January 2016

Creation Of A Conditional Reasoning Task For Hostile Sexism

Travis Daryl Clark

CREATION OF A CONDITIONAL REASONING TASK FOR
HOSTILE SEXISM

by

Travis Daryl Clark
Bachelor of Science, East Tennessee State University, 2011
Masters of Arts, University of North Dakota, 2013

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

August
2016
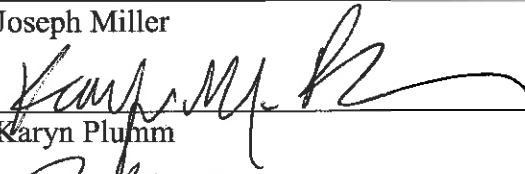
This dissertation, submitted by Travis Daryl Clark in partial fulfillment of the requirements for the Degree of Doctorate of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Heather Terrell

Joseph Miller

Karyn Plumm

Joelle Ruthig

Elizabeth Legerski

This dissertation is being submitted by the appointed advisory
committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Grant McGimpsey
Dean of the School of Graduate Studies

7/5/16

Date

PERMISSION

Title   Creation of a Conditional Reasoning Task for Hostile Sexism

Department   Psychology

Degree   Doctor of Philosophy

       In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in her absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

<div align="right">
Travis Daryl Clark
August 5, 2016
</div>

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

I wish to thank my advisory committee for their time and help during my doctoral

studies at the University of North Dakota.

Dedicated to the memory of my brother, Jordan Carter.

ABSTRACT

The Ambivalent Sexism Inventory (ASI) is a widely used measure of sexism. The current project describes the creation and administration of an alternative measurement of sexism that utilizes a conditional reasoning task (CRT) while supporting the theoretical underpinnings in Ambivalent Sexism Theory. Arguments in favor of creating a new scale are presented along with arguments for the specific use of a CRT. By creating a conditional reasoning task for hostile sexism (CRT-HS), the current project takes the first step toward creating a new indirect alternative for measuring sexism.

CHAPTER I

INRODUCTION

The Ambivalent Sexism Inventory (ASI) is a widely used measure of sexism based on the benevolent and hostile sexism distinction developed by Glick and Fiske (1996) and later adopted by other researchers (for a review, see Glick & Fiske, 2011). The goal of the current project is to create and validate an alternative measurement of sexism that utilizes a conditional reasoning task (James & Mazzerole, 2002) based on the theoretical underpinnings of Ambivalent Sexism Theory (Glick & Fiske, 1996). Conditional reasoning tasks have been developed to assess several implicit aspects of personality (James & LeBreton, 2010) when direct behavioral measurements are unavailable (i.e., aggression) or when self-report would be problematic (i.e., motive to achieve). The proposed Conditional Reasoning Task for Ambivalent Sexism (CRT-AS) would broaden the available horizons for research on hostile and benevolent sexism. The current project represents the first step in the creation of a CRT-AS by creating and validating a measurement for one of its subcomponents, hostile sexism.

What is a conditional reasoning task?

The conditional reasoning task (CRT) was designed by James (1998) as a methodology that explores latent motivations indirectly. James and Mazerolle (2002) argue that the CRT takes advantage of individuals' tendencies to believe their own behavior is rationally motivated while at the same time being subtly influenced by

previously held beliefs and values. The CRT construction process is part of an effort by many researchers to overcome the limitations of self-report and other research methods. The CRT approach was designed, in part, to fit expert suggestions for future directions in assessing personality in work organizations (James & Mazerolle, 2002; Chapter 4).

Conditional reasoning tasks use items that on the surface appear to measure logical reasoning and cognitive ability; in reality, one criterion for a good CRT is that it does not correlate with intelligence (James & Mazerolle, 2005). Instead, items contain some responses that are more likely to be chosen based on the cognitive biases of the survey taker. The cognitive biases are referred to as justification mechanisms (James, 2008). In a typical CRT item, participants are given a logic problem with four multiple-choice outcomes and tasked with providing the correct, logical conclusion. The format of these questions will be familiar to any participant who is familiar with the ACT, SAT, LSAT, or other standardized tests. One of the four choices is logically consistent and more likely to be chosen by a biased individual (e.g. aggressive or justifying substance abuse), one item is logically consistent and more likely to be chosen by an unbiased individual (e.g. not aggressive, not substance-abusing), and the remaining items are not logically consistent with the premises of the question (James & LeBreton, 2010; Bowler, Blowler, & James, 2011).

How is a Conditional Reasoning Task developed?

James and LeBreton (2012; p. 59) suggested steps for identifying justification mechanisms for motivations. These steps were employed to develop the CRT presented in the current studies. The final goal of these steps is to develop and validate a new measurement of benevolent and hostile sexism that uses conditional reasoning items.

The first step toward developing a new measure is to clearly match the proposed items with an established psychological theory. This process begins by identifying how individuals who are high, moderate, and low on a trait may differ in behavior. With regards to ambivalent sexism, individuals high, moderate, or low in hostile and/or benevolent sexism may engage different cognitive strategies to appraise situations that are relevant to gender relationships.

James and colleagues (2005) summarized their own steps to follow Ozer's (1999) best practice principles for making a good measuring instrument:

"These principles are (1) the content of the instrument should relate rationally to a psychological theory, (2) the item characteristics, scale characteristics, and factor structure of the instrument should be consistent with the psychological theory, and (3) the instrument should possess demonstrably high validities for the most theoretically relevant inferences." (James et al., 2005; p. 70)

It is appropriate, if these conditions are met, to use the CRT methodology to assess sexism. After this justification, the process of developing a CRT (James et al., 2005) will be outlined to elucidate the theoretical underpinnings of the proposed new scale.

This process of developing a CRT has been successfully implemented for several personality traits. The most prominent CRT is the Conditional Reasoning Task for Aggression (CRT-A; James & McIntyre, 2000). In many regards, the steps to validate a new CRT run parallel with the steps used to validate the CRT-A (James & McIntyre, 2000). LeBreton and colleagues (2007) tested the CRT-A (James & McIntyre, 2000) for several methodological issues and found the scale gives reliable data in several populations, unless participants are explicitly told the purpose of the instrument.

What is Ambivalent Sexism Theory?

Glick and Fiske (1997) developed Ambivalent Sexism Theory as an alternative framework for sexism toward women in response to changes in self-reported sexism over the previous few decades as well as changes in how similar topics, such as racism, were being studied. As Glick and Fiske (2011) retrospectively explained, Ambivalent Sexism Theory was born of several needs in sexism research. Popular scales to measure sexism such as the Attitudes Toward Women Scale (AWS; Spence, Helmreich, & Stapp, 1973) contained items that in the 1990s were becoming dated. Research on racism at the time was also concerned with modern forms of racism which tended to be subtler. Modern racism could also contain a pronounced ambivalence toward minority groups, such that pervasive negative stereotypes now coexisted with "White guilt" and other positive but possibly condescending feelings (Dovidio & Gaertner, 1986).

On a continuum between extremely direct ("How much do you dislike women?" being on this extreme) and extremely oblique (e.g., an implicit association task wherein "bad" words may be associated with women milliseconds faster than with men) measurements of antipathy, the study of sexism followed the study of racism in a paradigm shift toward less direct methods. Many authors saw a connection between racism research and something missing from sexism research. If researchers studying racism were improving their measurement tools by involving subtlety, ambivalence, and updated language, could researchers interested in sexism start using these same techniques? This thread of thought can be found in all of the introductory works to the sexism scales discussed below. These measurement instruments, such as the AWS,

incorporated elements of subtler forms of sexism while other authors focused on issues such as ambivalence.

By redesigning a measure of sexism to include both negative and positive stereotypes of women and women's roles in society, Glick and Fiske (1997) created a modernized scale that identified two related but distinct attitudes. The Ambivalent Sexism Inventory (ASI) is based on assumptions about gender roles that are more contemporary and fall under either hostile or benevolent categories. Hostile sexism includes traditionally prejudicial attitudes, corresponding to items on the scale such as "women seek to gain power by getting control over men." This conceptualization of hostile sexism had most in common with previous measures intended to study sexism. The more novel contribution of Ambivalent Sexism Theory was the introduction and validation of a benevolent sexism subscale. Benevolent sexist attitudes are subjectively positive and may lead to positive behaviors, but these positive attitudes and behaviors are directed at keeping women in a subservient or "weaker" (literally and metaphorically) social role. A person with benevolent sexist attitudes would be more likely to endorse the ASI item "A good woman should be set on a pedestal by her man."

Could a CRT be made to assess sexism?

To develop a CRT for ambivalent sexism, items were constructed to represent a variety of facets of sexism. As with the CRT-A, no one item was used as a representation of participants' sexist attitudes. Items were constructed that present participants with a variety of situations designed to tap into sexist bias; the sum of participants' sexist responses approximated a distribution with more sexist respondents indicating greater proclivity to select logically correct but sexist answers. From the inception of the original

ASI, the goal was to assess aspects of the sexist attitudes that less subtle instruments would miss. The current project took one step further by couching the self-report items in a logical reasoning task instead of presenting them to participants as statements to agree or disagree with.

In devising new conditional reasoning items to assess Ambivalent Sexism Theory, the literature on conditional reasoning tasks to assess implicit personality (e.g. James & LeBreton, 2012) as well as LSAT preparatory materials were used to write and rewrite convincing logic problems. These logic problems do not need to be *too* convincing, as previous research on aggression shows. For the conditional reasoning task of aggression, the "aggressive" option choices on specific questions do not necessarily appear to be logical to non-aggressive individuals; as James and colleagues (2005; p. 73) summarize, the literature on aggression predicts that aggressive individuals will be subject to biases that increase the likelihood they will view the aggressive options as correct choices. In some sense, this is the crux of the CRT procedure. It is being reiterated here because the proposed item choices on the CRT-AS were similarly designed to appear to an objective observer to be clearly illogical. These answer choices appeal to individuals with sexist biases who are not objective. In some sense, rational thinking comes after a decision has been made and is used to justify the decision. This inversion of rational thinking has been proposed for other domains in social psychology; Haidt (2001) likened it to an "emotional dog" wagging its "rational tail."

In creating a CRT, the steps followed in this project followed the suggestions of James and LeBreton (2012; Chp. 3, The development of conditional reasoning problems). These guidelines include exemplary (p. 75) and poor (p. 78) examples of conditional

reasoning items as well as guidelines to construct the former while avoiding the latter. An analysis of the pitfalls often made in designing CR questions is presented. Suggestions for distraction items, which comprise two out of four options for every item, are presented. Potential outcomes of biased reasoning (p. 107), an example of a differential framing table (p. 108), and problem ideas with example structures (p. 110) are all presented to facilitate question generation.

James and colleagues (2005) outlined many of the difficulties in assessing a measure that is explicitly unconscious. Self-report measurement of an implicit effect is not feasible (Nisbett & Wilson, 1977). James (1998) recommended the collection of data by inductive reasoning as an effective alternative way to gather data on implicit biases. In past research, participants completing inductive reasoning problems believed that their intelligence or verbal reasoning skills are being tested. This is likely because inductive reasoning problems follow the same format as true reasoning tasks found in intelligence and aptitude tests such as the Scholastic Aptitude Test (SAT) and others (James et al. 2005).

Is a new measurement of ambivalent sexism necessary?

This paper elucidates the development steps of a CRT with the assumption, thus far, that a CRT could be legitimately developed to assess sexism, so the next step is to justify why a new measure is necessary. While the ASI has been empirically successful and has helped to advance our understanding of sexism, it is a self-report measure that would benefit from additional validation from other data sources. Much of the literature on hostile and benevolent sexism uses self-report data, which in itself has its merits but it is not without its limitations. Homogenous data collection techniques introduce the

possibility of systemic error or response biases inherent to the data collection technique, not the measurement instrument (Williams, Hartma, & Cavozotte, 2010). This problem of shared variance can be addressed in several ways, but perhaps the simplest is to make use of a variety of measurement techniques.

Another weakness of the ASI is its applicability to different subcategories of women. The ASI was designed to measure attitudes toward women in general, but new frontiers of research and theory in feminism and psychology involve acknowledging and studying the distinctions among different groups of women. Sexist prejudice differs for minority women (and differs for each ethnic identity) and stereotypes about different occupations (or lack thereof) have been shown to change women's sexist beliefs about other women (Becker, 2010). While there are studies (e.g. Becker, 2010; Berdahl & Moore, 2006) that have used the ASI to assess attitudes toward different groups, it is argued a new scale could make studying women with intersecting identities easier.

Self-report measure critique. Self-report measures are a well-supported and useful methodological choice, but implicit or indirect measures do offer advantages. Attitudes, such as sexist ones, can be measured directly and explicitly, such as asking participants how positively or negatively they feel about women; attitudes can be measured indirectly by asking (as the ASI does) oblique questions concerning women. In a step farther from direct self-report, attitudes can be measured based on implicitly expressed notions, such as the biases captured in the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). The contemporary study of prejudice has shifted in recent years from more direct measurement tools to ones that are subtle, less direct, or measure implicit attitudes (Cunningham et al., 2001; e.g. Dovidio, Kawakami,

& Gaertner, 2002). Indirect measurements of prejudicial attitudes on behavior have several advantages. Prejudicial attitudes may be subliminal, influencing behavior, but not explicitly accessible to conscious awareness (Chen & Bargh, 1997). In other words, some participants may not be aware of the biases they hold. Other participants may be aware of their biases, may be unwilling to disclose their biases due to social desirability effects. As overt prejudice has become less socially acceptable, discriminatory behavior may become subtler (Sue et al., 2007). Implicit and explicit discriminatory attitudes are highly correlated, but distinct (Cunningham et al., 2001). Implicit measures may be better suited to capturing forms of racism such as microaggressions or aversive racism (Sue et al., 2007).

Research on prejudice against minority groups and research on sexism have strongly influenced each other as each field explores similar questions (Glick & Fiske, 2011). As feminist principles and gender equality become more socially acceptable, paternalistic or sexist attitudes may persist while participants feel social pressure not to express these attitudes. Some feminist critics argue that media narratives of female success lull women into a false sense of equality that impedes their motivation to seek equal opportunities (Douglas, 2010). Indirect measurements of sexist attitudes, like indirect measurements of racist attitudes, have the benefit of being less influenced by demand characteristics or social pressure. Researchers must be careful to avoid response bias when dealing with sensitive issues such as race, and using measurements less prone to response bias is one easy step in the right direction (Morgeson, et al. 2007). The ASI, along with other measures, was created in response to a need for less direct measures of sexist attitudes.

The CRT is one option for an even less direct form of measurement. The most direct way to assess sexism would be to ask participants "How prejudiced are you against women?", but answers to this question would be highly influenced by the social acceptability (or lack thereof) of being prejudiced against women. The ASI sidesteps this and other issues by asking participants about their sexist attitudes indirectly. This feature of the ASI is no accident, of course—the original authors found that items aligning with Hostile Sexism that were too direct ("Women are inferior to men") would not be endorsed by anyone, even respondents that would highly endorse several similar, but indirect, assertions about female inferiority (Glick & Fiske, 2001, p. 125). The CRT takes this indirect approach a step further—a well-designed CRT is not detectably about the attitudes in question at all, but merely another test of intelligence.

As time passes, cultural shifts—including increased awareness of the ASI itself—may contaminate assumptions about participants' naiveté to the purpose of the measurement instrument. As many of the participants in modern psychology research are college undergraduates (Henrich, Heine, & Norenzayan, 2010) and descriptions of hostile and benevolent sexism have made their way into prominent psychology textbooks (see for example Kenrick, Neuberg, & Cialdini, 2015 p. 353), this assumption comes into question more and more as a scale achieves prominence. If undergraduates are aware of the goals or hypotheses of a study due to prior information or piecing it together during the study procedures, they are unlikely to reveal this to experimenters (Blackhart et al, 2012).

Two decades old critique. As Glick and Fiske recalled their time creating the ASI and originally writing about ambivalent sexism theory (in Glick & Fiske, 2011), the

authors speculated upon the reasons that several new measures of sexism were in development simultaneously. For example, the Modern Sexism Scale (Swim, Aiken, Hall, & Hunter, 1995), among others, was created around the same time independently. Glick and Fiske (2011, p. 530) speculate that one possible reason for this movement is that various researchers all independently recognized a need for updated measures. Older measures of sexism or "appropriate" sex roles were at the time becoming outdated, as participants were less willing to agree to statements about cultural norms that increasingly reflected the past, not the present. While the criteria for what constitutes sexist language may have changed, sexist language will still be used; participants high in modern sexism may be unaware that their less poignant sexist comments qualify as sexist at all (Swim, Mallett, & Stangor, 2004).

In the decades before the ASI's creation, a shift toward feminist or more liberal values about women was observed in American samples (Twenge, 1997). In the twenty years since the ASI and other similar scales were developed, American culture has steadily shifted even farther from traditional sex roles. Several women (from all sides of the political spectrum) were prominent contenders for the American presidency in the 2000s, with increased public support for the idea of a female President of the United States (Streb, Burrell, Frederick, & Genovese, 2008). Workplace inequalities based on sex are a prominent issue (Korpi, Ferrarini, & Englund, 2013). Research on prejudice and bias has flourished in the last two decades, the ASI being one successful part of this research movement. Just as some the authors of the ASI recognized 20 years ago, it is again time for a new measurement tool for sexism.

Intersectionality critique. Feminist theory in the 20th century faced several challenges from within the feminist movement, including critiques that modern feminist movements were not adequately addressing the different realities experienced by minority women (Crenshaw, 1991). The concept of intersectionality is an acknowledgement and awareness that race, class, and sex cannot be meaningfully separated from an individual's experiences (Cole, 2009). Empirical studies in psychology often avoid the intersection of race and sex; this omission is due to many factors ranging from benign simplification to negligence (Cole, 2009). Research in ambivalent sexism theory has been conducted in many cross-cultural samples and appears to be broadly applicable to very diverse international cultures (Glick et al., 2004). There is no reason to disbelieve similar advancements in theory could not be achieved for how ambivalent attitudes toward women differ in various racial minorities and social classes. Research on instances of workplace harassment has found that sex alone and ethnicity alone both predict workplace harassment and have an additive effect on instances of harassment reported (Donelson & Moore, 2006; Raver & Nishii, 2010).

One current element missing from the ASI is the ability to collect data on attitudes toward women of particular categories such as women of a particular race or social standing. While the ASI could be retrofitted to this purpose by specifying a particular classification of woman the questionnaire is asking respondents to make judgments about, this would significantly alter the original intent and language of the questionnaire. Some sidestep this limitation through various means including the use of additional separate questions such as "What is your overall evaluation or general attitude toward [subtype]?" (e.g. Glick et al., 2015). These temporary solutions detract from the

standardization of the measurement instrument. Measurements like the ASI have been rigorously tested for validity and reliability.

Is a new measurement of sexism necessary? As it is argued in this paper: self-report limitations, the age of the scale, and intersectional research limitations all point to yes. The next steps would be examining the literature on ambivalent sexism and related topics to identify what cognitive biases might be active in individuals high in hostile and benevolent sexism. Then, several checkpoints must be met that the proposed CRT-AS must pass in order to be considered a rigorous application of the principles of a conditional reasoning task in service to Ambivalent Sexism Theory. These two steps (suggested and undertaken by James et al., 2005) follow below.

What cognitive biases do hostile and/or benevolent sexists have?

Stereotype content model approach. Where did the need for ambivalent sexism theory originate? Glick and Fiske (2011) recalled the origins of the scale (Glick & Fiske, 1996, 1997) being related to understanding the unique nature of prejudice against women. Ambivalent sexism theory proposes that sexism is composed of two dimensions that support each other. Under classic definitions of prejudice which necessarily link prejudice with negative beliefs (Allport, 1954), benevolent actions toward women would not be classified as prejudice. Being the recipient of benevolent actions like the attentions of an overly charming man or being helped more quickly in an emergency are positive— but after listing these responses in a condition asking about the advantages of being a woman, women subsequently were less likely to endorse a measure of collective action (Becker & Wright, 2011).

This is just one example of the undermining effect of benevolently sexist attitudes and actions; not only does benevolent sexism endorsement lower the likelihood that women will take action (e.g. Becker & Wright, 2011) but it decreases the likelihood that men will view gender inequality as unjust (e.g. Jost & Kay, 2005). Ambivalent Sexism Theory has now accumulated decades of disparate studies that all point to the powerful influences of both overt hostile sexism and covert benevolent sexism. Ambivalent Sexism Theory does not indicate that hostile or benevolent sexism are unilateral approaches taken by men or women, but rather that certain situations or contexts will evoke the need for these responses to stabilize the system of gender inequality (Glick & Fiske, 2001).

In what ways would benevolent sexism contribute to a superficial elevation of women's status that in reality undermines them in the workplace? When power differences exist, the cultures of the superior and inferior status groups adapt to this climate. Superior groups justify their status by attributing superior characteristics to themselves; even low-power group members may justify their position this way (e.g. Jost & Banaji, 1994; Jost & Kay, 2005). When a low-power group member is displaying the qualities associated with a high-power group member—such as a woman displaying status-relevant traits like ambition or intelligence—hostile sexism may be employed as a justification for behavior that denigrates the low-power group member (Glick & Fiske, 2001, p. 118). In situations where the low-power group members are not a threat, it should be expected that hostile attitudes are not needed to keep the group in its place but benevolent attitudes could be employed to reinforce that position.

For women, feeling a greater sense of entitlement predicts endorsement of benevolent sexism (Grubbs, Exline, & Twenge, 2014; Hammond, Sibley, & Overall,

2013). This relationship is much weaker or not present at all in men, suggesting that when women are enjoying the individual benefits of a patriarchal system they are more likely to employ benevolent sexist attitudes to justify this system. By participating in the patriarchal system that suppresses them, women gain benefits (such as the benevolent sexist attitudes discussed in this paper). By refusing to participate in this system, women may gain freedom and equality but be viewed as endangering the harmony of the gender dynamic. When women seek success at traditionally male jobs they are viewed more negatively by other women (Garcia-Retamero & Lopez-Zafra, 2006; Heilman et al, 2004).

The stereotype content model approach uses two distinct dimensions that outgroups can be judged on—warmth and competence (Fiske et al, 1999; Fiske et al., 2002). Perceptions of warmth include social indications of intentions while competence indicates the ability to act out one's intentions (Heflick et al., 2010). The judgement of others along these dimensions is supported by neuroscience indicating how social groups are judged at the level of the brain; heightened activity in the medial prefrontal cortex (mPFC) occurs during social cognition (Ochsner, et al., 2004). While the mPFC is triggered by socially active objects such as other humans, it is not activated by objects; the mPFC shows less activation when participants view pictures of extreme outgroups, such as drug addicts, than when participants are viewing disturbing objects (Harris & Fiske, 2006). Activation of the mPFC occurs when viewing scenes or pictures involving social objects or agents; lack of activation of this area when encountering low-agency humans supports the self-reported distinction made between those with high and low agency or competence (Heflick et al., 2010). Being objectified, or viewed as an object, is

15

not just a rhetorical device or theoretical claim by feminist scholars (see Nussbaum, 1999); women are *literally* viewed as having less agency by the mPFC when subjects are asked to focus on their appearance (Heflic, et al, 2010). The evidence from neuroscience that indicates women are viewed as less agentic converges with evidence from social psychology, providing a possible causal route for the systematic disadvantages women face when being treated as objects.

An outgroup that is viewed high in warmth is viewed as less threatening—women, for example, are more likely to be viewed as nurturing caretakers, a positive profile that does not threaten their lower social status (Fiske, et al., 1999). When women excel in non-status seeking domains such as warmth, there is no resultant punishment (Jackman, 1994). When women try to gain status through traditional pathways such as looks and a nurturing attitude, this can backfire if they are encroaching upon "male" avenues of success. Attractiveness is disadvantageous for women in managerial positions, but this effect does not extend to men in managerial or non-managerial positions (Heilman & Stopeck, 1985). Status-seeking traits are discouraged as they benefit the advancement of women without compensatory benefits that favor interdependency. Warmth is one social dimension that is positive but is non-threatening to the dominant group. When women display friendliness (Spears & Manstead, 1989), warmth (Fiske, Xu, Cuddy, & Glick, 1999), or behaviors that are consistent with traditional gender roles (Birnbaum, Ein-dor, Reis, & Segal, 2014) they are evaluated more positively or as more attractive.

Sibley and Wilson (2004) manipulated a vignette about a woman (enjoying or not enjoying casual flings) and found that men with a high sexual self-schema showed

increased hostile sexism toward the vignette subject who was non-traditional (i.e. enjoyed casual flings). When women fail to follow traditional gender norms, they are punished for it. Futhermore, women may not even need an outside observer to experience the effects of negative sentiments from others. Women experience negative effects of objectification even alone with their own thoughts, as self-objectification leads to negative feelings, body dissatisfaction, or poor cognitive performance (Fredrickson et al., 1998).

If a functionalist perspective of hostile and benevolent sexism is taken with regard to perceived warmth or competence of women, sexism would only be used when women are not fulfilling their role as an interdependent (but subordinate) social group. Thus, sexists would be motivated to select an answer less favorable to women but only in scenarios in which they are too high in agency or too low in warmth. Specific agency-related questions will address common workplace scenarios involving advancement and competency. Specific warmth-related questions will address interpersonal and social skills. Those high in benevolent sexism will tend to favor women in scenarios depicting social warmth; those high in hostile sexism will tend to punish women in scenarios depicting agency.

System justification. Jost and Kay's (2005) system justification approach to gender relationships utilizes the framework presented on warmth or communal characteristics versus competence or agency but explicitly adds a third point—the complementary nature of gender stereotypes. By being involved in a system in which every group has perceived advantages and disadvantages, the unequal distribution of power and wealth experienced by the groups can be perceived to be fair (Kay & Jost, 2003). We prefer to believe that the world is just and fair and are distressed by

17

indications to the contrary; a just world in which suffering is equitable is preferable to a world where victims suffer to no purpose (Lerner & Simmons, 1966).

High belief in a just world may predispose one to victim blaming (Furnham, 2003). Victimized groups may engage in victim blaming of others similar to them to maintain the belief that the victim deserved it; thus, group members are not at the same risk of harm unless they, too, behave in ways to deserve the treatment. Powerful groups may engage in victim blaming to maintain the belief that those in power are just and do not needlessly inflict suffering on those weaker than them; weak victims must have done something to deserve the punishment inflicted on them.

Victims of wife abuse are blamed more by women with a high positive attitude toward women and a high belief in a just world; men with an unfavorable attitude toward women were also more likely to blame the victim of wife abuse (Kristiansen & Giuletti, 1990). Sexist attitudes toward women also tend to be related to more permissive attitudes toward abuse against women. A recent study found that men self-reporting less hostility toward women were more likely to endorse using force to obtain sex without endorsing rape (Edwards et al., 2014). Based on their findings that men at different levels of hostility toward women differently endorse rape or forced intercourse, the authors suggest that different rape prevention educational programming is needed for the vastly different attitudes men have toward rape. Forced sexual intercourse is logically equivalent to rape, indicating that men who do not display a hostile affective response to women may see their forceful sexual advances as something that is approved or desirable. Edwards and colleagues suggest that education clarifying the definition of rape, among other things, could be necessary for this group (Edwards et al., 2014, p. 192).

To reduce cognitive dissonance, people rationalize their behaviors as supporting a worldview they can handle—whether this is the rationalization of "forced intercourse" as more socially acceptable than rape or the rationalization of benevolent sexism as more acceptable than direct acts of aggression against women. This approach of balance informed the development of the CRT-AS items. Men that are higher in benevolently sexist attitudes toward women should be more likely to mark more aggressive logical choices as correct, as the standards for hostile aggression are lower for these individuals. Women will be more likely to select hostile answers but only when their tendency for hostile sexism is greater.

Benevolent sexism can be truly beneficial at the individual level, a form of selfish choice benefitting individual social interactions—women may stand to gain material benefits from individuals as a result of benevolent actions. At the societal level, however, cultural reinforcement of benevolence towards women may stymie the efforts of women to receive equal treatment in non-traditional domains. Individuals may reward women's gender-conforming behavior, reinforcing a cycle of benevolent sexist attitudes in women (Becker, 2010). Women receiving the subjective benefits of benevolent sexism are then less prone to action (Becker & Wright, 2011). Benevolent sexism is a self-perpetuating set of beliefs. Men in the U.S. and China, for example, have highly correlated endorsement of hostile sexism and desire for marriage partners with traditional gender role beliefs (Chen, Fiske, & Lee, 2009). Families that are high in endorsement of traditional gender roles are more likely to carry these ideas into future generations.

Women may endorse hostile sexism as well. Becker (2010) discussed and empirically examined three possible reasons that women would endorse sexism. First,

women may endorse hostile sexist beliefs but only towards certain groups of women (such as feminists) while endorsing benevolent sexist beliefs only toward certain groups of women (such as housewives); second, women's internalization of hostile and benevolent beliefs may predispose them to endorsement. Another possibility is that women who had not internalized these beliefs may still be nudged to express them if reminded of non-traditional groups that the sexist beliefs target. The two studies conducted suggest that when women endorse hostile sexism, they are thinking of certain subtypes such as "career women and feminists" (p. 460). When women think of "housewives" they are more likely to endorse benevolent sexism. The author performed an experimental follow-up wherein women were directed to answer items about the common subtypes that were spontaneously called to mind in the first study; women directed to think of certain subtypes expressed different levels of sexism toward different groups. Again, hostile sexism was endorsed for career women and feminists while benevolent sexism was endorsed for housewives.

Current project goals and hypotheses

Using the ASI to study the interplay between sexist beliefs and sexist behavior has been informative for the study of individual behavior and the study of societal forces (Glick et al., 2004). A new measure of sexism that is complementary to Ambivalent Sexism Theory can further research in this area. With different types of measurement tools, researchers can control for the possibility that a particular type of data collection, such as self-report on Likert-type scales, is biasing our results (Williams, Hartman, & Cavazotte, 2010). The particular type of measurement tool presented here, a conditional reasoning task, has unique advantages and disadvantages which have been discussed.

These instruments should have high shared explanatory power with existing instruments measuring ambivalent sexism but also capture unique variance due to its unique method of measurement.

The current project is the first step in this research program, to create and validate a new measurement tool for hostile sexism. The long-term goal of this research program will be to fully validate a measurement that uses conditional reasoning items to assess both elements of Ambivalent Sexism Theory. The current project is focused on just one component, hostile sexism. Hostile sexism is much more susceptible to the problems outlined above; particularly, hostile sexism is more difficult to ask participants about directly. For the remainder of this paper a Conditional Reasoning Task – Hostile Sexism (CRT-HS) will be the focus of discussion. Future projects will incorporate successful results of creating a CRT-HS into further efforts to make a complete CRT-AS.

> Hypothesis 1: Items on the preliminary CRT-HS will cluster into one single factor assessing hostile sexism, similar to the hostile sexist subscale of the ASI.
>
> Hypothesis 2: Scores on the CRT-HS will be highly correlated with scores on the ASI subscales with a stronger correlation to the hostile subscale.
>
> Hypothesis 3: Responses on the conditional reasoning items will not be related to responses on actual reasoning items such as would be found in standardized testing situations.

Study 1 and Study 2 are reported separately. Study 1 includes a description of the steps used in item generation and item revision including a pilot study of revised items. Study 2 tests each of the hypotheses listed here with a representative sample of undergraduate students at a large Midwestern university.

CHAPTER II

STUDY 1 METHODS AND RESULTS

Study 1 Overview

The first step in the overall project was to create the items for the CRT-HS. Next, these items were reviewed for logic, legibility, and organization by a small set of raters. Finally, these items were given to a naïve sample along with the Ambivalent Sexism Inventory in order to generate item characteristics and examine the relationship with sexism.

Study 1 Methods

Item Generation

Items were generated based on several sources. First, items from the ambivalent sexism inventory itself were used as the basis of conditional reasoning questions. News stories and current events shaped other questions. Colleagues, primarily social psychologists or members of the Gender and Social Psychology (GASP) lab at the University of North Dakota, provided input on the first draft of items. A preliminary first draft of questions was established wherein each question contained a prompt, two logically incorrect answers, and two logically correct answers. One of the logically correct answers was a statement that could be considered sexist, promoting "traditional ideas of femininity", or supporting "traditional gender roles." The second logically correct response was created to be neutral in regards to gender roles, non-sexist, or slightly liberal with regards to femininity (feminist).

Next, members of the GASP lab were given a set of questions and asked to indicate, for each of four multiple choices, the extent to which each option could be correct. Research assistants were naïve to the nature and intentions of the conditional reasoning questions during this step; however, all members were knowledgeable about current and past literature on sexism and Ambivalent Sexism Theory. In this first draft process, research assistants could indicate all or none of the possible options were logically correct. This feedback was used to revise illogical options that looked superficially correct and revise logical options that looked incorrect. One weakness of this initial stage is that members of the GASP lab were much more likely to report that the neutral/feminist response was logically correct (relative to the sexist response). This pattern would be expected if the logically correct but sexist statements were unconvincing, but would also be expected if the logically correct but sexist statements *looked incorrect* to survey-takers that are biased by pro-feminist attitudes. Direct experience with members of the lab informs my decision to believe the latter explanation.

GASP lab members were also given the opportunity to provide feedback about each question. As lab members had all previously encountered research on sexism, many identified questions that were too obviously assessments of sexism. Questions were deleted entirely if lab members did not correctly identify the logically correct answer, expressed extreme skepticism about the question, or identified the question as too obviously assessing sexism or gender attitudes. The list of questions after this process was narrowed down to ten.

After narrowing down the list to potential candidate questions (found in Appendix F), the next step was to test these questions on a small, naïve sample. To be an accurate

measurement of individuals with implicit sexist motives, a small effect of hostile sexist (HS) items tested across a large sample should be found; follow-up analysis would show a subsection of the sample showing particularly high ratios of agreement with HS items. This is expected to match previous empirical findings that there is a distribution of hostile sexist values that centers on a moderate level with some in the extreme high range. Scores on the CRT-HS should be positively skewed.

Participants and Procedures

An undergraduate class (N = 134) at the University of North Dakota completed a survey packet for extra credit. This survey packet contained the final list of ten CRT-HS questions and the Ambivalent Sexism Inventory. Participants were instructed to select the "most logically correct" option on the CRT questions.

Study 1 Results

With perfectly written conditional reasoning questions, participants should avoid the logically incorrect question choices unless they are carelessly responding. Thus, the first check of the items in this step was to see which answer choices were most likely to be selected by participants. As Table 1 indicates, several questions had higher response rates for illogical answers than for logically correct answers. There are several explanations for this finding, but the most likely is that these options were appealing for the same reason that a well-placed incorrect distractor option on the S.A.T. is appealing—these options were *too superficially correct-looking*.

The range of scores on the CRT-HS was limited, with scores from 0 to 6 (out of a possible 10). The majority of participants (84%) had scores of 3 or less with a mean score of 2.157, *SD* = 1.320. Scores on the Hostile subscale of the ASI were normally

Table 1

Study 1 CRT-HS Answer frequencies

	Illogical Answer	Neutral	Sexist
CRT-HS 1	52 (38.8%)	25 (18.7%)	56 (41.8%)
2	58 (43.2%)	19 (14.2%)	57 (42.5%)
3	55 (41.0%)	29 (21.6%)	49 (36.6%)
4	23 (17.2%)	83 (61.9%)	28 (20.9%)
5	32 (23.9%)	59 (44.0%)	43 (32.1%)
6	108 (80.6%)	3 (2.2%)	23 (17.2%)
7	31 (23.1%)	93 (69.4%)	10 (7.5%)
8	97 (72.4%)	30 (22.4%)	7 (5.2%)
9	120 (89.6%)	8 (6.0%)	6 (4.5%)
10	71 (52.9%)	53 (39.6%	10 (7.5%)

distributed ($M = 25.05$, $SD = 7.80$). A small correlation was found between the CRT-HS and Hostile sexism, $r = .199$, $p = .021$, $r^2 = .04$.

Study 1 Discussion

The correlation between a CRT and a traditional survey of the same attitude is not expected to be perfect, however the correlation found in Study 1 was not convincing evidence that the CRT-HS could be used to supplement other measures of sexism. Based on Study 1, several revisions were made to the CRT-HS. Participants in Study 1 were very likely to select logically incorrect answers on several questions. If the goal of this project was to test participants' logical reasoning abilities, perhaps participants' variability in selecting illogical responses would be beneficial as it would allow discrimination between participants with high or low logical reasoning abilities. Discriminating between participants with high and low logical reasoning ability is actually a weakness to a CRT, therefore the appealing but logically incorrect responses were selectively edited to make them less appealing. The decision was made at this point to edit all illogical answers to be obviously incorrect, nonsensical, or incoherent (for the list of revised questions, see Appendix G).

CHAPTER III

STUDY 2 METHODS AND RESULTS

Study 2 Overview

With items narrowed down by a set of reviewers and edited to cover the

weaknesses identified in Study 1, the CRT-HS was administered to a new sample. Study

1 was a pilot test that included the CRT-HS and the ASI. Study 2 was a full survey that

also included other measurement instruments to establish convergent and divergent

validity for the CRT-AS.

Study 2 Methods

Participants

Participants (N = 266) were recruited from the University of North Dakota subject

pool via Sona-Systems. Students using Sona-Systems receive course participation or

extra credit for completing studies. Participants were primarily female (n = 200) with a

median age of 19 (M = 19.85, SD = 2.47). Participants were primarily freshmen (n = 107)

or sophomores (n = 81). The sample was primarily Caucasian (n = 246).

Materials and Procedures

After arriving at the study location at a specified time, participants were given a

packet of surveys and instructed they would have up to an hour to complete all study

procedures. Participants were seated at desks in a classroom and instructed to turn off all

cell phones for the duration of the study. The survey situation was, deliberately, similar in

feel to a standardized testing situation such as what many students encounter when taking the S.A.T.

After reading an informed consent page, participants encountered all of the measures listed below. After completing all survey procedures, participants dropped their anonymous responses into a collection box, were given a Debriefing form, and were thanked for their time.

Measurement Instrument Descriptions

Attitudes Toward Women Scale. The Attitudes Toward Women scale (Spence, Helmreich, & Stapp, 1973) is the oldest measure discussed here and is the least endorsed by participants in later decades (Twenge, 1997). This scale measures antipathy toward women and endorsement of traditional, restrictive gender roles. The relationship between this scale and the ASI is not straightforward; because the ASI measures both positive and negative beliefs about women but the Attitudes Toward Women scale measures endorsement of traditional gender roles, it is possible to highly endorse negative (or positive) views of women while either endorsing (or not) traditional gender roles. Past studies show a weak but significant correlation between the ASI and the Attitudes Toward Women scale for both sexes (Glick & Fiske, 2001), which is expected to replicate with any new scale based on Ambivalent Sexism Theory. We used the short version of the AWS for convergent validity testing (Spence, Helmreich, & Stapp, 1973; see Appendix B).

Neosexism Scale (Tougas et al, 1995). The Neosexism Scale, similar to the ASI, was designed to measure sexism expressed in subtle negative feelings toward women in egalitarian societies in which overt sexist acts are increasingly frowned upon (scale items

are reported in Appendix D). The Neosexism Scale measured negative affect only, unlike the ambivalence of the ASI. The relationship between neosexism and hostile sexism is strong and positive; neosexism is also strongly positively correlated with benevolent sexism, but only for women (Masser & Abrams, 1999).

Modern Sexism Scale (Swim et al., 1995). The Modern Sexism Scale was designed to capture facets of sexism beyond simple discriminatory behaviors. As overt racist behaviors become frowned upon, they may decrease in frequency; a new or previously ignored set of beliefs and behaviors may then rise. These beliefs, such as resentment for minorities getting "special treatment," interact with behaviors, such as early termination from a workplace, to support new forms of racism that are more subtle. The Modern Sexism Scale was designed to assess new forms of sexism in analogue to the changing ways of assessing racism; just like with racist behavior, sexist behaviors that were socially acceptable decades ago are now frowned upon. As with racist behaviors, however, sexist behaviors that are subtler but just as pervasive may continue (Swim et al., 1995). See Appendix C.

Social Dominance Orientation (Sibley, Wilson, & Duckitt, 2007). To show that the new measure of sexism is valid, it should also show the same relationship with other measurements as the ASI. One important measure is Social Dominance Orientation (SDO, see Appendix E; Pratto, Sidanius, Stallworth, & Malle, 1994). SDO is a measurement of intergroup competition. Individuals with high endorsement of SDO are more likely to believe that inequality between groups in society is normal and acceptable. Researchers have found that SDO is associated with hostile sexism (Sibley, Wilson, & Duckitt, 2007), perhaps because hostile sexists believe and accept that women are

inferior. This same pattern of relationships should be found between this scale and the proposed CRT-HS.

Study 2 Results

Item Analysis of the CRT-HS

While some illogical responses were marked by participants in the second sample, the rates of illogical responding were dramatically reduced (see Table 2). A well-written question will ask participants to judge between two equally correct alternatives with their choice reflecting subtle unconscious biases. High response rates to illogical questions implies that other features of the multiple choices, rather than logic, helps determine participant responses. In the Study 2 version, only three items had illogical responding rates higher than the response rate for one logical option; this is an improvement over the eight problematic items seen in Study 1.

Factor Analysis of the CRT-HS

James and colleagues (2005) used an exploratory factor analysis to determine the empirically derived factor structure of previous CRTs, acknowledging that the combination of items in a CRT may follow a latent structure that is more complex or deviant from the one proposed by theory. With proper theoretical backing, the factor structure should match the proposed Justification Mechanisms which match the factors reported in previous research on ambivalent sexism. Confirmatory factor analysis could be used to examine the factor structure first, but a data-driven preliminary approach could discover issues in data interpretation that a CFA would suppress. The analysis proceeded based on a principal components analysis of polychoric correlations, which are estimates of a linear variable, because the data collected from a conditional reasoning task is

30

Table 2

Study 2 CRT-HS Answer frequencies

	Illogical Answer	Neutral	Sexist
CRT-HS 1	65 (24.4%)	97 (36.5%)	104 (38.7%)
2	11 (4.1%)	182 (68.4%)	71 (26.7%)
3	60 (22.6%)	149 (56%)	57 (21.4%)
4	43 (16.4%)	53 (19.9%)	168 (63.2%)
5	72 (27.1%)	91 (34.2%)	102 (38.2%)
6	16 (6.0%)	129 (48.5%)	120 (45.1%)
7	91 (34.3%)	150 (56.4%)	24 (9.0%)
8	116 (43.6%)	89 (33.5%)	60 (22.6%)
9	27 (10.2%)	203 (76.3%)	34 (12.8%)
10	20 (7.6%)	191 (71.8%)	53 (19.9%)

ordinal (Holgado-Tello, et al. 2010).

In order to test hypothesis 1, that items on the CRT-HS will cluster into one single factor, a polychoric correlation matrix was created for analysis. As James and Lebreton (2012) reported, the factor structure of a CRT may be more complex than in traditional measurement instruments. Polychoric correlations and the subsequent EFA were calculated using the Polymat-c syntax provided by Lorenzo-Seva and Ferrando (2015).

Using Polymat-c, polychoric correlations between all items were calculated. These correlations are extrapolated estimates assuming that the dichotomously scored items represent an underlying factor structure. The polychoric correlation matrix (see Table 3) was then used to calculate an exploratory factor analysis in the same way that a Pearson r correlation matrix can be used in EFA.

Items on the CRT-HS did not cluster into one identifiable factor. As can be seen in Table 4, several items have weak (and two have negative) factor loadings, while others load strongly on multiple factors. More than two factors were indicated according to a cutoff Eigenvalue of 1.00, but SPSS terminated the calculation for a third factor after failing to converge on an appropriate solution over 100 iterations. When EFA is employed for a scale with Likert-type responses, many of the results seen for the CRT-HS would be warning signs for measurement inaccuracy.

Because the conditional reasoning task is not a measure of attitudes but is an estimation of implicit logical biases, it is reasonable that a factor structure be more complicated, even for a reliable measurement instrument. For this reason, the complicated factor structure of the CRT-HS has several possible interpretations. The most pessimistic, and parsimonious, interpretation is that the CRT-HS does not represent

a single construct well. The complicated factor structure could be a reflection of the type of data collection method. In a unidimensional measurement instrument, we would expect a single factor to emerge in EFA because participants are directly asked about different facets of the same construct (e.g. sexism, self-esteem, etc). In a conditional reasoning task, participants are necessarily not directly asked to evaluate the construct in question and in fact participants knowing the construct ahead of time is indicative of a weak CRT. Conditional reasoning questions are also scored dichotomously, which is problematic for several reasons.

In addition to the obvious estimation and interpretation errors with creating a scale and performing factor analysis on dichotomous responses, conditional reasoning questions truly have three possible answers. For data analysis, participants marking filler or illogical answers are scored the same way as participants marking the non-sexist but logically correct response. This means that a 0 indicates that the participant answered illogically or in a non-sexist manner while 1 indicates the participant answered in a sexist manner. This analysis rests on the assumption that illogical answers are rare. As can be seen in Table 2, this assumption did not hold for all CRT-HS items.

Scale Characteristics

    Ambivalent Sexism Inventory. The ASI is a well-established scale with two theoretically supported factors that are mildly correlated with each other. As such, a maximum likelihood Factor Analysis with a direct oblimin rotation was used. Allowing SPSS to empirically derive factors based on Eigenvalues greater than one returns a four-factor solution. Theory and past research both suggest a two-factor solution would be more appropriate. Analysis of the scree plot supports a two-factor solution as well; there

Table 3

Polychoric correlation matrix for CRT-HS in Study 2, created using Polymat-c

	CRT-HS 1	2	3	4	5	6	7	8	9	10
CRT-HS 1	1									
CRT-HS 2	-.150	1								
CRT-HS 3	-.047	.240	1							
CRT-HS 4	.028	.163	.008	1						
CRT-HS 5	-.128	.130	.193	.122	1					
CRT-HS 6	-.110	-.021	.059	.017	.160	1				
CRT-HS 7	-.019	.102	.124	.283	.144	-.109	1			
CRT-HS 8	-.009	.067	.120	.414	.168	.008	.113	1		
CRT-HS 9	-.021	.223	.071	.080	.165	-.076	.387	-.060	1	
CRT-HS 10	.066	.145	.058	.119	.197	.111	.162	-.041	.272	1

Table 4

Factor Analysis of the CRT-HS. Factor analysis created by analyzing the polychoric correlation matrix using Polymat-c

	Factor 1	Factor 2
CRT-HS 1	-.019	-.142
CRT-HS 2	.229	.416
CRT-HS 3	.226	.251
CRT-HS 4	.762	-.035
CRT-HS 5	.333	.476
CRT-HS 6	-.025	.200
CRT-HS 7	.361	.254
CRT-HS 8	.331	-.018
CRT-HS 9	.014	.489
CRT-HS 10	.022	.324

is a steep drop in additional variance explained after the second factor. A two-factor solution fits the hostile and benevolent subscales established theoretically. In Study 2, the Cronbach's alpha for the ASI is .938, the Benevolent subscale only is .901, and the Hostile subscale only is .935.

Attitudes Toward Women Scale (ATWS). The ATWS is scored such that higher scores indicate a profeminist response, therefore the CRT-HS is expected to be negatively correlated to the ATWS. The ATWS 25-item version had an essentially unifactorial structure in the sample analyzed by Spence, Helmreich, and Stapp (1973), validating this shorter version of the original 55-item measure. In our sample, however, the factor structure was more complex. A maximum likelihood exploratory factor analysis was performed with no rotation and determined that a solution as complex as six-factor is possible. The first factor explains 24% of the variance and additional factors add diminishing amounts of variance starting at 7%. According to decades of research on sexism, attitudes toward traditional femininity have been changing since the 1972 publication of these items (Buss, Shackelford, Kirkpatrick, & Larsen, 2001), which could explain why a scale this age did not follow its original theoretical factor structure. Despite this, the Cronbach's alpha was relatively high at .846.

Neosexism Scale. Neosexism is measured with 11 items, two of which are reverse-coded. A maximum likelihood exploratory factor analysis supported a single-factor solution. In the sample from Study 2, the Cronbach's alpha = .781.

Modern Sexism Scale. The Modern Sexism Scale (Swim et al., 1995) is an 8-item measure with five reverse-worded items. A maximum likelihood exploratory factor analysis revealed two factors with Eigenvalues above 1.0, but analysis of the scree plot

and the support of theory both back a one-factor solution. The second factor's Eigenvalue of 1.1 and low additional variance explained are not enough evidence to break from theory and suggest a two-factor solution. In the sample fro Study 2 the MSS has a Cronbach's alpha = .794.

Social Dominance Orientation. A maximum likelihood exploratory factor analysis with no rotation reveals two factors with Eigenvalues greater than 1.0. This 16-item scale has eight reverse-coded items and inspection reveals that the second factor has high loading for all reverse-worded items. Because the second factor appears to be an artefact of methodology and is not theoretically derived, the single-factor solution is more appropriate. In the sample from Study 2, the Cronbach's alpha = .909.

CRT-HS Item characteristics

It would be expected that CRT results for a trait that is not socially acceptable (such as aggression or, in the present study, sexism) would follow a positively skewed distribution. Participants could score anywhere from zero to 10 on the CRT-HS. The average score was M = 2.97, SD = 1.60, with a skewness value = .476. The skew of the distribution is in the correct direction (positive) but slight. With only ten questions, low possible variability may make it difficult to determine a positive skew if one existed. Only a small minority of participants (N = 15) gave the sexist answer to six or more questions.

Based on previous CRTs, it would be expected that there is a small subsample of highly sexist individuals. To determine whether there was any evidence for a small subsample of highly sexist participants, those who answered >50% of CRT-HS questions with the sexist response were compared to participants with <=50% responding.

Independent groups t-tests comparing high-sexist with low-sexist responders on Hostile, Benevolent, and Ambivalent sexism are reported in Table 5. As can be seen in Table 5, participants with high-sexist responding on the CRT-HS also have significantly higher rates of hostile (but not benevolent) sexism. While this pattern of results is what was expected based on design and theory, these results should be viewed with skepticism since the comparison groups were created post-hoc with disparate sample sizes.

Validity

      Correlations between the CRT-HS and other measurements. Finally, relationships among the CRT-HS and other measurements were explored. It is imperative that our CRT does not measure actual intellectual ability. The American College Test (ACT) is intended to measure intellectual ability and reasoning skills. LeBreton and colleagues (2007) received permission to obtain the ACT scores of several samples of undergraduates and found, supporting previous findings, no correlation between intelligence as measured by these standardized tests and scores on a conditional reasoning task. In a summary of the results from previous research, the average correlation between a measure of intelligence and a CRT was below .10. There is no reason to suspect that sexism is related to intellectual ability, so a relationship between the CRT-HS and intellectual ability would indicate a confound. Convergent validity must also be established, therefore it is predicted that sexism is related to the CRT-HS. Measurements on the ASI should be highly (but not perfectly) correlated by subscale with measurements on the CRT-HS. It was expected that gender differences found using the CRT-HS would reflect gender differences found using the ASI. Relationships between the CRT-HS and

Table 5

t-tests comparing High-sexist to Low-sexist responders

	t-value	Df	two-tailed p	95% CI
Hostile subscale	-2.819	245	.005	[-7.322, -1.298]
Benevolent subscale	-1.421	245	.156	[-5.407, 0.874]
ASI total	-2.605	243	.010	[-12.016, -1.668]

Neosexism, the Modern Sexism Scale, the Attitudes Toward Women Scale were also investigated.

The CRT-HS should have many overlapping correlates with the Ambivalent Sexism Inventory (Glick & Fiske, 1996). Before a discussion of the correlates between these measurements and the CRT-HS, the relationship between these measurements and the ASI should be looked at in some detail.

Correlates of the ASI. First, the Ambivalent Sexism Inventory itself will be discussed. Following will be descriptions of measurement tools the ASI is conceptually or empirically related to.

Ambivalent Sexism Inventory. Ambivalent Sexism Theory includes three subdimensions for the ASI, capturing three different types of prejudice against women. Factor analysis supports these subdimensions for benevolent sexism but hostile sexism is unidimensional. The three factors underlying benevolent sexism are protective paternalism, the notion that women are the weaker sex and thus need help; complementary gender differentiation, the idea that women's purity or other differences support a complementary view of the sexes that make up for each other's' weaknesses; and heterosexual intimacy, the belief that men need a woman to love to be fulfilled (see Glick et al., 2000 for a detailed example of factor analysis on the ASI; see Glick & Fiske, 1996 for discussion of why these factors are likely to emerge). Items addressing hostile sexism do include elements of paternalism, support for gender differences, and heterosexual bias; in different samples, however, factor analysis does not support the finding that sexist individuals have separate reporting patterns for different subsets of hostile sexism (Glick et al., 2000). For the full questions, see Appendix A.

It is often thought that self-report of consciously held negative traits is unreliably linked to implicit, unconscious measurements of said trait because individuals withhold reporting that they hold traits that they themselves or society views as bad. James and colleagues (2005) argued that "self-reports and projective techniques measure complementary aspects of traits, motives, and need states" (p. 93); self-reported sexism is expected to have a low or moderate relationship with all implicit measures of sexism, but the information garnered is not useless. The relationship between self-acknowledged and implicitly gathered information about sexist tendencies may itself give important clues to how and when sexist behavior manifests. Data were collected on this issue but adding context to this effect will necessarily be explored in future research after the CRT-HS has been successfully validated.

Pearson *r* correlations between all scales, including the CRT-HS, are presented in Table 6. It was expected that the CRT-HS would be moderately correlated with the ASI and follow a similar pattern of correlations to other measures as the ASI. As can be seen in Table 6, this pattern was supported, lending initial evidence that the CRT-HS can be used alongside traditional measurements of sexism.

Intelligence. Next, it was important to determine whether the CRT-HS is related to intelligence. A CRT should measure implicit aspects of personality but not intelligence, so a moderate correlation with a standardized test of intelligence would be a weakness of the CRT-HS. To test this association, participants were asked to report their ACT or SAT scores. For participants that reported they have taken the ACT (n = 247), there was no evidence of a correlation between ACT scores and CRT-HS scores, $r = -.078$, $p = .333$. Only three participants reported their SAT score, so correlations between

SAT and CRT-HS are unreliable estimates of a population value. Hypothesis 3 was supported—according to this evidence, there is no relationship between CRT-HS answers and intelligence.

Study 2 Discussion

Study 2 tested all three hypotheses using an improved version of the CRT-HS. Support for Hypothesis 1 would have been evident by simple, unifactorial factor structure. Support for Hypothesis 2 would have been evident by high correlations between the CRT-HS and sexism measured by several different scales. The factor structure of the CRT-HS in Study 2 did was not simple or unifactorial, meaning Hypothesis 1 was not supported. However, the CRT-HS showed the exact pattern of relationships we would expect if it this type of implicit measure were successful at measuring sexism (supporting Hypothesis 2) but was not related to intelligence (supporting Hypothesis 3).

Table 6

Correlations between all scales included in Study 2

	HSCRT-HS	Hostile	Benev	ASI Total	MSS	ATWS	Neo	SDO
HSCRT-HS	1							
Hostile	.368**	1						
Benevolent	.242**	.406**	1					
ASI Total	.368**	.835**	.841**	1				
MSS Total	.298**	.410**	.238**	.389**	1			
ATWS Total	-.295**	-.507**	-.449**	-.573**	-.463**	1		
Neosexism	-.367**	-.537**	-.417**	-.570**	-.631**	.652**	1	
SDO	-.349**	-.392**	-.391**	-.471**	-.413**	.455**	.547**	1

**p < .01

CHAPTER IV

GENERAL DISCUSSION

At this time, sound research methodologies (or lack thereof) are in the spotlight in psychological science (Open Science Collaboration, 2015). With reinvigorated interest in producing reliable research results, researchers cannot afford to rely on any one measurement technique alone. With an interest in creating a new measurement instrument for an established theory, I turned to the literature on conditional reasoning tasks. James and colleagues (2005) expressed caution and optimism about the creation of new CRTs:

> "…it is possible to contemplate how the conditional reasoning approach might be extended beyond the present tests for aggression and achievement motivation. This system is theoretically generalizable to any behavior that is subject to justification (rationalization) by at least some individuals. Included here are negative traits such as antisocial actions, hidden motives such as when a search for excellence is engendered by an obsessive quest for perfection, and biases in favor of one side of a dialectic, such as when a leader consistently chooses, and justifies, personal decision making over delegation of authority. Others will surely have creative ideas for how conditional reasoning might be used in the field and in research." (James et al. 2005, concluding paragraph)

In this series of investigations, this suggestion is being embraced—it is believed that a CRT can investigate the proclivity for endorsing sexist responses along lines carved by cognitive justification mechanisms. The first milestone in this project was to

pass the theoretical and empirical checkpoints established by James and Lebreton (2010). In an effort to pass this first milestone, several iterations of item generation, item refinement, and data collection were conducted. The final list of ten items was not unidimensional according to the exploratory factor analysis, so hypothesis 1 was not supported. The poor fit of the conditional reasoning items to a single factor is partially explained by the method. Conditional reasoning questions are scored dichotomously and are not explicitly worded survey items. The poor fit of the conditional reasoning items is also partially explained by poor specification of items. The current project incorporated several item specification steps, but due to the fickle responding pattern to logical reasoning questions, the subset of created questions that were empirically viable was small. Future projects would necessarily involve a more intensive item generation and pruning process. One prominent suggestion from James and Lebretion (2010) is hiring a professional logician to create or revise items.

The second milestone would be for the CRT-HS to be a valid reflection of the theory outlined in work by Glick and Fiske on Ambivalent Sexism Theory (1996, 2001). For this checkpoint, the CRT-HS cannot merely replicate results that could be achieved using self-report measures—it must, as suggested by James and colleagues (2005), capture elements that are not discoverable merely through self-report. In addition to this, the CRT-HS must fit into the theoretical framework already established by similar measures for sexism in general. In regards to this checkpoint, hypothesis 2 was fully supported. Scores on the CRT-HS were highly correlated with scores on the ASI, but had a stronger relationship with the Hostile subscale. Additionally, the pattern of relationships between the Hostile subscale of the ASI and other measurements of sexism was

replicated for the CRT-HS. Potentially, the CRT-HS could be used as an alternative methodology that shows similar relationships to sexist attitudes.

Another important aspect of a CRT is that items appear to measure logical reasoning ability but do not in reality measure intelligence. In Study 2, the hypothesis was tested that scores on the CRT-HS would not be related to participants self-reported SAT and/or ACT scores. Hypothesis 3 was supported for the relationship between ACT scores and the CRT-HS. No reliable relationship between ACT scores and the CRT-HS was detected, but future projects should confirm this result by administering logical reasoning tasks in addition to the CRT-HS so that the relationship can be examined without relying on self-report. In the Midwestern United States, the ACT is the preferred option for high school students, so we were unable to collect enough SAT data to reliably judge the relationship between the SAT and CRT-HS.

Limitations

Poor specification and poor items cripple many projects before data are even collected (Podsakoff, MacKenzie, & Podsakoff, 2016). While the current project was successfully at creating a scale with theoretically-backed relationships to other measurements, the scale also failed to pass several statistical checkpoints. At each step of the process in creating a conditional reasoning task, there are additional steps that could be implemented. When generating items, a logician could be hired to generate and review items to begin with a wider and deeper selection of items to begin testing. Better specification of the justification mechanisms that support hostile sexism could unify the conditional reasoning questions better which would result in a simpler factor structure.

The path ahead, for a version of the CRT-HS to be used in research, would include revisions and retesting in different populations. The sample presented in this paper was of college students at a Midwestern University and therefore was narrow in terms of ethnicity, age, and socioeconomic status. Younger generations of Americans tend to be more liberal on issues of gender equality, so it would be expected that an older sample would be have higher endorsement of sexist options. College students are more familiar than the general population with testing procedures since they are subjected to tests to enter and maintain their education; the CRT-HS may not be correlated with sexism in non-students if they are put off by its likeness to tests like the S.A.T.

Conclusions

The goal of this project was to test whether a conditional reasoning task is a viable way to measure sexism, then to take the first steps toward doing so. Based on this work, it can be seen that inn principle hostile sexism can be measured using a conditional reasoning task. With additional validation of the CRT-HS, this measurement instrument would facilitate a wide variety of additions to the literatures on implicit measurement and sexism.

APPENDICES

Appendix A
ASI

Ambivalent Sexism Inventory. Hostile items are marked with H, Benevolent items are marked with B.

1. No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman. (B)

2. Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for "equality." (H)

3. In a disaster, women ought not necessarily to be rescued before men. (B)

4. Most women interpret innocent remarks or acts as being sexist. (H)

5. Women are too easily offended. (H)

6. People are often truly happy in life without being romantically involved with a member of the other sex. (B)

7. Feminists are not seeking for women to have more power than men. (H)

8. Many women have a quality of purity that few men possess. (B)

9. Women should be cherished and protected by men. (B)

10. Most women fail to appreciate fully all that men do for them. (H)

11. Women seek to gain power by getting control over men. (H)

12. Every man ought to have a woman whom he adores. (B)

13. Men are complete without women. (B)

14. Women exaggerate problems they have at work. (H)

15. Once a woman gets a man to commit to her, she usually tries to put him on a tight leash. (H)

16. When women lose to men in a fair competition, they typically complain about being discriminated against. (H)

17. A good woman should be set on a pedestal by her man. (B)

18. There are actually very few women who get a kick out of teasing men by seeming sexually available and then refusing male advances. (H)

19. Women, compared to men, tend to have a superior moral sensibility. (B)

20. Men should be willing to sacrifice their own well-being in order to provide financially for the women in their lives. (B)

21. Feminists are making entirely reasonable demands of men. (H)

22. Women, as compared to men, tend to have a more refined sense of culture and good taste. (B)

Appendix B
AWS

Attitudes Towards Women Scale short version (Spence, Helmreich, & Stapp, 1973).

1. Swearing and obscenity are more repulsive in the speech of a woman than of a man.

2. Women should take increasing responsibility for leadership in solving the intellectual and social problems of the day.

3. Both husband and wife should be allowed the same grounds for divorce.

4. Telling dirty jokes should be mostly a masculine prerogative.

5. Intoxication among women is worse than intoxication among men.

6. Under modern economic conditions with women being active outside the home, men should share in household tasks such as washing dishes and doing the laundry.

7. It is insulting to women to have the "obey" clause remain in the marriage service.

8. There should be a strict merit system in job appointment and promotion without regard to sex.

9. A woman should be as free as a man to propose marriage.

10. Women should worry less about their rights and more about becoming good wives and mothers.

11. Women earning as much as their dates should bear equally the expense when they go out together.

12. Women should assume their rightful place in business and all the professions along with men.

13. A woman should not expect to go to exactly the same places or to have quite the same freedom of action as a man.

14. Sons in a family should be given more encouragement to go to college than daughters.

15. It is ridiculous for a woman to run a locomotive and for a man to darn socks.

16. In general, the father should have greater authority than the mother in the bringing up of children.

17. Women should be encouraged not to become sexually intimate with anyone before marriage, even their fiancés.

18. The husband should not be favored by law over the wife in the disposal of family property or income.

19. Women should be concerned with their duties of childbearing and house tending, rather than with desires for professional and business careers.

20. The intellectual leadership of a community should be largely in the hands of men.

21. Economic and social freedom is worth far more to women than acceptance of the ideal of femininity which has been set up by men.

22. On the average, women should be regarded as less capable of contributing to economic production than are men.

23. There are many jobs in which men should be given preference over women in being hired or promoted.

24. Women should be given equal opportunity with men for apprenticeship in the various trades.

25. The modern girl is entitled to the same freedom from regulation and control that is given to the modern boy.

Modern Sexism Scale (Swim et al., 1995)

1. Discrimination against women is no longer a problem in the United States.

2. Women often miss out on good jobs due to sexual discrimination.

3. It is rare to see women treated in a sexist manner on television.

4. On average, people in our society treat husbands and wives equally.

5. Society has reached the point where women and men have equal opportunities for achievement.

6. It is easy to understand the anger of women's groups in America.

7. It is easy to understand why women's groups are still concerned about societal limitations of women's opportunities.

8. Over the past few years, the government and news media have been showing more concern about the treatment of women than is warranted by women's actual experiences.

Appendix D
Neosexism

Neosexism (Tougas et al., 1995)

1. Discrimination against women in the labor force is no longer a problem in the United States. (originally "in Canada")

2. I consider the present employment system to be unfair to women.[1]

3. Women shouldn't push themselves where they are not wanted.

4. Women will make more progress by being patient and not pushing too hard for change.

5. It is difficult to work for a female boss.

6. Women's requests in terms of equality between the sexes are simply exaggerated.

7. Over the past few years, women have gotten more from government than they deserve.

8. Universities are wrong to admit women in costly programs such as medicine, when in fact, a large number will leave their jobs after a few years to raise their children.

9. In order not to appear sexist, many men are inclined to overcompensate women.

10. Due to social pressures, firms frequently have to hire underqualified women.

11. In a fair employment system, men and women would be considered equal.[1]

12. [1]Reverse coded items.

Social Dominance Orientation (Pratto, Sidanius, Stallworth, & Malle, 1994)

1. Some groups of people are simply inferior to others.

2. In getting what you want, it is sometimes necessary to use force against other groups.

3. It's OK if some groups have more of a chance in life than others.

4. To get ahead in life, it is sometimes necessary to step on other groups.

5. If certain groups stayed in their place, we would have fewer problems.

6. It's probably a good thing that certain groups are at the top and other groups are at the bottom.

7. Inferior groups should stay in their place.

8. Sometimes other groups must be kept in their place.

9. It would be good if groups could be equal.

10. Group equality should be our ideal.

11. All groups should be given an equal chance in life.

12. We should do what we can to equalize conditions for different groups.

13. Increased social equality is beneficial to society.

14. We would have fewer problems if we treated people more equally.

15. We should strive to make incomes as equal as possible.

16. No group should dominate in society.

Conditional Reasoning Task – Hostile Sexism
For each question below, options (a) and (d) are illogical. Option (b) is logically correct but sexist. Option (c) is logically correct and neutral/non-sexist.

1. Corporate hiring policies reflect what human resources specialists believe is best for the corporate environment. Many such companies are, by law, prohibited from discriminating against women or various minority groups. Interview procedures may exclude potentially valuable hires when they display poor social skills. Hiring decisions are sometimes based on likeability, personal charm, or favoritism. Applicants can use their sex, minority status, or social skills to increase their chances of being hired.

Which of the following is the most reasonable conclusion based on the above?

a) Human resources departments' day-to-day decisions are important to the company.

b) Underqualified female or minority applicants may have an advantage due to anti-discrimination laws.

c) Anti-discrimination laws even the playing field among underrepresented groups.

d) Interviewing is an important process for human resources employees.

2. Shirley, a personal stylist, is interested in finding a romantic partner. When Shirley uses online dating sites, she receives many messages that she views as sexist or rude. Her latest date Brad made many comments that Shirley found

offensive and wore an outfit that she found shabby. Brad believed the date went

well and believes he was being polite, but Shirley did not seem interested in

conversation. Shirley did not ask for a second date with Brad.

Which of the following is the most reasonable conclusion based on the

above?

a) All messages on dating sites are vulgar.

b) Women may interpret innocent remarks as sexist or rude.

c) Women use clothing style as a way to judge men.

d) Shirley met her latest date, Brad, on a dating site.

3. A cleaning product company is assessing consumers' reaction to their new

indoor cleaning product. A survey was distributed to a diverse group of

individuals. On average, those that returned the survey were concerned about the

product's safety around pets and small children. The company decided to test a

new ad campaign showing a smiling mother using the product around her home.

A follow-up survey showed that men, but not women, were more likely to

endorse the product after the ad campaign.

Which of the following is the most reasonable conclusion based on the

above?

a) Ad campaigns should avoid the image of a smiling mother.

b) Women who responded to the survey were offended by the ad

campaign.

c) The ad campaign appealed to men concerned about the safety of pets.

d) Cleaning products must be safe for use around pets and small children.

4. Action movies in the 21st century are more likely to have females in lead roles than action movies in the 20th century. Executives in the film industry often refuse projects written by men or women new to the business. New writers are more likely to use female leads, lower budgets, and untested ideas. Many film studios will refuse to produce movies based on high-quality scripts unless they believe the movie will also bring in a profit.

Which of the following is the most reasonable conclusion based on the above?

a) Films in the 21st century have higher budgets than in the past

b) Film executives believe movies with female leads will be less profitable.

c) Film executives discriminate against new, unknown writers.

d) Action movies are likely to have low-quality scripts.

5. Military historians are hotly debating the inclusion of women in combat units. Active combat duty requires high physical strength, endurance, and mental ability. Men outcompete women in tests of physical strength. Usually, men outcompete women in tests of endurance.

Which of the following is the most reasonable conclusion based on the above?

a) Women outcompete men in tests of mental ability.

b) Women are not fit for active combat roles.

c) Military units rely on more than physical strength.

d) Military history is highly controversial.

6. Studies in many animal species show that females will refuse to mate unless males provide them with benefits such as food or shelter. In birds, males often build a nest. In modern human societies, men and women have different roles. Some men and some women are the primary breadwinners in relationships.

Which of the following is the most reasonable conclusion based on the above?

a) Information about animal mating behavior always comes from observation, not inference.

b) Men are, by nature, providers.

c) Human sex differences are less strict than animal sex differences.

d) Men display instincts identical to "nesting."

7. The office recently started using a feedback box where employees can anonymously leave written complaints about office life. Many complaints center on a distracting work environment. These distractions include personal conversations in the workplace, female employees wearing short skirts, and resentment between employees and managers. The office manager wants to create a distraction-free workplace. The manager believes that implementing a new dress

policy and initiating penalties for personal conversations in the workplace will both lead to a distraction-free workplace.

Which of the following most weakens the manager's argument?

a) Managers always try to ignore feedback from employees.

b) The new dress policy may not stop women from wearing revealing clothing.

c) Penalizing personal conversations may lead to more resentment toward managers.

d) Resentment between employees and managers is never resolved.

8. Michael B. and Samantha W. were both running for a City Council position. Michael B. chose to make his campaign primarily about city tax laws. Samantha W. produced campaign ads in opposition to a recent city ordinance that local businesses supported but she claims hurts local communities. Michael B. credits his winning the election to his likeability and his views on local taxes. Samantha W. blames her loss on local business owners spreading bad information about city ordinances.

Which of the following is the most reasonable conclusion based on the above?

a) Personal charm is always important than the issues in local elections.

b) When women lose to men, they complain about unfair practices.

c) Small businesses hold a lot of power in local elections.

d) Taxes are the most important issue in local elections.

9. Nicole is a narcissist. Nicole likes to flirt heavily with men then ignore them afterwards. After Nicole ignores someone, she feels good and in control. Therefore….

Which of the following best logically completes this analogy?

a) Narcissists need to feel in control.

b) Women tease men to feel good.

c) Narcissists flirt to feel in control.

d) Nicole likes to feel narcissistic.

10. When men view pornography more than four times per week, studies indicate they then show less respect to women. Feminist groups have petitioned the government to require a warning to accompany pornographic material. When asked if the government would put a warning on pornographic material, a representative said it would not because the government would lose support if these studies were to be contradicted in the future.

Which of the following most strengthens the representative's argument?

a) Warnings on pornographic material will hurt the adult industry.

b) Feminist groups are making an unreasonable demand of the government.

c) The government has updated its warnings in the past and then lost support.

d) The government should stay out of the personal affairs of citizens.

Conditional Reasoning Task – Hostile Sexism
after revisions based on feedback from Study 1.
For each question below, options (a) and (d) are illogical. Option (b) is logically correct
but sexist. Option (c) is logically correct and neutral/non-sexist.

1. Corporate hiring policies reflect what human resources specialists believe is

best for the corporate environment. Many such companies are, by law, prohibited

from discriminating against women or various minority groups. Interview

procedures may exclude potentially valuable hires when they display poor social

skills. Hiring decisions are sometimes based on likeability, personal charm, or

favoritism. Applicants can use their sex, minority status, or social skills to

increase their chances of being hired.

Which of the following is the most reasonable conclusion based on the

above?

a) Human resources departments' day-to-day decisions are always final.

b) Underqualified female or minority applicants may have an advantage

due to anti-discrimination laws.

c) Anti-discrimination laws even the playing field among

underrepresented groups.

d) Interviewing is not an important part of the hiring process.

2. Shirley, a personal stylist, is interested in finding a romantic partner. When

Shirley uses online dating sites, she receives many messages that she views as

sexist or rude. Her latest date Brad made many comments that Shirley found

offensive and wore an outfit that she found shabby. Brad believed the date went well and believes he was being polite, but Shirley did not seem interested in conversation. Shirley did not ask for a second date with Brad.

a) All messages on dating sites are vulgar.

b) Women may interpret innocent remarks as sexist or rude.

c) Women use clothing style as a way to judge men.

d) Brad was wearing a shabby outfit.

3. A cleaning product company is assessing consumers' reaction to their new indoor cleaning product. A survey was distributed to a diverse group of individuals. On average, those that returned the survey were concerned about the product's safety around pets and small children. The company decided to test a new ad campaign showing a smiling mother using the product around her home. A follow-up survey showed that men, but not women, were more likely to endorse the product after the ad campaign.

Which of the following is the most reasonable conclusion based on the above?

a) Ad campaigns should avoid the image of a smiling mother.

b) Women who responded to the survey were offended by the ad campaign.

c) The ad campaign appealed to men concerned about the safety of pets.

d) Consumers do not care if products are safe for use around children.

4. Action movies in the 21st century are more likely to have females in lead roles than action movies in the 20th century. Executives in the film industry often refuse projects written by men or women new to the business. New writers are more likely to use female leads, lower budgets, and untested ideas. Many film studios will refuse to produce movies based on high-quality scripts unless they believe the movie will also bring in a profit.

Which of the following is the most reasonable conclusion based on the above?

a) Films in the 21st century are more likely to be dramas.

b) Film executives believe movies with female leads will be less profitable.

c) Film executives discriminate against new, unknown writers.

d) New, unknown writers are more likely to request higher budgets.

5. Military historians are hotly debating the inclusion of women in combat units. Active combat duty requires high physical strength, endurance, and mental ability. Men outcompete women in tests of physical strength. Usually, men outcompete women in tests of endurance.

Which of the following is the most reasonable conclusion based on the above?

a) Women outcompete men in tests of mental ability.

b) Women are not fit for active combat roles.

c) Military units rely on more than physical strength.

d) Active duty combat units are highly controversial.

6. Studies in many animal species show that females will refuse to mate unless males provide them with benefits such as food or shelter. In birds, males often build a nest. In modern human societies, men and women have different roles. Some men and some women are the primary breadwinners in relationships.

> Which of the following is the most reasonable conclusion based on the above?
>
> a) Information about animal mating behavior is pure speculation.
>
> b) Men are, by nature, providers.
>
> c) Human sex differences are less strict than animal sex differences.
>
> d) Modern human societies all have the exact same gender roles with no cultural differences.

7. The office recently started using a feedback box where employees can anonymously leave written complaints about office life. Many complaints center on a distracting work environment. These distractions include personal conversations in the workplace, female employees wearing short skirts, and resentment between employees and managers. The office manager wants to create a distraction-free workplace. The manager believes that implementing a new dress policy and initiating penalties for personal conversations in the workplace will both lead to a distraction-free workplace.

> Which of the following most weakens the manager's argument?

a) Managers never get feedback from employees.

b) The new dress policy may not stop women from wearing revealing clothing.

c) Penalizing personal conversations may lead to more resentment toward managers.

d) Resentment between employees and managers is beneficial.

8. Michael B. and Samantha W. were both running for a City Council position. Michael B. chose to make his campaign primarily about city tax laws. Samantha W. produced campaign ads in opposition to a recent city ordinance that may impact small businesses. Michael B. credits his winning the election to his likeability and his views on local taxes. Samantha W. blames her loss on local business owners spreading bad information about city ordinances.

Which of the following is the most reasonable conclusion based on the above?

a) Likeable city councilors are not effective.

b) When women lose competitions, they blame the influence of others.

c) When male politicians win elections, they credit their own charm.

d) Keeping streets safe is the most important issue in local elections.

9. Nicole is a female narcissist. Nicole likes to flirt heavily with men then ignore them afterwards. After Nicole ignores someone, she feels good and in control. Therefore….

Which of the following best logically completes this analogy?

a) Narcissists go to bars to flirt.

b) Women tease men to feel good.

c) Narcissists flirt to feel in control.

d) Nicole likes to feel narcissistic.

10. When men view pornography more than four times per week, studies indicate they then show less respect to women. Feminist groups have petitioned the government to require a warning to accompany pornographic material. When asked if the government would put a warning on pornographic material, a representative said it would not because the government would lose support if these studies were to be contradicted in the future.

Which of the following logically follows from the argument?

a) Warnings on pornographic material will destroy the adult industry.

b) Feminist groups are making an unreasonable demand of the government.

c) The government has updated its warnings in the past and then lost support.

d) The government should invade the personal affairs of citizens.

REFERENCES

Allport, G. W. (1954). ''The Nature of Prejudice''. Addison-Wesley, Reading, MA.

Altemeyer, B. (1981). Right-wing authoritarianism. Winnipeg, Canada: University of Manitoba Press.

Becker, J. C. (2010). Why do women endorse hostile and benevolent sexism? The role of salient female subtypes and personalization of sexist contents. *Sex Roles, 62*, 453–467. doi:10.1007/s11199-009-9707-4

Birnbaum, G. E., Ein-Dor, T., Reis, H. T., & Segal, N. (2015). Why do men prefer nice women over bad ones? The mediating role of gender typicality in the effect of responsiveness on perceived attractiveness in initial acquaintanceships. *Personality and Social Psychology Bulletin. 40(10),* 1341–1353

Blackhart, G. C., Brown, K. E., Clark, T., Pierce, D. L., & Shell, K. (2012). Assessing the adequacy of postexperimental inquiries in deception research and the factors that promote participant honesty. *Behavior Research Methods, 44(*1), 24-40.

Bowler, J. L., Bowler, M. C., & James, L. R. (2011). The cognitive underpinnings of addiction. *Substance Use and Misuse, 46*(8), 1060-1071.

Buss, D. M., Shackelford, T. K., Kirkpatrick, L. A., & Larsen, R. J. (2001). A half century of mate preferences: The cultural evolution of values. *Journal of Marriage and Family*, 63(2), 491-503.

Chen, Z., Fiske, S. T., & Lee, T. L. (2009). Ambivalent sexism and power-related gender-role ideology in marriage. *Sex Roles, 60*(11-12), 765-778.

Douglas, S. J. (2010). Enlightened sexism: The seductive message that feminism's work

    is done. Macmillan.

Edwards, S. R., Bradshaw, K. A., & Hinsz, V. B. (2014). Denying rape but endorsing

    forceful intercourse: Exploring differences among responders. *Violence and*

    Gender, 1(4), 188-193.

Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis) respecting versus (dis) liking:

    Status and interdependence predict ambivalent stereotypes of competence and

    warmth. *Journal of Social Issues, 55*(3), 473-489.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed)

    stereotype content: competence and warmth respectively follow from perceived

    status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-

    902.

Fredrickson, B. L., Roberts, T. A., Noll, S. M., Quinn, D. M., & Twenge, J. M. (1998).

    That swimsuit becomes you: sex differences in self-objectification, restrained

    eating, and math performance. *Journal of Personality and Social Psychology,*

    75(1), 269-284.

Furnham, A. (2003). Belief in a just world: Research progress over the past decade.

    Personality and Individual Differences, 34(5), 795-817.

Garcia-Retamero, R., & López-Zafra, E. (2006). Prejudice against women in male-

    congenial environments: Perceptions of gender role congruity in leadership. *Sex*

    Roles, 55(1-2), 51-61.

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, Adetoun, B., Soagie, J., Akande, A., Alao, A., Brunner, A., Willemsen, T., Chipeta, K., dardenne, B., Dijksterhuis, A. Wigboldus, D., Eckes, T., Sit-Materna, I., Exposito, F., Moya, M., Foddy, M., Kim, H., Lameiras, M., Sotelo, M.J., Mucchi-Faina, A., Romani, M., Sakalh, N., Udegbe, B., Yamamoto, M., Ui, M., Ferreira, M. & López, W. L. (2000). Beyond prejudice as simple antipathy: hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology, 79*(5), 763-775.

Glick, P., and Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*, 491–512.

Glick, P., & Fiske, S. T. (1997). Hostile and benevolent sexism measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly, 21*(1), 119-135.

Glick, P., and Fiske, S. T. (2001). Ambivalent sexism. In ''Advances in Experimental Social Psychology'' (M. P. Zanna, ed.), Vol. 33, pp. 115–188. Academic Press, Thousand Oaks, CA.

Glick, P., Wilkerson, M., & Cuffe, M. (2015). Masculine identity, Ambivalent Sexism, and attitudes toward gender subtypes. *Social Psychology*, Advance online publication. http://dx.doi.org/10.1027/1864-9335/a000228

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464-1480.

Grubbs, J. B., Exline, J. J., & Twenge, J. M. (2014). Psychological Entitlement and

      Ambivalent Sexism: Understanding the Role of Entitlement in Predicting Two

      Forms of Sexism. *Sex Roles, 70*(5-6), 209-220.

Gulliksen, H. (1950). Theory of mental tests. New York: John Wiley.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to

      moral judgment. *Psychological Review, 108*(4), 814.

Hammond, M. D., Sibley, C. G., & Overall, N. C. (2013). The allure of sexism:

      Psychological entitlement fosters women's endorsement of benevolent sexism

      over time. *Social Psychological and Personality Science*, *5*(4), 422-429.

Heflick, N. A., Goldenberg, J. L., Cooper, D. P., & Puvia, E. (2011). From women to

      objects: Appearance focus, target gender, and perceptions of warmth, morality

      and competence. *Journal of Experimental Social Psychology, 47*(3), 572-581.

Heilman, M. E., & Stopeck, M. H. (1985). Being attractive, advantage or disadvantage?

      Performance-based evaluations and recommended personnel actions as a function

      of appearance, sex, and job type. *Organizational Behavior and Human Decision

      Processes*, *35*(2), 202-215.

Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for

      success: reactions to women who succeed at male gender-typed tasks. *Journal of

      Applied Psychology, 89*(3), 416.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?.

      Behavioral and brain sciences, 33(2-3), 61-83.

Holgado–Tello, F. P., Chacón–Moscoso, S., Barbero–García, I., & Vila–Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity, 44*(1), 153-166.

Jackman, M. R. (1994). ''The Velvet Glove: Paternalism and Conflict in Gender, Class, and Race Relations.'' University of California Press, Berkeley.

James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., Frost, B., Russell, S., Sablynski, C., Mitchell, T. & Williams, L. J. (2005). A conditional reasoning measure for aggression. *Organizational Research Methods, 8*(1), 69-99.

James, L. R., & LeBreton, J. M. (2010). Assessing aggression using conditional reasoning. *Current Directions in Psychological Science, 19*(1), 30-35.

James, L. R., & LeBreton, J. M. (2012). Assessing the implicit personality through conditional reasoning. American Psychological Association.

Jost, J. T., and Banaji, M. R. (1994). The role of stereotyping in system justification and the production of false consciousness. *British Journal of Social Psychology, 22*, 1–27.

Jost, J. T., and Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology. 88*, 498–509.

Kay, A. C., and Jost, J. T. (2003). Complementary justice: Effects of ''poor but happy'' and ''poor but honest'' stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology, 85*, 823–837.

Kenrick, D., Neuberg, S., Cialdini, R. (2016). *Social Psychology: Goals in interaction* (6th ed). Pearson.

Korpi, W., Ferrarini, T., & Englund, S. (2013). Women's opportunities under different family policy constellations: gender, class, and inequality tradeoffs in western countries re-examined. *Social Politics: International Studies in Gender, State & Society*, *20*(1), 1-40.

Kristiansen, C. M., & Giulietti, R. (1990). Perceptions of wife abuse. *Psychology of Women Quarterly, 14*(2), 177-189.

Lerner, M. J., & Simmons, C. H. (1966). Observer's reaction to the" innocent victim": compassion or rejection*?. Journal of Personality and social Psychology, 4*(2), 203.

Lorenzo-Seva, U., & Ferrando, P. J. (2015). POLYMAT-C: a comprehensive SPSS program for computing the polychoric correlation matrix. *Behavior Research Methods*, 47(3), 884-889.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*(3), 683-729.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84*(3), 231.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down-and up-regulation of negative emotion. *Neuroimage*, 23(2), 483-499.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological

science. *Science,* 349, 943-952. DOI: 10.1126/science.aac4716

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2016). Recommendations for

Creating Better Concept Definitions in the Organizational, Behavioral, and Social

Sciences. *Organizational Research Methods*, 1094428115624965.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social Dominance

Orientation: A personality variable predicting social and political attitudes.

Journal of Personality and Social Psychology, 67, 741-763.

Sibley, C. G., & Wilson, M. S. (2004). Differentiating hostile and benevolent sexist

attitudes toward positive and negative sexual female subtypes. *Sex Roles*, *51*(11-

12), 687-696.

Spears, R., and Manstead, A. S. (1989). The social context of stereotyping and

differentiation. *European Journal of Social Psychology. 19*, 101–121.

Spence, J. T., Helmreich, R., & Stapp, J. (1973). A short version of the Attitudes Toward

Women Scale (AWS). *Bulletin of the Psychonomic Society, 2*(4), 219-220.

Streb, M. J., Burrell, B., Frederick, B., & Genovese, M. A. (2008). Social desirability

effects and support for a female American president. *Public Opinion

Quarterly*, *72*(1), 76-89.

Tougas, F., Brown, R., Beaton, A. M., & Joly, S. (1995). Neosexism: Plus ca change,

plus c'est pareil. *Personality and Social Psychology Bulletin, 21*(8), 842-849.

Twenge, J. M. (1997). Attitudes Toward Women, 1970–1995 A Meta-Analysis.

Psychology of Women Quarterly, 21(1), 35-51.

Williams, L. J., Hartman, N., & Cavazotte, F. (2010). Method variance and marker

variables: A review and comprehensive CFA marker technique. *Organizational*

Research Methods, *13*(3), 477-514.