

Running Title: LABORATORY MEASURE OF SEXUAL HARASSMENT

A Refined Computer Harassment Paradigm:

Validation, and Test of Hypotheses about Target Characteristics

Frank Siebler, Saskia Sabelus, and Gerd Bohner

Universität Bielefeld

This is the final author version of an article published in *Psychology of Women Quarterly* (The Society for the Psychology of Women, Division 35 of the American Psychological Association; Blackwell Publishing). The definitive version is available at [www.blackwell-synergy.com](http://www.blackwell-synergy.com).

**Bibliographic details:** Siebler, F., Sabelus, S., & Bohner, G. (2008). A refined computer harassment paradigm: Validation, and test of hypotheses about target characteristics. *Psychology of Women Quarterly*, 32, 22-35.

**Link:** <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1471-6402.2007.00404.x>

Author Note

Frank Siebler, Saskia Sabelus, and Gerd Bohner, Department of Psychology, University of Bielefeld, Bielefeld, Germany. Frank Siebler is now at the Department of Psychology, University of Tromsø, Tromsø, Norway. Saskia Sabelus is now at the Department of Psychology, Goldsmiths College, University of London, London, United Kingdom. The reported research was funded by a Deutsche Forschungsgemeinschaft grant (BO 1248/5-1) to Gerd Bohner. Study 2 was conducted by Saskia Sabelus as part of her diploma thesis at the University of Bielefeld under the supervision of Gerd Bohner and Frank Siebler. Preliminary reports of these data were presented at the 18th Annual Conference of the Feminist and Women's Studies Association (UK & Ireland) on "Gender and Violence", Aberdeen, UK, 9-11 September 2005, and the 48th "Tagung experimentell arbeitender Psychologen", Mainz, Germany, 26-29 March 2006. We thank Anne Maass for helpful advice and generously sharing the original stimulus materials of the Italian Likelihood to Sexually Harass Scale; Rita Boeing, Antje Huber, Stephanie Maibaum, Ute Michaelis, and Xenia Polotzek for their help in data collection; as well as Marc Becher and Friederike Eyssel for their helpful comments on a previous draft. Address correspondence and reprint requests to: Frank Siebler, Department of Psychology, University of Tromsø, 9037 Tromsø, Norway. E-mail: [frank@psyk.uit.no](mailto:frank@psyk.uit.no)

## Abstract

A refined computer paradigm for assessing sexual harassment is presented, validated, and used for testing substantive hypotheses. Male participants were given an opportunity to send sexist jokes to a computer-simulated female chat partner. In Study 1 ( $N = 44$ ), the harassment measure (number of sexist jokes sent) correlated positively with self-reported harassment proclivity.

Study 2 ( $N = 77$ ) included a more elaborate cover story, variations of the female target's attitude (feminist vs. traditional) and physical attractiveness (low vs. high), and additional measures for construct validation. Results showed that harassment correlated positively with self-reported harassment proclivity, hostile sexism, and male identity. Feminist targets were harassed more than traditional targets, whereas target attractiveness had no effect. Theoretical and applied implications are discussed.

**Keywords:** sexual harassment, computerized measurement, intergroup behavior, feminist attitude, physical attractiveness

## A Refined Computer Harassment Paradigm:

## Validation, and Test of Hypotheses about Target Characteristics

Sexual harassment affects people at work, in academic settings, in the military, and in informal social contexts (Paludi & Paludi, 2003). In recent decades, sexual harassment over the Internet and other channels of telecommunication (e.g., Sczesny & Stahlberg, 2000) has become increasingly common (for a review, see Barak, 2005). In the present paper we present a computerized paradigm simulating such online harassment, to be used for studying sexual harassment behavior in the laboratory. Although both men and women can experience sexual harassment, the majority of incidents involve female victims and male perpetrators (Pryor & Fitzgerald, 2003; Schuster, Sczesny, & Stahlberg, 1999; Tangri, Burt, & Johnson, 1982). Our research therefore focused on the “male perpetrator – female target” constellation; the proposed paradigm, however, can be adapted to other gender constellations.

*Sexual Harassment*

Definitions of sexual harassment encompass various behaviors and emphasize different aspects. Some definitions include particular locations (e.g., the workplace) or victim and perpetrator characteristics (e.g., assuming that only women can be victims of sexual harassment), and some even suggest that sexist behavior in general should be labeled sexual harassment (for a review, see O’Donohue, Downs, & Yeater, 1998). It may be useful to emphasize psychological rather than legal definitions of harassment (see, e.g., Fitzgerald, Swan, & Magley, 1997; Ilies, Hauserman, Schwochau, & Stibal, 2003), because many types of behavior do not rise to a legal definition but have damaging consequences nonetheless. Most definitions focus on the victim’s perspective. Accordingly, a behavior is described as sexually harassing if it is unwanted by the target of the behavior, and if the target communicates this perception to the perpetrator. German law, for example, defines as sexual harassment a variety of sexually motivated behaviors which are “recognizably rejected by the person affected” (see Zippel, 2006). In the present paper, we

adopted this widely accepted feature of communicated disapproval by implementing it in our own operational definition of harassing behavior.

A literature review yields mainly three forms of sexually harassing behavior, which vary in frequency and severity (Brannon, 2002; Fitzgerald et al., 1997). *Quid pro quo* harassment, the most serious form, includes behaviors such as demands for sexual favors in return for employment, promotion or simply to keep one's job. It usually involves a power differential, e.g., employers or teachers abuse their power by blackmailing or threatening their employees or students. The second form, *hostile environment*, includes behaviors that can interfere with a person's job performance. Typical behaviors in this category are sexual touching, sexualized remarks, or displaying pornographic materials in a place that the victim cannot avoid (Brannon, 2002; Fitzgerald et al., 1997).<sup>1</sup> The third form, *gender harassment*, is not primarily about sexuality. Behaviors falling into this category include degrading remarks, sexist jokes and statements, and are aimed at insulting women (or men) as a group. This form of sexual harassment was the target behavior we implemented in our studies, where sexually harassing behavior was defined as repeatedly sending sexist jokes to a (computer-simulated) female chat partner who consistently communicated her disapproval of this behavior.

Although reported prevalence rates for sexual harassment vary enormously (see Ilies et al., 2003), the results of a survey conducted by the United States Merit Systems Protection Board (1995) confirmed earlier findings that women experience sexual harassment more frequently than do men (44% vs. 19%). The most common behaviors were "unwanted sexual teasing, remarks, or questions" and "unwanted sexually suggestive looks or gestures" (Pryor & Fitzgerald, 2003; p. 82). Interestingly, sexual harassment occurred more often between co-workers than between superior and subordinate. Our present research focuses on this most common situation of equal institutional power between harasser and target, as we investigated the behavior of male students toward a female target who was introduced as a fellow student.

*Individual Differences in the Likelihood to Sexually Harass*

Whereas earlier models of sexual harassment each focused on one class of explanatory constructs (e.g., biological or socio-cultural variables; see Tangri et al., 1982, for a review), more recent theories take a multi-dimensional approach. According to one prominent theory, the person x situation model (Pryor, LaVite, & Stoller, 1993), some men possess a chronic predisposition to harass; however, these men will show harassment only when situational factors *allow* them to do so (e.g., Pryor, 1987). More recent research has shown, furthermore, that men with a proclivity to harass will show harassing behavior to a greater extent if features of the situation (e.g., the salience of male-female competition) or characteristics of potential targets (e.g., a feminist attitude) *motivate* them to do so, for example by challenging their male identity (Dall'Ara & Maass, 1999; Maass, Cadinu, Guarnieri, & Grasselli, 2003).

Pryor (1987) developed a questionnaire measure of harassment proclivity: the Likelihood to Sexually Harass (LSH) scale. It consists of hypothetical scenarios in which men are described who hold powerful positions (e.g. manager) and have the opportunity to take advantage of an attractive female subordinate. Respondents are asked to place themselves in the role of the male protagonist and to rate how likely they would be to engage in particular behaviors, some of which constitute quid pro quo harassment. The LSH scale has been shown to be a reliable predictor of sexually harassing behavior. This is true for behavior in laboratory tasks (Dall'Ara & Maass, 1999; Maass et al., 2003; Pryor, 1987), self-reported behavior in workplace situations (Barak & Kaplan, 1996 – as cited in Driscoll, Kelly, & Henderson, 1998), and behavior unobtrusively recorded in a staged “waiting room” situation where participants were unaware of being observed (Schmidt, Gerger, Kley, Siebler, & Bohner, 2003). There is also ample evidence for the person x situation (interaction) model provided by studies using this scale (e.g., Pryor, 1987; see Pryor & Fitzgerald, 2003, for a review).

Interestingly, although LSH scores are based on items reflecting quid pro quo

harassment, they predict behaviors that fall into other categories of sexual harassment, such as unwanted sexual attention (Pryor, 1987; Schmidt et al., 2003) and gender harassment (Dall'Ara & Maass, 1999; Maass et al., 2003). This suggests that the different types of harassment that have been identified in descriptive surveys represent manifestations of a unitary behavioral disposition. Whether a man possessing this disposition will show harassing behavior at all, and if so, which type of behavior, would then depend on situational constraints or opportunities. We used the LSH scale in both studies reported here as a predictor of gender harassment.

#### *Assessing Sexually Harassing Behavior in the Laboratory*

Studying sexual harassment in vivo has theoretical advantages over survey methodologies, because it avoids reporting biases and enables researchers to introduce experimental manipulations. It is limited, however, by ethical constraints that proscribe exposing unsuspecting research participants to harassing behavior (but see Woodzicka & LaFrance, 2005). Therefore, previous studies have featured interactions between male participants and female confederates (see, e.g., Pryor, 1987). Although this procedure reduces the number of females who are subjected to potential harassment and enables full prior consent, it typically increases the amount of unpleasant interactions that a target (i.e., the female confederate) is exposed to.

To overcome these problems, Maass and her colleagues (Dall'Ara & Maass, 1999; Maass et al., 2003) introduced a paradigm that allows the experimenter to study sexual harassment in the laboratory in a highly realistic setting. The purpose of this “computer harassment paradigm” is “to simulate a prototypical form of sexual harassment without actually exposing female participants (or collaborators ...) to sexual harassment, which may be a rather unpleasant experience and, hence, ethically problematic” (Maass et al., 2003, p. 856). The procedure unfolds as follows: A male participant is made to believe that he participates in a computer chat in which his task is to exchange images with a (fictitious) female chat partner. In each trial, participants had the option of choosing images from several folders, which were labeled with

different category names (e.g., “nature”, “animals”). One critical folder was labeled “porno” and contained pornographic images. In Dall’Ara and Maass’ (1999) study, a male confederate tried to persuade the participant to send these pornographic images to the female chat partner. It was assessed whether the participant followed these suggestions and, if so, how many persuasion attempts were needed. Maass et al. (2003) modified this procedure by either using a virtual confederate or completely omitting this collaborator.

The advantages of the computer harassment paradigm are obvious. First, it represents an experimental method that allows researchers to manipulate various theoretically interesting variables (e.g., situational norms, victim characteristics, presence vs. absence of role models etc.) and thus to examine causal factors that may affect harassment. Second, sexually harassing behavior can be directly measured within ethical limits.

#### *The Present Research: Refining the Paradigm*

The main purpose of our research was to refine the computer harassment paradigm by (a) using less blatant stimulus materials, and (b) removing cues to the offensive nature of some of the materials from the experimental situation. Specifically, we replaced pornographic images with sexist jokes as the material to be sent via the “computer chat line”. Sexist jokes (here always targeting women) fall in the category of gender harassment, which is the most frequent form of sexual harassment. Mitchell, Hirschman, Angelone, and Lilly (2004) adopted a joke-telling paradigm to study peer sexual harassment in the laboratory. They asked male participants to select five jokes from a list of fifteen jokes, and to tell the selected jokes to a female confederate. The list comprised three categories of jokes: clean, gross, and sexist. Perhaps the most striking result was the high rate of participants who told one or more sexist jokes (80%). In contrast, in both experiments by Maass et al. (2003), a substantial number of participants did not send a single pornographic picture at all in five critical trials. In line with these findings, we assumed that the general threshold of mailing moderately sexist jokes would be lower than that

of posting patently pornographic pictures.

By arranging images into labeled folders (one of them called “porno”), Maass and her colleagues (2003) provided participants with a highly salient cue to the potentially undesired and harassing nature of some of the images. This may not be a valid model of typical decision situations in sexual harassment. Mitchell and colleagues’ (2004) paradigm omits such cues by using a simple, apparently unstructured list of jokes for participants to select from, thus increasing the ecological validity of the behavioral measure. Taking this approach one step further, we adopted a dual-choice paradigm in which participants were repeatedly asked to select one joke out of a pair. The jokes in critical pairs were matched for funniness, but one of them was sexist whereas the other was not (see Method section for detail).

We developed a fully computerized script where the female target’s responses (which communicated her disapproval of sexist material sent) were pre-programmed. This design feature obviated the need of using a female confederate as target. Also, compared to procedures where live interactions between a participant and a confederate need to be staged, the precision of measurement would be increased by avoiding chance variation in the confederate’s behavior.

To sum up, the refined computer harassment paradigm was developed as a research instrument for the unobtrusive and ethical assessment of sexually harassing behavior. The paradigm’s task (choosing one out of two jokes) was designed such that a participant could justify any choice by apparent task demands. Thus, in terms of a person x situation model of sexual harassment, the experimental situation was designed to constantly *allow* for sexual harassment. In two studies we investigated conditions that may motivate men to actually *use* this behavioral option.

### Study 1

In our first study, we examined the viability of the refined computer harassment paradigm and collected initial data on its validity. The study was also aimed at further validating a German



version of Pryor's (1987) LSH scale (Schmidt et al., 2003). We hypothesized that male students' LSH scores would be correlated positively with harassing behavior. To explore potential effects of impression management on the measures used, we manipulated participants' outcome intentions. Specifically, participants were instructed either to answer honestly, or to try to make a favorable impression (for a similar procedure, see Fiedler & Bluemke, 2005). If impression management concerns affect participants' responses, then (a) both the LSH scores and the harassment scores should be lower in favorable impression conditions than in honest answer conditions, and (b) the correlation between LSH scores and harassment scores might be lower in the favorable impression conditions than in the honest answer conditions. In terms of the measures' construct validity, of course, neither effect would be desirable.

### *Method*

#### *Participants*

Forty-four male participants volunteered for a study allegedly dealing with the development of new research instruments in occupational psychology. Their mean age was 25.3 years ( $SD = 7.15$ ). Specifically, men who appeared to be students were individually contacted by research assistants on the campus of the University of Bielefeld, for instance in waiting rooms and cafeterias. Three participants did not reveal their course of study in the experimental materials, whereas the others reported a variety of courses including law, business administration, and pedagogics. Volunteers received 2 EUR (approx. 2.44 US \$) for their participation. The laboratory floor section was prepared such that on their way to the assigned room, participants passed several cubicles that were apparently occupied, as indicated by closed doors and lit "busy" signs. This was done to lend credibility to the subsequent cover story whereby other persons were participating simultaneously (see below).

#### *Materials*

*Selection of sexist and non-sexist jokes.* In a pilot study, twelve male participants

indicated for each of 23 sexist and nonsexist jokes (a) how sexually harassing the joke would be for a female recipient (scale from 0, *not at all harassing*, to 5, *very harassing*), and (b) how funny they personally found the joke (scale from 0, *not at all funny*, to 5, *very funny*). Eight joke pairs were formed based on these pilot ratings such that a non-sexist joke was matched with an equally funny sexist joke. For the selected jokes, averaged funniness ratings did not differ between non-sexist jokes ( $M = 2.16$ ,  $SD = .79$ ) and sexist jokes ( $M = 2.19$ ,  $SD = .80$ ),  $t < 1$ , *ns*. In contrast, as intended, the sexist jokes were rated as clearly more harassing ( $M = 2.18$ ;  $SD = 1.38$ ) than the non-sexist jokes ( $M = 0.14$ ;  $SD = .20$ ),  $t(11) = 5.26$ ,  $p < .001$ . Overall then, both jokes in the critical pairs were rated as equally funny; importantly, however, only the sexist jokes were rated as moderately harassing as well, whereas the non-sexist jokes were rated as not harassing at all. For example, one of the critical joke pairs read: "What do you hear when you hold a Döner Kebab<sup>2</sup> to your ear? – The silence of the lambs" (non-harassing) / "Why do women not need an umbrella? – Because it doesn't rain between kitchen and bedroom" (harassing).

*Likelihood to Sexually Harass scale.* The German Likelihood to Sexually Harass scale (Schmidt et al., 2003) assesses men's proclivity to sexually harass. It is based on a modified version of the LSH scale (Pryor, 1987) that was first introduced in Italian by Dall'Ara and Maass (1999; see also Maass et al., 2003). A cover story describes the purpose of the questionnaire as assessing decision making in work situations. After completing several filler items addressing previous work experiences, the participant reads and responds to nine scenarios where a male person in a work context is described in the second person; participants are instructed to imagine that they are the person described. Five scenarios are fillers without any relation to sexual harassment, and four are critical scenarios that include behavioral response alternatives representing quid pro quo harassment. These are direct translations of scenarios from Pryor's original scale. For each scenario, participants complete three items indicating the likelihood that they would engage in a certain behavioral option, along a scale from 1, *not at all likely*, to 7, *very*

*likely*. In each critical scenario, one item represents a strong form of sexual harassment, a second item represents a milder form of harassment, and the third item describes a neutral behavior (for an example, see the Appendix). Mean scores of the two harassment items across the four critical scenarios constitute a participant's LSH score.

*Computer chat paradigm.* After starting the program that controlled the computer chat, the participant was asked to enter his first name. Then, he was made to believe that the program tried to find a peer in a network of ten computers. Having passed apparently occupied cubicles on the way to his own, each participant could plausibly infer that other persons were participating in experiments at the same time. A screen display suggested that one of the networked computers that "responded" was occupied by a person named "Karin" (who was unequivocally identifiable as a woman by this first name). After ostensibly connecting to "Karin's" computer, the program presented the screen mask that was used for the rest of the task. In its top half, the mask comprised two blank boxes with the labels "[XXX] sends:" and "KARIN's vote:", respectively. "[XXX]" was replaced by the participant's first name. The bottom half of the mask was labeled "Selection area (visible for sender)". It featured two blank boxes; each of these was accompanied by a push-button with the caption "Send this one".

Instructions stated that the participant's task was to select and send one of two jokes to his partner. The partner would evaluate the selected joke and would send zero, one, or two credits in return. Participants were encouraged to select jokes such that they would collect as many credits as possible. A permanently visible label provided a verbal interpretation (0 credits: *not funny*, 1 credit: *quite nice*, 2 credits: *really funny*). Then, a matched pair of jokes was shown in the boxes of the mask's selection area. When the participant selected one of them by clicking on the corresponding "Send this one" button, that joke was shown in the box labeled "[XXX] sends". After a variable delay, the preprogrammed feedback (i.e., "Karin's" alleged rating of the funniness of the joke) was shown in the box labeled "KARIN's vote". For non-sexist jokes,

participants received one credit if the joke was low in funniness according to pilot results, and two credits if the joke was high in funniness. For sexist jokes, participants always received zero credits, independent of the joke's funniness rating in the pilot test. Thus, whereas "Karin's" feedback reflected the actual quality of non-sexist jokes, she consistently disapproved of sexist jokes. Five seconds after the feedback, the next pair of jokes was shown. The total number of sexist jokes selected (possible range: 0 to 8) served as our measure of harassing behavior.

One might argue that the first sexist joke sent does not constitute harassment because the target has not yet had a chance to communicate her disapproval. However, because pilot testing had shown that all sexist jokes were seen as moderately harassing, we considered it appropriate to include the first sexist joke sent in our behavioral index. The same scoring strategy has been used in the literature (e.g., Dall'Ara & Maass, 1999, Maass et al., 2003; Mitchell et al., 2004). Nonetheless, we conducted additional analyses with a behavioral index that did not include the first sexist joke sent (see below).

Importantly, by ensuring that the chat partner never rewarded a sexist joke, we implemented a highly conservative test of our main hypothesis. This feature of the design rules out the possibility that participants would select sexist jokes in order to maximize credits. On the contrary, if a participant selected sexist jokes under the given circumstances, he would do so *in spite of* their providing zero rewards.

### *Procedure*

On arrival, participants learned that the study's purpose was to test several new instruments in occupational psychology. The last sentence of the written instructions asked participants either to "respond honestly" in the experimental tasks, or to present themselves "in a good light". To make sure that participants would not miss this crucial part of the instructions, it was verbally repeated by the experimenter.

The study consisted of two separate parts that were each presented on a computer screen. The first part comprised the LSH scale as well as other instruments that we piloted in this study

(these will not be discussed further). The second part comprised the computer chat paradigm. Both parts were controlled by computer programs that were completed in a self-administered fashion. Once participants had completed all parts of the study, they were thoroughly debriefed, thanked, and dismissed.

## *Results*

### *Likelihood to Sexually Harass*

Internal consistency of the LSH scale was satisfactory, Cronbach's  $\alpha = .74$ . Scores were averaged to form an overall index of self-reported harassment proclivity ( $M = 2.49$ ,  $SD = 0.93$ , range 1.13 to 4.75). Importantly, the mean of LSH scores was not significantly affected by instructions (favorable impressions condition:  $M = 2.39$ ,  $SD = 0.78$ ; honest answers condition:  $M = 2.60$ ,  $SD = 1.07$ ),  $t < 1$ ,  $p = .47$ .

### *Behavioral Measure of Sexual Harassment*

The number of sexist jokes sent was used as a summary score of harassing behavior. Each participant sent at least one sexist joke, and 70.5% of participants sent more than one sexist joke ( $M = 2.27$ ;  $SD = 1.07$ ). The means for the behavioral measure did not differ between the favorable impressions condition ( $M = 2.23$ ,  $SD = 1.11$ ) and the honest answers condition ( $M = 2.32$ ,  $SD = 1.04$ ),  $t < 1$ ,  $p = .78$ .

### *Relationship between LSH and Harassing Behavior*

We conducted a hierarchical multiple regression analysis where LSH score, instruction condition, and the interaction of these two variables were used as predictors of harassing behavior. Before creating the interaction term, LSH score and instruction condition were centered (cf. Cohen, Cohen, West, & Aiken, 2003). In step 1, we simultaneously entered the centered LSH score and the instruction condition (coded  $-.5$ , *honest answers*, and  $.5$ , *favorable impression*) as predictors. In step 2, we entered the multiplicative product of the predictors from step 1. The results revealed that LSH significantly predicted the number of sexist jokes sent,  $\beta = .34$ ,  $p = .03$  (step 1). As was expected, the higher participants' LSH score, the higher was the

number of sexist jokes sent. Instruction condition, in contrast, did not predict that number, either alone ( $\beta = -.005$ , *ns*, step 1) or in interaction with LSH ( $\beta = .006$ , *ns*, step 2). Thus, importantly, instructions to answer honestly versus to make a favorable impression did not moderate the relation between LSH scores and harassing behavior. Repeating the analysis with a behavioral index that did not include a participant's first sexist joke sent yielded identical results.

### *Discussion*

The results of the first study show that we had successfully created a modified version of the computer harassment paradigm that used sexist jokes as harassing material. We also found that sexually harassing behaviors, as assessed by the new measure, could be significantly predicted from scores on the modified Likelihood to Sexually Harass scale, which speaks to the external validity of the behavioral measure. Importantly, neither the means nor the intercorrelation of LSH and the number of sexist jokes sent were significantly affected by instructions to answer honestly versus to present oneself in a positive light. This attests to the discriminant validity and contextual robustness of both the modified LSH scale and our behavioral measure of harassment.

### Study 2

The main purposes of Study 2 were to refine our behavioral measure of harassment and to use it for testing substantive hypotheses about the impact of (a) individual difference variables on the part of potential perpetrators, and (b) characteristics of potential targets on sexual harassment. The target characteristics we studied were potential targets' gender-role related attitudes (feminist vs. traditional) and physical attractiveness (high vs. low). The individual-difference variables we studied were potential perpetrators' LSH, identification with the male gender, and hostile sexism.

One potential concern regarding the procedure of Study 1 may be that the LSH scale was administered prior to the computer chat task (although separated by other materials). Completing the scale may thus have influenced participants' harassing behavior. Other research indeed

suggests that the correlation between two concepts may be strengthened if the concept that is assumed to causally affect the other is assessed first rather than last (Bohner, Jarvis, Eyssel & Siebler, 2005; Bohner, Reinhard, Rutz, Sturm, Kerschbaum & Effler, 1998; Schwarz & Strack, 1981). Taking these considerations into account, the order of assessment of LSH and the behavioral measure was reversed in our second study.

Finally, to provide a conceptual replication of the previous study with different materials, we changed various aspects of the experiment. Specifically, we (a) embedded the computer harassment paradigm in a more elaborate cover story, (b) adopted a new set of sexist and non-sexist jokes, and (c) had the computer-simulated “female partner” communicate her disapproval of sexist jokes in natural language.

#### *Targets’ Feminist vs. Traditional Attitudes*

Previous studies have shown that the perceived gender-role related attitude of a potential target influenced the extent to which men engaged in sexual harassment (Dall’Ara & Maass, 1999; Maass et al., 2003). According to social identity theory, a positive self-concept originates from a person’s membership in valued in-groups, including their gender group (see, e.g., Bohner & Sturm, 1997). Maass and her colleagues argued that if men’s gender-related social identity was threatened they would derogate or punish the out-group (i.e., females) in order to protect or re-establish their threatened social identity. One type of threat that was manipulated by Maass and her colleagues was threat to the legitimacy of pro-male status differences. According to Maass et al., such status differences are increasingly challenged in contemporary society, e.g. by women competing for jobs and positions that were traditionally male-dominated. As a consequence, threatened males (the high-status group) try to “defend their privileged status through out-group derogation, including sexual harassment” (Maass et al., 2003, p.855).

In Maass et al.’s (2003) study, social identity threat was induced by manipulating self-descriptions of the fictitious female chat partner: The target expressed either a clearly feminist

attitude (e.g., by aiming to become a manager and supporting a union that defends women's rights) or a traditional gender-role attitude (e.g., by emphasizing her concern for family and children). Maass et al. found that men interacting with a feminist chat partner were more likely to send pornographic images than men interacting with a traditional partner (see also Dall'Ara & Maass, 1999). We set out to replicate this finding using our refined computer harassment paradigm. Specifically, with respect to target attitude, we tested

*Hypothesis 1:* Men send more sexist jokes to a woman who is described as holding feminist attitudes than to a woman who is described as holding traditional attitudes.

#### *Target Attractiveness*

We were further interested in potential effects of the target's physical attractiveness. Investigating this characteristic might yield insights regarding the motivations underlying sexual harassment. Gutek (1985) reports that men were less likely than women to perceive and label particular sexual behaviors in the workplace (e.g., sexual touching, complimentary looks) as harassing. It has therefore been suggested that some forms of sexual harassment might be interpreted as "miscommunication" (Brannon, 2002), which could be motivated by sexual attraction (O'Hare & O'Donohue, 1998). Since men generally prefer physically attractive females as sexual partners (Buss & Barnes, 1986), one might therefore predict that attractive (vs. unattractive) females would be more likely targets of harassing behavior to the extent that it is sexually motivated. However, physical attractiveness is a powerful feature in social interactions in more general terms. People ascribe more desirable traits to attractive others, applying the stereotype that "what is beautiful is good" (see Eagly, Ashmore, Makhijani, & Longo, 1991). Based on these findings, one might alternatively predict that highly attractive women would be treated with greater respect than unattractive women and, hence, would be harassed to a lesser extent. With the computer harassment paradigm, it was possible to investigate the role of target attractiveness on behavior directly, by systematically varying target attractiveness. Two rival



hypotheses concerning physical attractiveness were entertained:

*Hypothesis 2a:* Men send more sexist jokes to more attractive females than to less attractive females.

*Hypothesis 2b:* Men send more sexist jokes to less attractive females than to more attractive females.

### *Individual Difference Measures*

As in Study 1, we included the modified LSH scale as a predictor of sexual harassment. In addition, we administered a measure of *identification with the male gender*. Maass and her colleagues (2003) had found that males identifying highly with their gender showed greater harassment. This finding is well in line with the social identity framework of sexual harassment, which holds that greater salience of a relevant group identity will increase the likelihood of negative behavior toward an outgroup (see Maass et al., 2003, for further discussion). Study 2 also featured a measure of general negative attitudes toward women, the *hostile sexism* subscale of the Ambivalent Sexism Inventory (Glick & Fiske, 1996). Finally, we included a measure of impression management (Paulhus, 1998) to determine whether any of our measures would be prone to socially desirable responding. We formulated the following hypotheses:

*Hypothesis 3:* The higher a man's LSH score, the more sexist jokes he sends.

*Hypothesis 4:* The higher a man's hostile sexism score, the more sexist jokes he sends.

*Hypothesis 5:* The higher a man's gender identification score, the more sexist jokes he sends.

### *Method*

#### *Participants and Design*

Eighty-five male students were individually recruited, in the same way as in Study 1, for participation in two consecutive but allegedly independent studies. One of these studies would investigate memory abilities, whereas the other study would deal with the validation of personnel selection tests. The data of eight participants had to be excluded from analyses because of either

equipment failure (3) or doubts about the authenticity of their chat partner (5). The final sample thus consisted of 77 participants, whose mean age was 23.84 years ( $SD = 3.63$  years). The study featured a 2 (*attitude of chat partner*: feminist vs. traditional) x 2 (*attractiveness of chat partner*: high vs. low) factorial between-subjects design. Participants were randomly assigned to conditions; number of participants per condition ranged from 18 to 21.

### *Procedure*

Each session was presented as consisting of two independent studies, one on “memory performance”, the other on “the validation of several questionnaires”. In the memory study, participants would interact via a computer chat line with another participant at another university. Ostensibly, other participants who had consented to participate were simultaneously being instructed and ready to begin with the study at several locations. The computer would establish a connection with one of them. Each participant’s picture would be taken and sent to their respective chat partner so they would get an impression of him or her.

After a participant had agreed to this procedure, the experimenter took his photograph and then handed the camera to an assistant, ostensibly to upload the picture in the computer network. Participants were escorted to the laboratory and asked to start online with the interactive “memory study”. The remainder of the procedure was fully controlled by a Visual Basic computer program, which simulated a chat room over a university network.

After participants had “logged on”, the program pretended to connect them with another participant waiting at a different university. On the following screens, the alleged purpose of the study was explained: It would be of interest to investigate whether individuals’ performance on memory tasks was superior when they were passively exposed to material or when they could actively choose material. One participant’s role would be that of the “sender” who would choose the information (which was explained as being associated with active information processing), while the other participant would serve as the “receiver” and would only see the information

chosen by the “sender” (apparently associated with passive information processing). The information to be exchanged and remembered would be jokes. In each trial the “sender” would choose and send one joke from a pair of jokes, whereas the “receiver” would only see the joke that was selected by the sender. The receiver would be allowed to return a brief comment on each selected joke. Memory for the jokes would be tested subsequently. To induce a certain level of threat to participants’ male identity, all participants were informed that the aim of the study was to examine potential gender differences in memory performance (see Maass et al., 2003). One situational feature presumed to facilitate sexual harassment within the person x situation framework was thus present to some extent in all experimental conditions.

Next, target features were experimentally manipulated. Participants were asked to indicate their age, university, and subject of study, and to describe in a few sentences why they chose that subject and what their future plans were. This information together with their photo would be sent to their chat partner. After completing their own profiles, participants received the profile of the female chat partner “Julia” including her photo. This was where both target attractiveness and target attitude were manipulated (see below for detail). The computer then ostensibly performed a “random assignment” of sender and receiver roles. In fact, the participant was always assigned the sender role, and the female chat partner always the receiver role.

### *Selection of Jokes*

In each trial, participants had to choose one of two jokes that were presented in pairs. Every second pair represented a critical pair containing a sexist joke. The program was set up so that each sexist joke chosen would always be followed by a disapproving comment from the chat partner (e.g., “that’s a stupid joke”, or “quite offensive”). Participants would thus be made aware of the unwanted nature of their behavior. The number of sexist jokes sent was again defined as the main dependent variable. During the whole joke selection part the female chat partner’s photograph was displayed in the top left corner of the screen.

After completion of the last joke selection, participants were informed that the “second study” would now begin and that they would complete several computerized questionnaires. This would also serve as a filler task before their memory for the jokes would be assessed. The computer program then presented items measuring male identity, hostile sexism, the likelihood to sexually harass, and impression management (see below for detail).

At the end of this part, participants were asked to click a button to “return to the memory study”. Then they completed an open-ended suspicion check by indicating what they thought the purpose of the study was so far. Also, participants were asked to indicate how likable they thought their chat partner was (scale from 1, *not likable at all*, to 7, *very likable*) and how much they would like to meet her (scale from 1, *not at all*, to 7, *very much*). Then participants were thoroughly debriefed and were paid 4 EUR (approx. 4.88 US \$); psychology students had the option of earning course credit. A session took about 35 minutes in total.

### *Materials*

*Target attractiveness.* In two pilot studies ( $N = 8$  and  $N = 41$ ), male participants rated a number of color photographs showing head-and-shoulders views of women aged between 18 and 26 years on a scale from 1, *not at all attractive*, to 7, *very attractive*. Two photographs that best represented the category “high in attractiveness” ( $M = 4.80$  and  $4.85$ ) and two that best represented the category “low in attractiveness” ( $M = 1.93$  and  $2.17$ ) were chosen for the study. In the second pilot study, we also asked participants to indicate for each photograph how feminist they thought the depicted person was on a scale from 1, *not feminist at all*, to 7, *very feminist*. Analyses showed that feminism ratings were near the scale midpoint and did not differ between attractive ( $M = 3.89$ ;  $SD = 1.32$ ) and unattractive pictures ( $M = 3.77$ ;  $SD = 1.23$ ),  $t(40) = 0.42$ . The pictures thus were perceived as differing in attractiveness as intended, but did not in themselves convey information about the target’s feminism.

*Target’s gender-role related attitude.* Information about the chat partner’s feminist versus

traditional attitude was embedded in a short self-description that the target had allegedly typed in response to experimental instructions. Participants learned that the woman's name was Julia, that she was 22 years old and studied in Freiburg (a town about 600 km away from their own university). The further description differed depending on condition.

In the *feminist condition*, Julia had indicated that she was studying business administration and commented: "In my opinion studying business administration is ideal as you can demonstrate your skills, especially because I am aiming at a career in the management of a bank. Indeed, I do get to hear a lot that a management position isn't the right job for a woman because you hardly have enough time for family and children, but I think that women have many skills that are useful in management and that they can do a lot of things just as well as men or even better. This is also the reason why I've joined a group that campaigns for women's rights and equal opportunities in the job market."

In the *traditional condition*, Julia reported that she was studying education and commented: "In my opinion the job of a teacher, especially a primary school teacher, is ideal for a woman because you can have enough time for family and children. At first I intended to study law but I don't think the competition with all these men would have been right for me and that's why I've changed my mind. As for my plans, I will soon be working at a primary school for a couple of weeks. Other than that, I will just be finishing my studies. Later I also want to have children and so I probably won't be working for a while."

In a pilot study ( $N = 20$ ), male students had rated the feminist description at  $M = 3.03$  ( $SD = 1.02$ ) and the traditional description at  $M = 4.97$  ( $SD = 0.78$ ) on a 12-item scale (Cronbach's  $\alpha = .91$ ) ranging from "feminist" (1) to "traditional" (7),  $t(18) = 4.79$ ,  $p < .01$ . Thus, the two descriptions differed as intended in the gender-role related attitudes they conveyed.

*Selection of sexist and non-sexist jokes.* A final pilot study was carried out to obtain a new set of non-sexist and sexist jokes. Twenty-two male students (mean age 25.00 years,

$SD = 3.52$  years) read 83 jokes and answered the following questions for each joke: "How sexist / hostile toward women is this joke?", and "How funny is this joke?" (7-point scales anchored 1, *not at all*, and 7, *very*). Instructions emphasized that participants should judge each aspect independently. Similar to Study 1, we set out to match non-sexist jokes with equally funny sexist jokes. In addition, neutral pairs consisting of only non-sexist jokes were created. A total of 30 joke pairs (16 critical pairs, 14 neutral pairs) was selected for use in the main study. For critical joke pairs, pilot participants' averaged funniness ratings did not differ between non-sexist jokes ( $M = 2.59$ ,  $SD = .93$ ) and sexist jokes ( $M = 2.62$ ,  $SD = 1.12$ ),  $t < 1$ , *ns*. By contrast, the sexist jokes were rated as clearly more sexist / hostile toward women ( $M = 5.46$ ;  $SD = 1.57$ ) than the non-sexist jokes ( $M = 1.15$ ;  $SD = .31$ ),  $t(21) = 12.84$ ,  $p < .001$ . Thus, both neutral and sexist jokes were rated as low to moderate in funniness, whereas only the sexist jokes were rated as highly sexist / hostile toward women. None of the jokes had been used in Study 1. To provide an example, one critical joke pair read: "When does a woman lose 99% of her intelligence? – When her husband dies" (sexist) / "What do you get when you crossbreed a hedgehog and a tapeworm? – Six meters of barbed wire" (non-sexist).

### *Individual Difference Variables*

*Likelihood to sexually harass.* As in Study 1, the German version of the LSH Scale was used to assess a proclivity for sexually harassing behavior.

*Hostile sexism.* We used items from the German version (Eckes & Six-Materna, 1999) of the Ambivalent Sexism Inventory (Glick & Fiske, 1996) to measure hostile sexism (HS). The HS subscale contains 11 items, e.g., "Women seek power by gaining control over men".

*Impression management.* A tendency toward socially desirable responding was assessed using the 10-item impression management (IM) subscale of the German version (Musch, Brockhaus, & Bröder, 2002) of the Balanced Inventory of Desirable Responding (Paulhus, 1998). An item example is "I have done things that I don't tell other people about".

*Male identity.* To assess participants' male identity, we used 8 items from a German scale by Bohner and Sturm (1997) that is based on the Collective Self-Esteem (CSE) scale (Luhtanen & Crocker, 1992). Bohner and Sturm rephrased the original CSE items such that they pertain specifically to gender. An example item is: "I feel good about being male".

Each of the hostile sexism, impression management, and male identity items was accompanied by a scale from 1, *strongly disagree*, to 7, *strongly agree*. After reverse scoring negatively worded items, the internal consistency (Cronbach's  $\alpha$ ) of all self-report scales was examined. Results were satisfactory overall; LSH:  $\alpha = .71$ ; hostile sexism:  $\alpha = .90$ ; impression management:  $\alpha = .67$ ; male identity:  $\alpha = .74$ . Summary scores were formed for each scale by averaging across its items.

### *Results*

#### *Behavioral Measure of Sexual Harassment*

The number of sexist jokes sent was used as a summary score of harassing behavior. The potential range of the summary score was 0 to 16. Its observed range was 0 to 14 ( $M = 3.52$ ;  $SD = 3.12$ ), with 78% of participants sending at least one sexist joke.

#### *Intercorrelations Among Individual-Difference Measures*

Significant intercorrelations were found between LSH and hostile sexism,  $r(75) = .27, p = .018$ , as well as LSH and male identity  $r(75) = .33, p = .004$ , whereas hostile sexism and male identity were uncorrelated,  $r(75) = .09, ns$ . Both LSH and male identity were negatively correlated with impression management,  $r(75) = -.23, p = .043$ , and  $r(75) = -.29, p = .011$ , whereas hostile sexism was not,  $r(75) = -.08, ns$ .

#### *Attractiveness Manipulation Check*

To examine whether the manipulation of the chat partner's attractiveness was effective, participants' ratings of her likability and their desire to meet her ( $r[74] = .67, p < .001$ ) were averaged to form an index of liking. This index was subjected to a 2 x 2 analysis of variance

(ANOVA) with the factors target attractiveness (high, low) and attitude (feminist, traditional),  $MSE = 1.97$ . A strong main effect of target attractiveness emerged: Attractive targets were liked more ( $M = 4.70$ ) than unattractive targets ( $M = 3.50$ ),  $F(1, 72) = 13.87, p < .001$ . Feminist targets ( $M = 3.99$ ) and traditional targets ( $M = 4.18$ ) were liked to the same extent,  $F < 1$  for both the main effect of attitude and the interaction effect.

#### *Correlations Between Individual Difference Measures and Harassing Behavior*

We examined the correlations between each of the individual-difference variables on the one hand and the number of sexist jokes sent on the other. Maass and colleagues (2003) had assessed gender identification both before and after the experimental procedure. These authors found identification scores to increase significantly from pre- to posttest in harassers, but less so in non-harassers. In other words, harassment seemed to reinforce male identity. In the present study, the likelihood of harassment was predicted to differ between experimental conditions; further, all individual difference measures (including gender identification) were administered after the experimental procedure. Given Maass and colleagues' findings, correlations of individual-difference scores with harassing behavior might thus have been unduly affected in our study by a participant's experimental condition<sup>3</sup>, possibly resulting in spurious zero-order correlations in our data. To protect against spurious zero-order correlations, we chose a more conservative approach and computed partial correlations, with possible main and interaction effects of the experimental factors removed (cf. Cohen, Cohen, West, & Aiken, 2003). As was predicted, the number of sexist jokes sent correlated positively with LSH,  $r_{\text{partial}}(72) = .28, p = .015$ . Also as predicted, a strong positive correlation was found between the number of sexist jokes sent and hostile sexism,  $r_{\text{partial}}(72) = .46, p < .001$ . Our assumption that men high (vs. low) in male identity would show greater harassment was supported by a significant correlation between the jokes measure and the collective self-esteem measure,  $r_{\text{partial}}(72) = .26, p = .026$ . Overall, these results support our Hypotheses 3, 4, and 5.



Finally, a low but significant negative correlation was found between impression management and the number of sexist jokes sent,  $r_{\text{partial}}(72) = -.28, p = .016$ . This last correlation suggests that harassment scores in the present study may have been contaminated to some extent by a tendency to present oneself in a positive light.

We again repeated the analysis with a behavior index that did not include a participant's first sexist joke sent. In this analysis, the magnitude of correlations was almost identical.

#### *Ordinal Position of First Sexist Joke Sent*

We also examined the ordinal position of the first sexist joke sent in the sequence of joke pairs. This variable may serve as a further indicator of harassment proclivity (cf. Dall'Ara & Maass, 1999). Using only data from participants who sent at least one sexist joke, and partialling out possible main and interaction effects of the experimental factors, we found hostile sexism to be negatively correlated with the first sexist jokes' ordinal position,  $r_{\text{partial}}(55) = -.30, p = .023$ . Thus, the greater a participant's hostile sexism, the sooner he sent a sexist joke. Ordinal position was not related to male identity, social desirability, or LSH scores, partial  $r$ s < .17, *ns*.

#### *Effects of Target Attitude and Target Attractiveness on Harassing Behavior*

To examine the impact of target characteristics on men's harassing behavior, we conducted a hierarchical multiple regression analysis with the number of sexist jokes sent as the criterion variable. Following recommendations by Cohen, Cohen, West and Aiken (2003), we used a predetermined sequence of entering predictors such that the predictors of focal interest in this analysis (i.e., target attitude, and target attractiveness) were entered only after other variables (i.e., individual-difference scores) that may be a source of spurious relationships with the criterion. Specifically, in step 1, we entered our individual-difference variables (LSH, HS, male identity, and impression management) as predictors. This step merely served to remove variability due to individual differences from the behavioral measure. In step 2, we then tested main effects of the experimental factors, *target attractiveness* (low, high) and *target attitude* (traditional, feminist). To test for an interaction of the experimental factors, we entered their

multiplicative product in step 3. Predictor variables were again centered. Beta weights were recorded both in the step where a predictor was first entered (initial beta) and in step 3 (final beta). Initial beta weights reflect a given variable's unique contribution to the prediction of harassing behavior, over and above the effects of all other predictors from the same or any previous step. Results are shown in Table 1.

In step 1, hostile sexism emerged as a significant predictor of sexual harassment, initial  $\beta = .43$ ,  $p < .001$ , whereas other individual-difference measures did not make a significant unique contribution to the prediction of harassing behavior,  $ps > .10$  for the initial beta weights. Thus, whereas each of our individual-difference variables by itself had shown a significant partial correlation with harassing behavior, these variables' intercorrelations prevented most of them from becoming a significant predictor when they were entered simultaneously into the regression analysis as a set.

Of greater importance, in step 2, a significant main effect of target attitude emerged: As predicted, participants sent more sexist jokes to the target who expressed a feminist attitude than to the target who expressed a traditional attitude, initial  $\beta = .22$ ,  $p = .04$ . This result strongly supports Hypothesis 1. In contrast, target attractiveness did not affect the number of jokes sent significantly, either as a main effect (step 2) or in interaction with target attitude (step 3),  $ps > .31$  for the initial beta weights. The data thus did not support either Hypothesis 2a or Hypothesis 2b. We again repeated the analysis with a behavioral index that did not include a participant's first sexist joke sent. In this analysis, the pattern of significance remained unchanged for both the experimental factors and their interaction.

### *Discussion*

Positive correlations of several individual-difference measures, including the likelihood to sexually harass, hostile sexism, and identification with the male gender, with the number of sexist jokes sent showed that our modified harassment paradigm was again successful in

assessing sexual harassment. We could also demonstrate that this harassment paradigm can be embedded in an elaborate cover story that allows to investigate a range of possible factors, for example target characteristics such as feminist attitude or physical attractiveness.

### *Impact of Target Characteristics*

As predicted in our first hypothesis, the fictitious female chat partner who was described as holding feminist attitudes was harassed more than the female chat partner described as holding traditional attitudes. This target-specific strategy of joke selection rules out explanations of harassment that revolve exclusively around the perpetrator, for instance, that harassers might find sexist jokes in fact more funny than other jokes. Instead, it supports a social-identity model of sexual harassment and is consistent with previous findings (Dall'Ara & Maass, 1999; Maass et al., 2003). Females who are perceived to pose a threat to male dominance are thus more likely to become victims of sexual harassment. A potential alternative to this interpretation might be that participants harass feminist targets more because of a general negative attitude toward feminists, even in the absence of perceived identity threat. We suspect, however, that negative attitudes toward feminists may not have played a decisive role in the present study, given that participants' judgments of liking and dating interest were equally high for feminist and traditional targets in our attractiveness manipulation check. Nonetheless, the potential mediating role of perceived threat versus attitudes toward feminists could be studied in future research (see Maass et al., 2003).

Our second hypothesis concerned target attractiveness and consisted of two opposing predictions: The first assumed that highly attractive females would be harassed more than unattractive females, whereas the second predicted the opposite. None of these was clearly supported: Attractive and unattractive females were harassed to about the same extent. Previous research using third-person vignettes had revealed that, in the same ambiguous behaviors, observers perceive less harassment if the female target is less attractive (Golden, Johnson, &

Lopez, 2001). In contrast to naïve observers' intuitions, our results suggest that unattractive females may in fact face an equally high risk of being harassed as attractive females. Given that our results are based on a single study manipulating facial attractiveness via photographs, it would be desirable to employ other manipulations of physical attractiveness, such as variations of targets' interpersonal style or body silhouette, in future research.

#### *The Role of Individual Differences in Potential Perpetrators*

Our third hypothesis, that men high in LSH would show more harassment than men low in LSH was supported. Consistent with previous findings (the present Study 1; Maass et al. 2003; Pryor, 1987), the LSH scale, whose items address quid pro quo harassment, was again shown to correlate with a different form of sexual harassment, namely gender harassment. This result lends further credence to the notion of sexual harassment as a unitary construct.

Supporting our fourth hypothesis, another strong correlate of sexual harassment was hostile sexism. The stimulus materials for our behavioral measure of harassment consisted of jokes that conveyed a derogatory and sexist view of women as a group. Selecting and sending such jokes to a woman in a computer chat represents behavior that is consistent with the definition of hostile sexism as including negative evaluations of women and an overt antipathy toward women (Glick & Fiske, 1996; for related findings see Eyssel & Bohner, 2006). Interestingly, when entering all of our individual-difference measures simultaneously into a multiple regression analysis, only hostile sexism emerged as a significant predictor of sexual harassment. This may indicate that perceptions of women as an out-group (which are directly assessed by the hostile sexism scale) are a more important determinant of harassing behavior than the identification with men as an in-group (as assessed by the collective self-esteem scale), or than self-predicted behavior in scenarios of an apparently interpersonal nature (as assessed by the LSH scale). To shed further light on the relative importance of these predictors of harassment, future research may systematically vary the salience of gender groups by using

sexist jokes that capitalize either on gender ingroup strengths or on gender outgroup weaknesses, or that refer to gender only indirectly, for instance via the first names of protagonists.

A final prediction concerned participants' identification with the male gender. This individual-difference measure showed a significant correlation with harassing behavior, thus replicating previous results by Maass et al. (2003).

Finally, a significant negative correlation between impression management and the behavioral measure of harassment suggests that the latter may not be completely free from reflecting a tendency toward presenting oneself in a positive light. However, this correlation was small in magnitude, did not result in a unique contribution to the prediction of behavior in the multiple regression analysis, and also stands in contrast to the very encouraging findings of Study 1, where the harassment measure proved to be robust against instructions to present oneself favorably. Taken together, the findings of Study 2 thus yielded additional evidence for the feasibility of the computer harassment paradigm and for its convergent as well as discriminant validity.

### General Discussion

The results of two studies indicate that our refined computer harassment paradigm provides an economical and valid instrument for assessing sexual harassment in the laboratory. We observed rates of harassing behavior that were comparable in magnitude to results reported by Mitchell et al. (2004), who used a joke-telling paradigm where male students interacted face-to-face with a female confederate. Moreover, the task of choosing one of two jokes to be sent to a female chat partner seems to produce higher rates of harassing behavior overall, as compared to the task of selecting images from folders, which had been used in previous research (Dall'Ara & Maass, 1999; Maass et al., 2003). Since our paradigm does not use category labels indicative of the potentially undesired nature of some of the stimulus materials (like the folder label "porno" used by Maass and colleagues), it appears to be an ecologically more valid model of the

decision situations that may lead to sexual harassment outside of the laboratory. Finally, our results converged across two rather different formats of interaction between a participant and his virtual partner: Whereas Study 1 featured a rather minimalist format (i.e., sending jokes to, and receiving abstract credits from an almost anonymous female), Study 2 simulated a more realistic interaction which, in addition to an exchange of vivid personal background information, involved natural-language feedback about the jokes sent.

Given the high likelihood of observing the target behavior in our paradigm, it is important to note that not all men were equally likely to sexually harass. Instead, we find meaningful correlations with individual differences on the part of the perpetrators. Men high in hostile sexism, high in the likelihood to sexually harass, and identifying strongly with their gender showed higher rates of harassment compared to men scoring low in these individual differences. A slight setback is the fact that a tendency toward presenting oneself in a favorable light might affect responding to the computer harassment paradigm, resulting in somewhat lower rates of harassing behavior. This was only found at the correlational level in Study 2, however, whereas in Study 1 a direct instruction to present oneself favorably did not significantly affect harassment behavior. Nonetheless, it seems useful to further investigate the relationship between behavioral measures of sexual harassment and measures of impression management in future research.

Furthermore, the extent of harassing behavior also was clearly affected by target features. Women who hold and express feminist attitudes seem to be perceived as a threat by many men, and one reaction to this perceived threat consists of harassing these women in particular. This finding is by now well established (see Maass et al., 2003) and could be used to identify person-situation constellations that hold a high risk for sexual harassment to occur. The other target feature we investigated, physical attractiveness, did not have a clear-cut effect. Interestingly, however, the data do not fall in line with research that addressed the perception of attractiveness as a risk factor in others (e.g., Golden et al., 2001). This discrepancy should be further

investigated.

### *Considerations of External Validity*

Given the increasing prevalence of online harassment via the Internet, where gender harassment is one of the most typical forms of harassment observed (Barak, 2005), our paradigm may indeed be seen as directly emulating everyday forms of harassment. One aspect of the paradigm that could be fruitfully expanded and analyzed more closely is its “interactive nature”. In the present versions, the target person’s disapproving feedback for sending sexist jokes was rather limited and perhaps somewhat ambiguous. Although ambiguous reactions on the target’s part may not be very different from many real-life situations of harassment, it would be useful to vary the explicitness of the target’s feedback in future studies. If the correlation between LSH and harassing behavior could be replicated with highly explicit disapproving feedback, this would strengthen further our confidence in the computer harassment paradigm’s validity. The research by Maass and her colleagues (2003), who used more explicit feedback, is encouraging in this respect. From an applied perspective, it would be interesting to vary the explicitness of feedback communicated by the target as a means of discouraging men from showing harassing behavior. We believe that the computer harassment paradigm could be easily adapted for this purpose.

It might be objected that our experimental paradigm is lacking some of the attributes that characterize real-world sexual harassment as, for example, an expectation of future interaction between perpetrator and target, or the involvement of coworkers. However, as outlined in the Introduction, our operational definition of harassment includes many of the common features of definitions found in the literature: Participants were given repeated opportunities to engage in the critical behavior, and the target consistently communicated her disapproval. It should further be noted that studies of real-world sexual harassment have addressed behaviors that do *not* include the expectation of future interaction or the involvement of co-workers, but nonetheless show

high prevalence rates and are experienced by targets as highly stressful. Prominent examples are large-scale survey studies on harassment over the telephone (Sczesny & Stahlberg, 2000), or work on sexual harassment over the Internet (see Barak, 2005). Results from a study in our own lab, where women rated the severity of a range of potentially harassing behaviors in different contexts, suggest that it makes little difference whether a behavior is enacted by a co-worker or by someone whom the target encounters during her leisure time (Vanselow & Bohner, 2006).

A final limitation of our results may be seen in the use of German samples in both studies. Note, however, that our modified computer harassment paradigm is comprised of elements that, individually, have proven their utility in a range of cultures, including other European countries (i.e., Maass and colleagues' computer harassment paradigm) and the United States of America (i.e., Mitchell and colleagues' joke-telling paradigm). We are confident that our paradigm will lend itself readily to application beyond the present cultural context.

#### *Potential Applications*

What can be learned from our findings in terms of potential interventions? As harassment seems to be strongly linked to participants' gender-related beliefs, educational programs or persuasive messages aimed at reducing LSH or targeting hostile attitudes might be effective ultimately in reducing harassing behavior as well.

Furthermore, the finding of a causal link between threat to male identity and the likelihood of harassing behavior, which has now been repeatedly demonstrated in laboratory experiments (see Dall'Ara & Maass, 1999; Maass et al., 2003), has important implications. As Maass and her colleagues have noted, these findings suggest that the greater harassment against feminists, as observed in correlational studies in Italy (see Maass et al., 2003), does not just reflect differential reporting rates of feminists versus more traditional women. Instead, the available experimental data suggest that feminist women objectively are at greater risk of becoming victims of harassment. A highly interesting applied implication of this interpretation



would be that if gender harassment is aimed at defending a privileged status or male in-group identity, "then any strategy that reduces the power of categorization along gender lines may be effective. Rather than trying to change the male's attitudes, it may be considerably easier and more efficient to change those contextual aspects of work settings and other environments that are sources of gender categorization and identity threat and that may ultimately be conducive to sexual harassment." (Maass et al., 2003, p. 867). We believe that psychological research on intergroup relations has much to offer in this respect (see, e.g., Pettigrew, 1998), and it would be an interesting challenge to apply these insights to gender relations (for further discussion, see Maass et al., 2003).

Endnotes

<sup>1</sup> Distinctions between quid pro quo harassment and hostile environment harassment are no longer made in U.S. law. However, they may still provide a useful differentiation between forms of behavior with different psychological consequences.

<sup>2</sup> Döner Kebab is a popular fast-food, usually prepared from lamb.

<sup>3</sup> We thank an anonymous reviewer for pointing this out.

## References

- Barak, A. (2005). Sexual harassment on the Internet. *Social Science Computer Review*, 23, 77-92.
- Bohner, G., Jarvis, C. I., Eyssel, F., & Siebler, F. (2005). The causal impact of rape myth acceptance on men's rape proclivity: Comparing sexually coercive and noncoercive men. *European Journal of Social Psychology*, 35, 819-828.
- Bohner, G., Reinhard, M.-A., Rutz, S., Sturm, S., Kerschbaum, B., & Effler, D. (1998). Rape myths as neutralizing cognitions: Evidence for a causal impact of anti-victim attitudes on men's self-reported likelihood of raping. *European Journal of Social Psychology*, 28, 257-269.
- Bohner, G., & Sturm, S. (1997). Evaluative Aspekte sozialer Identität bei Frauen und Männern: Vorstellung einer Skala des Kollektiven Selbstwerts in bezug auf das Geschlecht (KSW-G) [Evaluative aspects of social identity in women and men: A German scale of collective self-esteem concerning gender]. *Psychologische Beiträge*, 39, 322-335.
- Brannon, L. (2002). *Gender: Psychological perspectives* (3rd ed.). Boston: Allyn and Bacon.
- Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, 50, 559-570.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.
- Dall'Ara, E., & Maass, A. (1999). Studying sexual harassment in the laboratory: Are egalitarian women at higher risk? *Sex Roles*, 41, 681-704.
- Driscoll, D. M., Kelly, J. R., & Henderson, W. L. (1998). Can perceivers identify likelihood to sexually harass? *Sex Roles*, 38, 557-588.
- Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but . . . : A meta-analytic review of research on the physical attractiveness

- stereotype. *Psychological Bulletin*, 110, 109-128.
- Eckes, T., & Six-Materna, I. (1999). Hostilität und Benevolenz: Eine Skala zur Erfassung des ambivalenten Sexismus [Hostility and benevolence: A scale measuring ambivalent sexism]. *Zeitschrift für Sozialpsychologie*, 30, 211-228.
- Eyssel, F., & Bohner, G. (2006). *It's only a matter of time: Individual differences in the endorsement of sexist and nonsexist jokes*. Unpublished manuscript, University of Bielefeld.
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology*, 27, 307-316.
- Fitzgerald, L. F., Swan, S., & Magley, V. (1997). But was it really sexual harassment? Legal, behavioral, and psychological definitions of the workplace victimization of women. In W. O'Donohue (Ed.), *Sexual harassment: Theory, research, and treatment* (pp. 5-28). Boston: Allyn and Bacon.
- Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70, 491-512.
- Golden, J. H., Johnson, C. A., & Lopez, R. A. (2001). Sexual harassment in the workplace: Exploring the effects of attractiveness on perception of harassment. *Sex Roles*, 45, 767-784.
- Guttek, B. A. (1985). *Sex and the workplace*. San Francisco: Jossey-Bass.
- Ilies, R., Hauserman, N., Schwochau, S., & Stibal, J. (2003). Reported incidence rates of work-related sexual harassment in the United States: Using meta-analysis to explain reported rate disparities. *Personnel Psychology*, 56, 607-631.
- Luhtanen, R. & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, 18, 302-318.
- Maass, A., Cadinu, M., Guarnieri, G., & Grasselli, A. (2003). Sexual harassment under social

- identity threat: The computer harassment paradigm. *Journal of Personality and Social Psychology*, 85, 853-870.
- Mitchell, D., Hirschman, R., Angelone, D. J., & Lilly, R. S. (2004). A laboratory analogue for the study of peer sexual harassment. *Psychology of Women Quarterly*, 28, 194-203.
- Musch, J., Brockhaus, R., & Bröder, A. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit [An inventory for the assessment of two factors of social desirability]. *Diagnostica*, 48, 121-129.
- O'Donohue, W., Downs, K., & Yeater, E. A. (1998). Sexual harassment: A review of the literature. *Aggression and Violent Behavior*, 3, 111-128.
- O'Hare, E. A., & O'Donohue, W. (1998). Sexual harassment: Identifying risk factors. *Archives of Sexual Behavior*, 27, 561-580.
- Paludi, M., & Paludi, C. (2003). *Academic and workplace sexual harassment: A handbook of social science, legal, cultural, and management perspectives*. Westport, CT: Praeger.
- Paulhus, D. L. (1998). *Manual for the Balanced Inventory of Desirable Responding*. Toronto, Canada: Multi-Health Systems.
- Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, 49, 65-85.
- Pryor, J. B. (1987). Sexual harassment proclivities in men. *Sex Roles*, 17, 269-290.
- Pryor, J. B., & Fitzgerald, L. F. (2003). Sexual harassment research in the United States. In S. Einarsen, H. Hoel, D. Zapf & C. L. Cooper (Eds.), *Bullying and emotional abuse in the workplace: International perspectives in research and practice* (pp. 79-100). London: Taylor & Francis.
- Pryor, J. B., LaVite, C. M., & Stoller, L. M. (1993). A social psychological analysis of sexual harassment: The person/situation interaction. *Journal of Vocational Behavior*, 42, 68-83.
- Schmidt, E. A., Gerger, H., Kley, H., Siebler, F., & Bohner, G. (2003, September). *Eine deutschsprachige Skala zur Erfassung der Neigung zu sexueller Belästigung* [A German

- scale for assessing sexual harassment proclivity]. Poster presented at the 9th Tagung der Fachgruppe Sozialpsychologie, Heidelberg, 21-24 September 2003.
- Schuster, B., Sczesny, S., & Stahlberg, D. (1999). Sexuelle Belästigung und Mobbing am Arbeitsplatz [Sexual harassment and bullying in the workplace]. In C. Graf Hoyos & D. Frey (Eds.), *Arbeits- und Organisationspsychologie: Ein Lehrbuch* (pp. 188-199). Weinheim, Germany: Beltz Psychologie Verlags Union.
- Schwarz, N., & Strack, F. (1981). Manipulating salience: Causal assessment in natural settings. *Personality and Social Psychology Bulletin*, 6, 554-558.
- Sczesny, S., & Stahlberg, D. (2000). Sexual harassment over the telephone: Occupational risk at call centres. *Work and Stress*, 14, 121-136.
- Tangri, S. S., Burt, M. R., & Johnson, L. B. (1982). Sexual harassment at work: Three explanatory models. *Journal of Social Issues*, 38 (4), 33-54.
- United States Merit Systems Protection Board (1995). *Sexual harassment in the Federal Government: Trends, progress, continuing challenges*. Washington, DC: US Government Printing Office.
- Vanselow, N., & Bohner, G. (2006, March). *Beurteilung sexuell belästigenden Verhaltens: Die Rolle von Kontextfaktoren* [Ratings of sexually harassing behaviors: The role of context factors]. Poster presented at the 48th Tagung experimentell arbeitender Psychologen, Mainz, Germany, 26-29 March 2006.
- Woodzicka, J. A., & LaFrance, M. (2005). The effects of subtle sexual harassment on women's performance in a job interview. *Sex Roles*, 53, 67-77.
- Zippel, K. S. (2006). *The politics of sexual harassment: A comparative study of the United States, the European Union, and Germany*. New York: Cambridge University Press.

Table 1

*Number of sexist jokes sent by target attitude and target attractiveness (Study 2)*

Step	Predictor	Beta		$R^2_{\text{Increment}}$
		Initial	Final	
1				.27***
	Likelihood to sexually harass	.04	.09	
	Hostile sexism	.43***	.41***	
	Male identity	.11	.15	
	Impression management	-.17	-.19 <sup>+</sup>	
2				.05 <sup>+</sup>
	Target feminist attitude	.22*	.22*	
	Target attractiveness	-.10	-.10	
3				< .01
	Target attitude X attractiveness	.05	.05	

*Note.* Hierarchical multiple regression of number of sexist jokes sent on experimentally manipulated target characteristics (steps 2 and 3), controlling for men's individual-difference scores (step 1).  $N = 77$ ; Overall  $R^2 = .33$ ,  $p < .001$ . Initial beta weights are from the step where a predictor was first entered; final beta weights are from step 3.

\*\*\*  $p < .001$

\*  $p < .05$

<sup>+</sup>  $p < .10$

## Appendix

## Example scenario from the modified Likelihood to Sexually Harass Scale

*Scenario 2* - Please read the following text carefully and imagine that *you* are the main character. Please rate how likely it is that you would perform each of the behaviors described. Please use the whole range of the scale for stating your personal opinion.

*Imagine that you are a famous Hollywood film director. You are casting for a minor role in a film you are planning. The role calls for a particularly stunning actress, one with a lot of sex appeal. How likely are you to do the following things in this situation?*

1) *You give the role to the actress whom you personally find most suitable for the role.*

not at all likely    1    2    3    4    5    6    7    very likely

2) *You give the role to the actress who agrees to have sex with you.*

not at all likely    1    2    3    4    5    6    7    very likely

3) *You ask the actress to whom you are personally most attracted to talk with you about the role over dinner.*

not at all likely    1    2    3    4    5    6    7    very likely