 30 items in the sentence unscrambling task did not replicate the results obtained in the original Srull and Wyer (1979). Participants in the current study were also tested in a lab-setting as was done in the original study. Similar to the findings of samples from different geographical regions (Washington, Ohio, and Virginia) in the same multi-site study, hostile priming effects were not found.

The current study also found that gender (male/female) and/or experimental condition

(experimental/control) had no effect on average hostility scores among participants. Additional analyses were conducted to determine if gender and/or group had an individual effect on each of the negative traits *hostile, unfriendly,* and *dislikable.* No significant effects were found for the trait *hostile,* but a significant effect of gender on the trait *dislikable* was found as well as an interaction effect of gender and group on the same trait in the Hunter sample and a significant main effect of the trait *unfriendly* was found in the Virginia sample. It was found that female participants had a significantly higher rating for the trait *dislikable* compared to male participants and that female participants in the control group had a significantly higher trait rating compared to the male participants in the control group. These results could be due to a methodological limitation, the number of male participants in the study. While, in the Hunter sample, the experimental group had nine male participants, the control group had only one. This limitation was also present in the Ohio, Virginia, and Washington samples where less than half of the participants in the hostile and control groups were male. In the Virginia sample, it was found that for the trait *unfriendly,* female participants ($N = 40$) had significantly higher trait ratings compared to males ($N = 7$). Although gender was not examined in the original study and in the RRR gender was discussed, with approximately 75% of the total number of participants being female.

Additional analyses included two positive controls which assessed test validity providing some assurance that the experiment was conducted properly. Positive controls were not included in the original study and the 2018 RRR, but were included here to help with interpretability of results. A significant correlation was found for the items, "*Reading books is a hobby of mine*" and "*How many books have you read for pleasure in the past year?*", results we expected to find and implies that the inability to detect a hostile priming effect may not be due

to participants not paying attention, errors in administering the questionnaire, or data collection. A significant effect was not found for gender differences in self-reported height in the Hunter sample, results we did not expect to find. Significant gender differences in self-reported height was found in the Ohio and Washington sample, a significant effect was also not found in the Virginia sample. These nonsignificant effects are most likely due to the low number of male participants vs. female participants within the main study (36 female and 10 male) and the Virginia sample (40 female and 7 male). The Washington sample (34 female and 14 male) and Ohio sample (35 female and 13 male) found significant effects, which is most likely due to the higher number of male participants. Failing to detect an effect in the positive controls in which we expected to find a result may be helpful when interpreting the results of not finding a hostile priming effect, suggests there may have been an error in the procedure of the study. Positive controls have a correct result which is very well known, and it is reasonable to believe that failure to have detected significant gender differences in self-reported height may have been caused by aspects of the study procedure, such as participants not paying attention. As an example, three participants in the Hunter sample had listed their height as 100 inches (8 ft 3 in), 82 inches (6 ft 8 in), and 40 inches (3 ft 3 in.). These heights are extremes within the sample; upon removing these participants from the sample, significant effects were found, $t(43) = 1.69$, $p < .05$, with males being significantly taller than female participants. It is possible that these three participants were not reading the question clearly and mis-entered their heights (in inches). Participants that reported extreme height values were not excluded from the study as failing to detect an effect can lend itself to interpreting study results.

As discussed, possible limitations in the current study included the number of male participants in the overall study (36 females, 10 males, 1 unidentified) and participants failing to

pay attention during the questionnaire. Another possible limitation of the current study was the environment in which participants were tested. Two enclosed rooms were used in the current study, each room having a table with a computer and chair. However, in one of the rooms there was an additional table with three chairs around it which did not serve as a highly controlled condition and may have influenced participants' responses. Sitting in an enclosed room such as a booth or cubicle with only a computer, table, and chair available where no distancing behavior is possible has been found to produce a higher accessibility of prime-related items (Cesario et al., 2010). Having extra furniture within the room may have distracted participants from the current task, as shown in a study by Haworth Human Performance Lab, where they found that visual distractions increase cognitive load, pulling resources from high focus work (Johnson, 2017).

Future research should include further evaluations of the effect of gender on finding an assimilative priming effect. Emotion literature suggests that, on average, females are more sensitive to emotional stimuli, such as emotional faces, compared to males (Thayer and Johnson, 2000) and that this sensitivity occurs automatically (Hatfield, Cacioppo, and Rapson, 1994). In a study conducted by Abbassi, Blanchette, Sirmon-Taylor, Ansaldo, Ska, and Joanette (2019), they examined the pattern of quick affective priming in female and male participants. It was found that there is rapid priming in both hemispheres of females contrary to males for whom rapid affective priming was lateralized to the lateral hemisphere, which seems to support the quick activation of emotional words in females. If females have quicker activation of emotional words than males, this may be a variable to explore when examining conditions to demonstrate hostile priming effects. Across the traits that were used to rate Ronald, five out of seven traits were rated more negatively among female than male participants.

References

Abbassi, E., Blanchette, I., Sirmon-Taylor, B., Ansaldo, A. I., Ska, B., & Joanette, Y. (2019). Lateralized Affective Word Priming and Gender Effect. *Frontiers in psychology*, *9*, 2591. https://doi.org/10.3389/fpsyg.2018.02591

Anderson, C. A., Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature. *Psychological Science*, 5, 353-359. doi:10.1111/1467-9280.00366

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893–912.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditionally automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*, 104–128.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–240.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. Reis & C. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 253–285). New York: Cambridge University Press.

Bargh, J., Gollwitzer, P., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals. *Journal of Personality and Social Psychology, 81*(6), 1014-1027.

Buchanan, T. (2015). Aggressive priming online: Facebook adverts can prime aggressive

cognitions. *Computers in Human Behavior*, *48*, 323–330. https://doi.org/10.1016/j.chb.2015.01.072

Cesario, J., Plaks, J. E., Hagiwara, N., Navarrete, C. D., & Higgins, E. T. (2010). The ecology of automacitiy: How situational contingencies shape action semantics and social behavior. *Psychological Science, 21*(9), 1311-1317

Chartrand, T. L., & Bargh, J. A. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology*, 71, 464–478.

David, E.J.R., & Okazaki, Sumie. (2010). Activation and Automaticity of Colonial Mentality.(Report). *Journal of Applied Social Psychology, 40*(4), 850.

Djiksterhuis, A., Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annual Review of Psychology*, 61, 467 – 490.

Gulan, T., & Valerjev, P. (2010). Semantic and related types of priming as a context in word recognition. *Review of Psychology*.

Gustafsson, A., & Lindenfors, P. (2006). Human size evolution: No evolutionary allometric relationship between male and female stature. *Journal of Human Evolution, 47*, 253–266.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*, 323–340.

Herr, P. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology, 51*(1986), 1106-1115.

Higgins, E., Bargh, J., & Lombardi, W. (1985). Nature of Priming Effects on

Categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 59-69.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins and A. W. Kruglan- ski (Eds.), *Social Psychology: Handbook of basic principles* (pp. 133-168). New York: Guilford.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, 13, 141–154.

Jacobs, A., Grainger, J., & Ferrand, L. (2015). The incremental priming technique: A method for determining within-condition priming effects. *Perception & Psychophysics*, *57*(8), 1101-1110. doi: 10.3758/bf03208367

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*,55, 893-905.

Kawakami, K. , Dovidio, J. , & Dijksterhuis, A. (2003). Effect of social category priming on personal attitudes. *Psychological Science*, 14, 315-319.

Kay, A. C., & Ross, L. (2003). The perceptual push: The interplay of implicit cues and explicit situational construal in the prisoner's dilemma. *Journal of Experimental Social Psychology*, 36, 634–643.

Komorita SS, Parks CD. Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology*. 1995;46:183–207.

Loersch, C., & Payne, B. K. (2012). On mental contamination: The role of (mis) attribution in behavior priming. *Social Cognition, 30,* 241-252. doi:10.1521/soco.2012.30.2.241

Mattler, U. (2006). On the locus of priming and inverse priming effects. *Perception & Psychophysics*, *68*(6), 975-991. doi: 10.3758/bf03193359

McConahay, J. B. (1986). Modern racism, ambivalance, and the modern racism scale. In J. E Dovidio & S. L. Gaertner (Eds.), Prejudice, discrimination, and racism (pp. 91-125). Oralando, FL: Academic Press.

Molden, D. C. (2014). Understanding priming effects in social psychology: What is "social priming" and how does it occur? *Social Cognition*.

Otten, S., & Stapel, D. (2007). Who is this Donald? How social categorization affects aggression-priming effects. *European Journal of Social Psychology*, *37*(5), 1000-1015.

Sperber, R.D., McCauley, C., Ragain, R.D. et al. Memory & Cognition (1979) 7: 339. https://doi.org/10.3758/BF03196937.

Srull, T.K, & Wyer, R.S (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37*, 1660-1672.

Srull, Wyer. "Category Accessibility and Social Perception: Some Implications for the Study of Person Memory and Interpersonal Judgments." *Journal of personality and social psychology*38.6 (1980): 841–856. Web.

Takarada, Y., Nozaki, D. Motivational goal-priming with or without awareness produces faster and stronger force exertion. *Sci Rep* 8**,** 10135 (2018). https://doi.org/10.1038/s41598-018-28410-0

Thayer, J. F., and Johnson, B. H. (2000). Sex differences in judgment of facial affect: a multivariate analysis of recognition errors. *Scand. J. Psychol.* 41, 243–246. doi: 10.1111/1467-9450.00193

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306. http://dx.doi.org/10.1126/science.2296719

Ven, M. (2019). Supplemental Material for Expectations and Perceptual Priming in a Visual Search Task: Evidence From Eye Movements and Behavior. *Journal Of Experimental Psychology: Human Perception and Performance*. doi: 10.1037/xhp0000618.supp

Zhang, Q., Espelage, D. L., & Zhang, D.-J. (2018). The Priming Effect of Violent Game Play on Aggression Among Adolescents. *Youth & Society*, 0044118X18770309.

Zhang, Q., Zhang, D., & Wang, L. (2013). Is Aggressive Trait Responsible for Violence? Priming Effects of Aggressive Words and Violent Movies. *Psychology*. https://doi.org/10.4236/psych.2013.42013

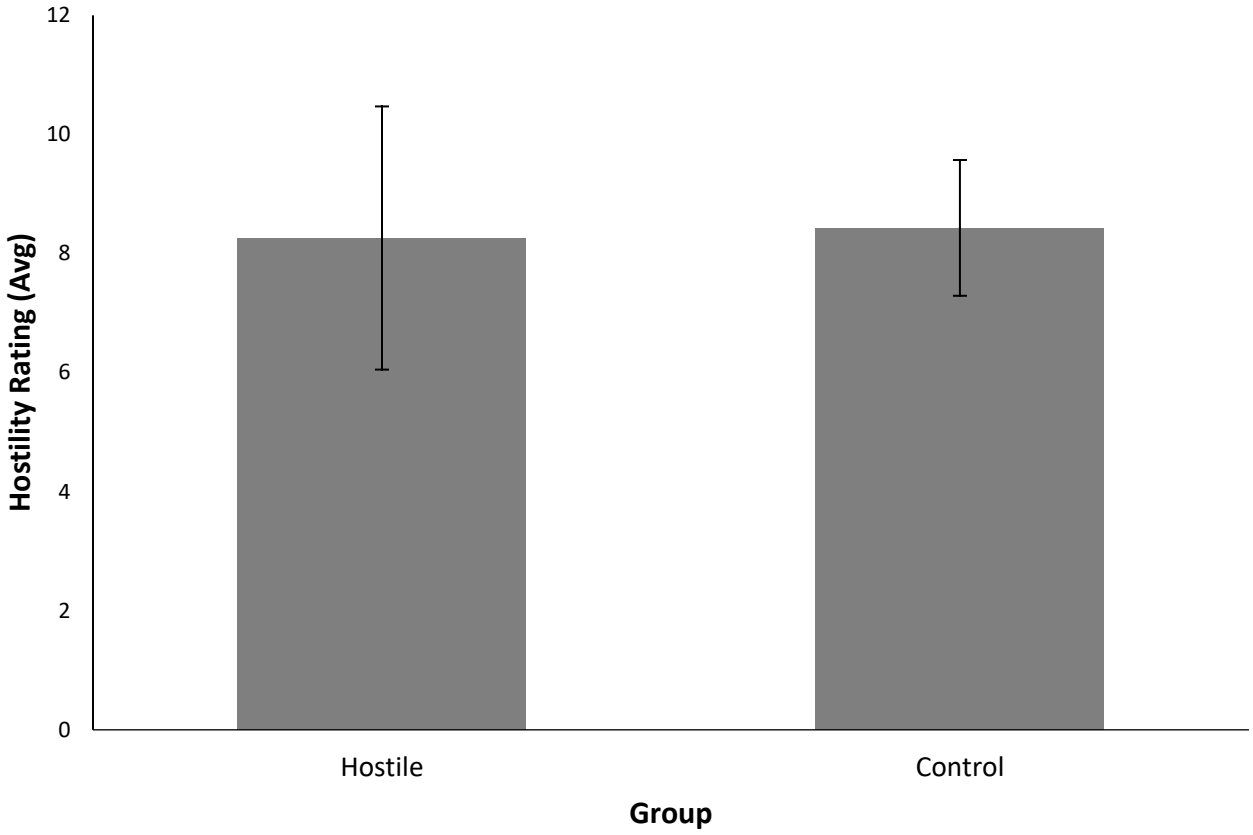

Figure 1. Main study: Average level of hostility rating in the hostile and control group with SD bars.

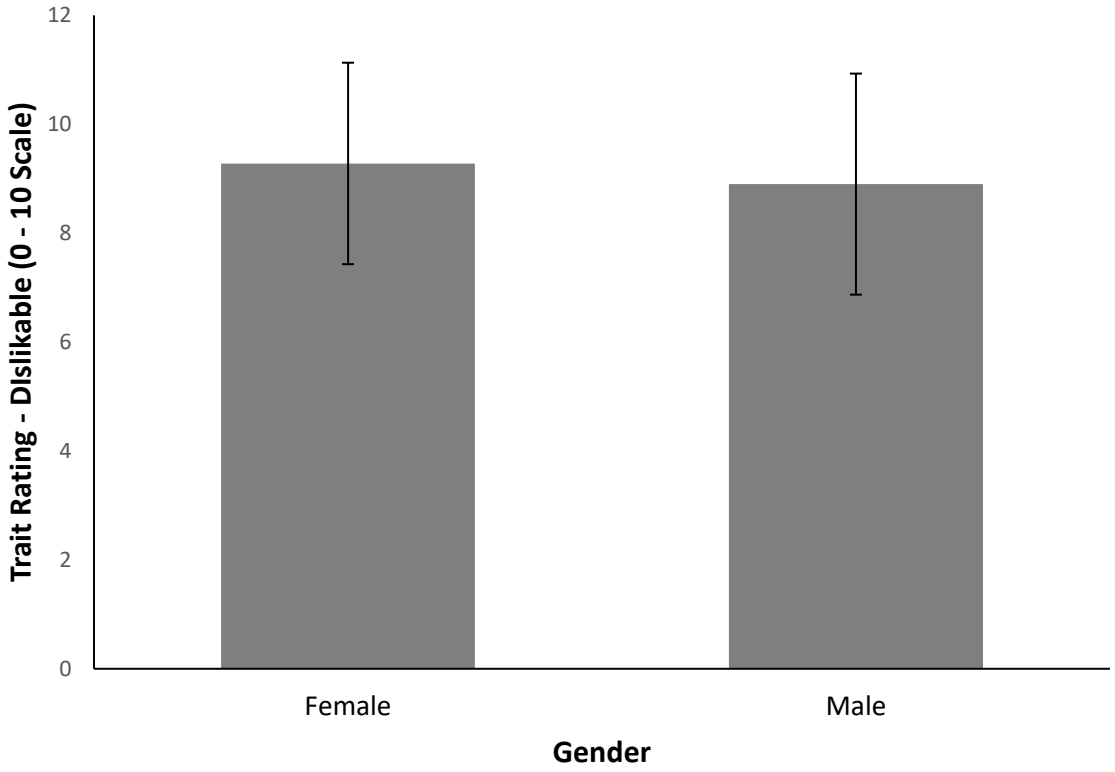

Figure 2. Average rating of trait *dislikable* among female and male participants with SD bars.

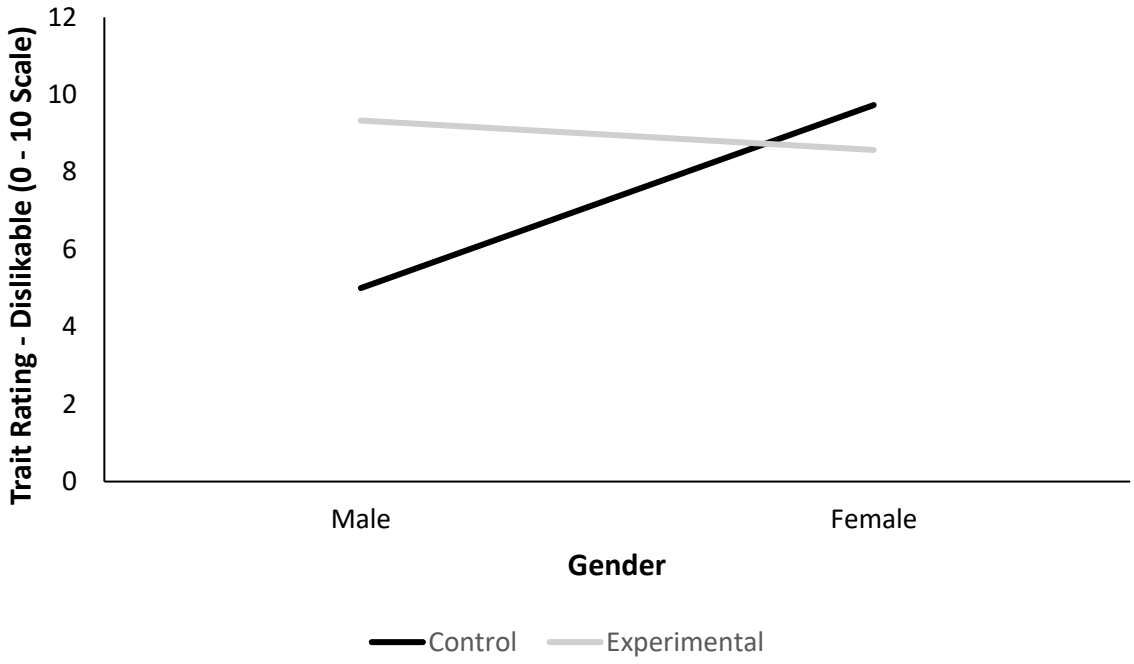

Figure 3. Average rating for trait *dislikable* in the experimental and control group separated by male and female participants.

Table 1
*Cell Means of Gender * Group on Trait Rating - Dislikable*

					95% CI	
Gender	Group	N	Mean	Standard Error	Lower Bound	Upper Bound
Male	Control	1	5.00**	1.74	1.48	8.52
	Hostile	9	9.33	0.58	8.16	10.51
Female	Control	22	9.73**	0.37	8.98	10.48
	Hostile	14	8.57	0.47	7.63	9.51

Note: Mean values marked with an asterisk (**) were significant at $p < 0.05$ level.

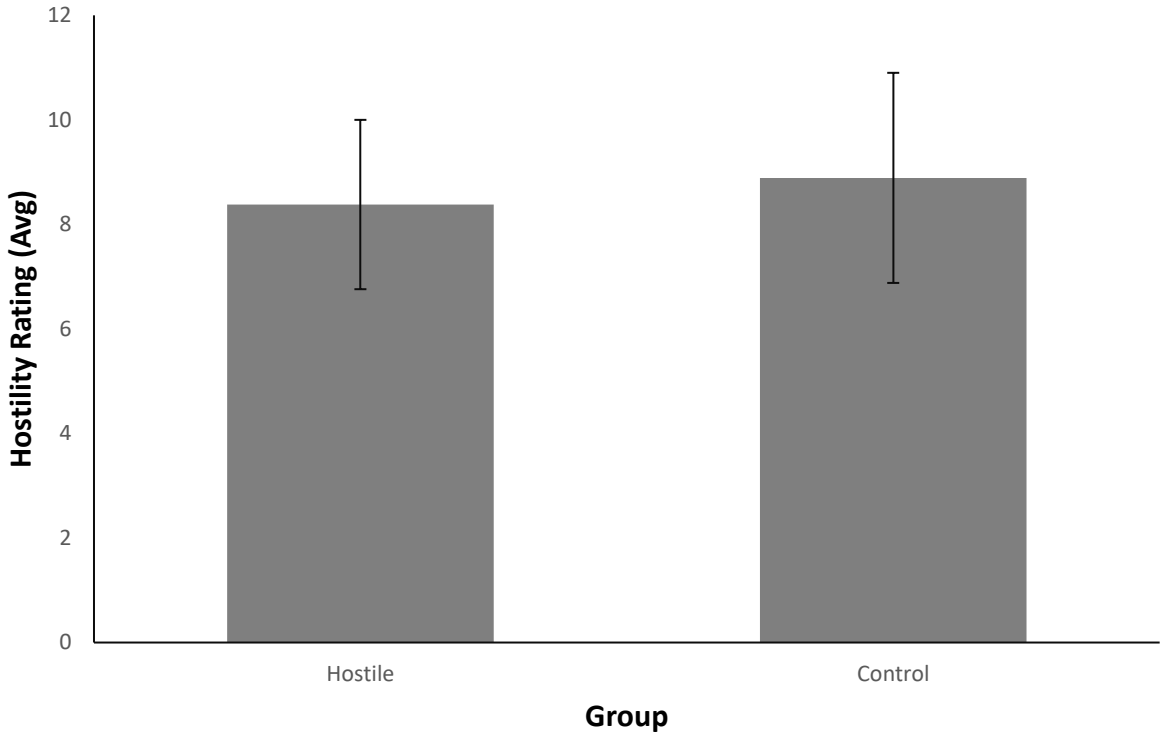

Figure 4. Ohio Sample - Average level of hostility rating in the hostile and control group with SD bars.

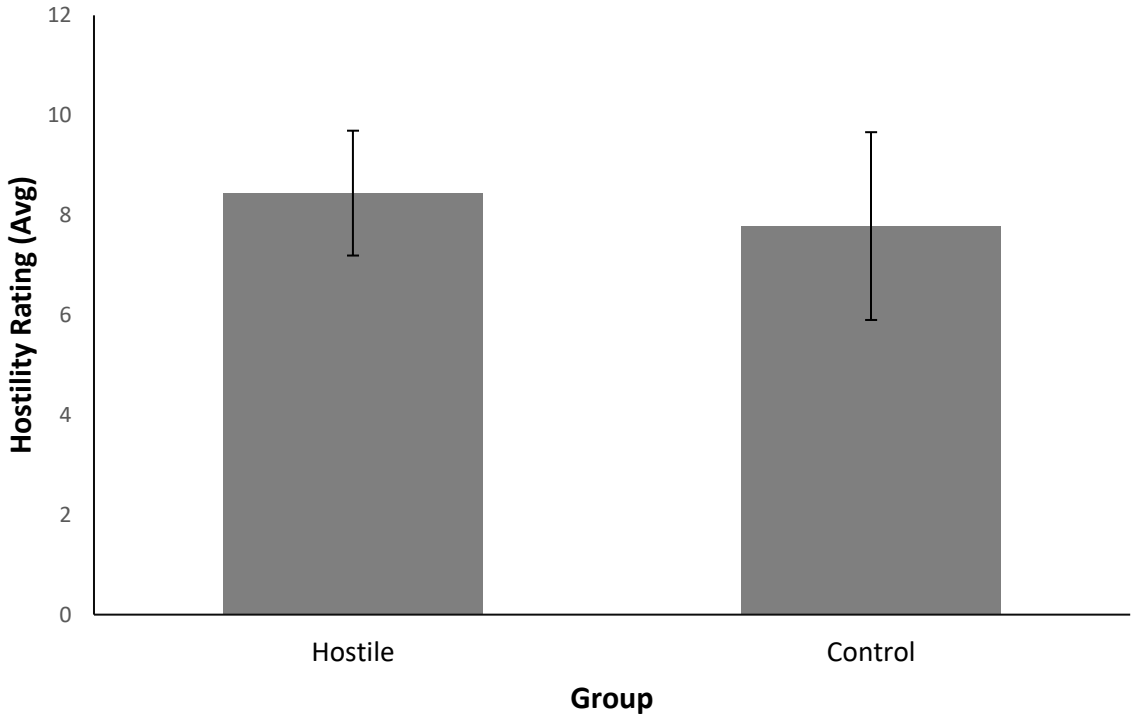

Figure 5. Washington Sample - Average level of hostility rating in the hostile and control group with SD bars.

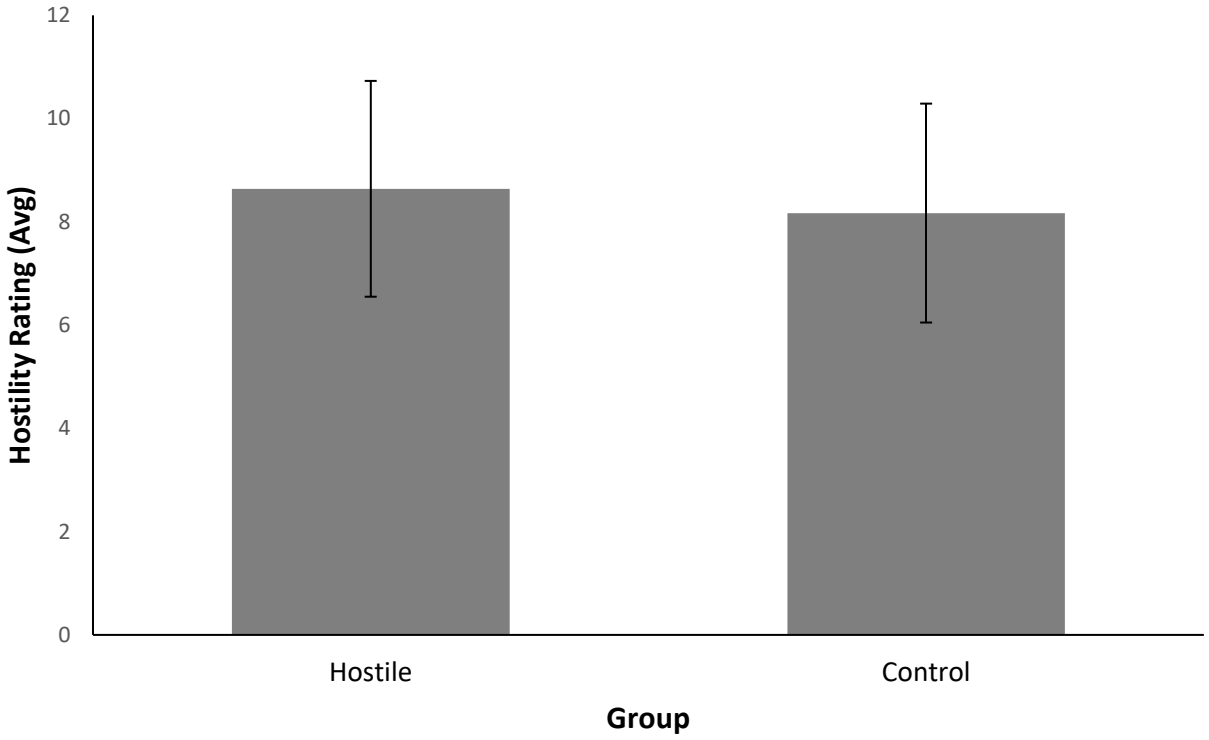

Figure 6. Virginia Sample - Average level of hostility rating in the hostile and control group with SD bars.

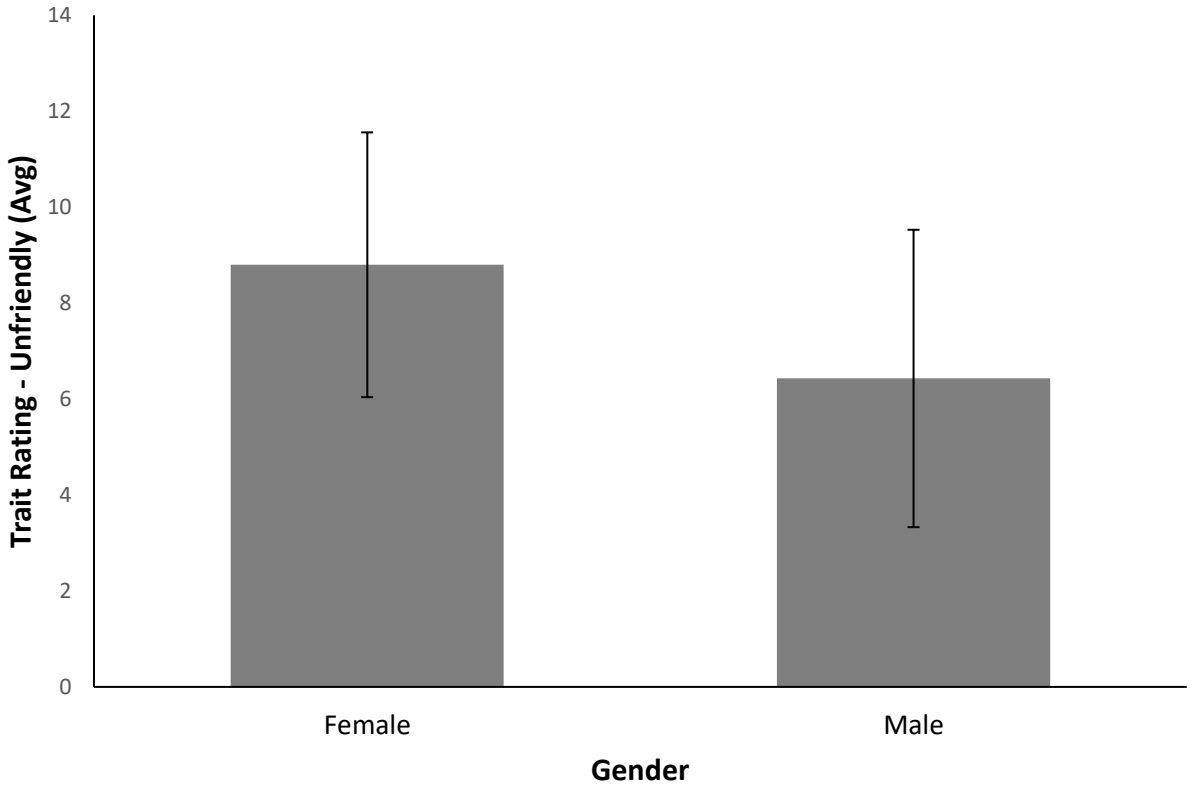

Figure 7. Virginia sample - Average rating of trait *unfriendly* among female and male participants with SD bars.

Appendix A

I ran into my old acquaintance Ronald the other day. On our way to lunch, he mentioned that he refused to discuss politics with one of his associates that morning although I know he's still very interested in foreign affairs. When we got to the restaurant, Ronald told the parking attendant not to scratch his car. Lunch was fine, although soon after we got a table, Ronald insisted that the waitress replace all the silverware because it was dirty. We must have spent a good two hours talking about old times. After lunch, we went downtown for a walk in the park. As usual, the park was very crowded and full of solicitors. Ronald told a beggar who asked for 25 cents to go get a job, and a little later, he refused to give money to the United Fund. For the most part, time seemed to pass pretty quickly and we agreed to meet again as soon as possible.

Appendix B

PRIMING AND AUTOMATICITY RESEARCH

Table 2: Example of funneled debriefing procedure for supraliminal priming task

The experimenter proceeds to ask the participant the following questions, and records the answers given:

1. What do you think the purpose of this experiment was?
2. What do you think this experiment was trying to study?
3. Did you think that any of the tasks you did were related in any way?

   (if "yes") In what way were they related?
4. Did anything you did on one task affect what you did on any other task?

   (if "yes") How exactly did it affect you?
5. When you were completing the scrambled sentence test, did you notice anything unusual about the words?
6. Did you notice any particular pattern or theme to the words that were included in the scrambled sentence test?
7. What were you trying to do while reading the behavioral phrases on the computer monitor? Did you have any particular goal or strategy?

(Source: Chartrand & Bargh, 1996, Experiment 1)