Accurately Detecting Flirting:
Error Management Theory, the Traditional Sexual Script, and Flirting Base Rate

Jeffrey A. Hall, Ph.D.

Chong Xing, MS

Seth Brooks

University of Kansas

Communication Studies Department

Bailey Hall

1440 Jayhawk Blvd., Rm 102

Lawrence, KS 66045-7574

(785) 864-1082

hallj@ku.edu

Funding: This research was supported by the University of Kansas General Research Fund (GRF Award #2301662).

Abstract

The present manuscript reports two studies on the accuracy of flirting detection. In Study 1, 52 pairs (N = 104) of opposite-sex heterosexual strangers interacted for 10-12 minutes, then self-reported flirting and perceived partner flirting. The results indicated that interactions where flirting did not occur were more accurately perceived than interactions where flirting occurred. In Study 2, 26 one minute video clips drawn from Study 1 were randomly assigned to one of eight experimental conditions that varied flirting base rate and the traditional sexual script. Participant observers (N = 261) attempted to determine if flirting occurred. Results indicated that base rate affected accuracy; flirting was more accurately detected in clips where flirting did not occur than in clips where flirting occurred. Study 2 also indicated that female targets' flirting was more accurately judged than male targets' flirting. Findings are discussed in relation to accuracy and courtship context.

Keywords: biological sex, courtship, flirting, perceptual accuracy

Accurately Detecting Flirting:

Error Management Theory, the Traditional Sexual Script, and Flirting Base Rate

Flirting is a form of communication that directly or indirectly signals attraction (Hall, Carter, Cody, & Albright, 2010; Lindgren, Parkhill, George, & Hendershot, 2008). It often occurs between individuals who have little relational history. Flirting is not always direct. Because of fear of rejection or appearing inappropriate (Kunkel, Wilson, Olufowote, & Robson, 2003), flirtatious behavior resembles other forms of communication (e.g., being friendly, joking). Flirting serves various goals, including playfulness, sexual invitation, and/or relational initiation (Henningsen, 2004). Consequently, observers or targets of flirting may not be able to accurately detect flirting when it occurs, particularly during initial interactions between strangers who lack prior information. How can one accurately determine if another person is flirting? The present investigation explores flirting detection accuracy in initial interactions between opposite-sex heterosexual strangers.

Two theories indicate when flirting will be accurately perceived. Due to ancestral mating challenges, error management theory (EMT) (Haselton & Buss, 2000) proposes that males adapted a bias that heightens the perception of sexual interest in females' behavior. This bias should increase detection of flirting in females who are actually flirting, thus reducing Type 2 error (i.e., false negative). Second, the traditional sexual script (TSS) is a cognitive conceptualization of the roles males and females play during all stages of courtship (Gagnon, 1990). The script follows that males are more interested in and pursuant of sex than females (Eaton & Rose, 2011). The present investigation will argue that the TSS influences both males' and females' flirting detection accuracy, wherein targets who behave in script consistent ways will be judged more accurately. That is, a lack of flirting will be more accurately detected in female targets and the presence of flirting will be more accurately detected in male targets.

Flirting detection accuracy may also be influenced by base-rate. Flirting is somewhat rare in initial interactions (Abbey, 1982), and is more difficult to judge from photos than other expressions (Farris, Treat, Viken, & McFall, 2008). Therefore, whether an observer can accurately detect flirting depends on whether the observer is watching a target who is flirting or not flirting. Utilizing both an in-person study of accuracy (Study 1) and a study of third-party observer accuracy (Study 2), the present investigation will argue that in zero-acquaintance interactions accuracy is a function of perceptual biases predicted by EMT, the TSS, and base

rate.

Theoretical perspectives

Error management theory. Evolutionary psychological theory suggests that there are sex differences in human mating preferences and strategies because males and females experience disparate reproductive realities (Schmitt, 2005). Due to asymmetric investment in offspring, behaviors in initial interactions are particularly prone to misdirection and ambiguity (Grammer, Kruck, Juette, & Fink, 2000; Haselton & Buss, 2000). EMT argues that evolutionary pressures have resulted in perceptual adaptations that are attuned to ancestral mating challenges (Haselton & Buss, 2000). For males, the reproductive benefits of successfully identifying females who are open to mating are outweighed by the reproductive costs of missing a chance to mate. EMT predicts that males overestimate females' sexual interest. A recent meta-analysis of sex differences in perceptions of sexual interest, including 13 samples and 3,631 participants, found that males perceived more flirtatiousness, seductiveness, and promiscuousness in evaluating the behavior of female targets (La France, Henningsen, Oates, & Shaw, 2009). However, a significant moderator of this effect suggested that this sex difference was larger when males in the interaction. This is consistent with EMT in that males interacting with females were at considerably greater risk of missing a potential mating opportunity than males observing videos or photos.

Traditional sexual script. The traditional sexual script is a cognitive conceptualization of the events that lead up to heterosexual sex, from the roles of those involved to the particular nonverbal and verbal actions taken in the early stages of courtship (Gagnon, 1990; La France, 2010). The sexual script is influenced by both cultural and interpersonal factors (Gagnon, 1990), and the script influences the perceptions of individuals who are engaging in courtship and evaluating how others engage in courtship (de Weerth & Kalma, 1995). The TSS links particular courtship behaviors to biological sex in ways that conform to traditional gender roles (Laner & Ventrone, 1998). Namely, the script asserts that males are sexual pursuers and females are yielding gatekeepers (La France, 2010). The script prevails despite societal changes in gender roles because it provides a familiar and culturally agreed upon heuristic from which to make decisions about ambiguous interactions (Eaton & Rose, 2011).

The TSS has implications for perceptions of flirting. The sexual script characterizes males as always interested in and ready for sex and females as uninterested in or

even actively hampering their own sexual desire (Impett & Peplau, 2003). There is a widespread perception that males and females in general adhere to the TSS, even when individual respondents themselves do not adhere to it (de Weerth & Kalma, 1995). This suggests that the script influences males' and females' perceptions of flirting – both as interactants and as observers of an interaction between strangers – wherein script accordant behavior will be more accurately judged than script discordant behavior.

Contrasting EMT and the TSS reveals differences and similarities. When judging female targets, the EMT and the TSS perspectives differ. EMT suggests that the perceptual biases of males adapted specifically for the purpose of reducing Type 2 error (i.e., a false negative). That is, males over-perceive sexual intent in females so as to not miss out on potential mating opportunities. A different bias is predicted from the TSS; targets that behave in a way that is consistent with the TSS will be more accurately judged. Therefore, females' lack of interest will be more successfully interpreted than males' lack of interest, and males' presence of interest will be more accurately judged than females' presence of interest. Consider the following investigation of cross-sex friends. The main effects that were consistent with EMT: males over-perceived their female friends' sexual interest and females under-perceived their male friends' sexual interest (Koenig, Kirkpatrick, & Ketelaar, 2007). However, once accounting for how interested each friend actually was, males *under-perceived* females' sexual interest and females *over-perceived* males' sexual interest. Koenig et al. explain this unexpected finding by arguing that "socially acquired knowledge that men have more sexual interest than women" may have influenced estimates of attraction (p. 426). Interpreted through the lens of the TSS, this suggests that flirting by males is more easily detected because it is presumed to be present, but flirting by females is presumed not to be present, which leads women who are not flirting being more accurately assessed. When testing flirting detection accuracy, EMT would predict that:

H1: When female targets are flirting, males' flirting detection accuracy will be higher than when female targets are not flirting.

Hypotheses drawn from the TSS predict:

H2a: When male targets are flirting, they will be more accurately judged than when male targets are not flirting.

H2b: When female targets are not flirting, they will be more accurately judged than when female targets are flirting.

Both EMT and the TSS suggest that males and females will perceive a male target as more flirtatious than a female target. When rating female targets, male and female observers judge females' flirtatiousness similarly (La France et al., 2009). Therefore:

H3: Male and female observers will perceive male targets to flirt more frequently than female targets.

EMT would suggest an additional difference in flirting perception, wherein compared to females, males will perceive more flirting by female targets. Meta-analytic data suggests that when the observer is a third-party outside of the interaction, male observers judge both individuals in a cross-sex interaction as more flirtatious compared to female observers (La France et al., 2009). Therefore:

H4a: Compared to females, males will perceive more flirting by female targets.

H4b: Compared to females, males will perceive more flirting by male targets.

Flirting base rate

One of the great insights of contemporary research on deception (i.e., Levine, Kim, Park, & Hughes, 2006) is that the ability to detect lies is a function of the base rate, or the proportion of truths and lies that observers evaluate. Due to the veracity effect (i.e., the truth-bias), it is considerably easier to accurately judge honesty than deception. Deception research has two critical implications for theorizing about the accurate detection of flirting: (a) the presence/absence of flirting in any interaction will influence the accuracy of detecting flirting, and (b) a bias similar to the veracity effect might exist for detecting flirting.

The examination of the mean perception of flirtatiousness in two studies of interest in a zero-acquaintance interaction (i.e., Abbey, 1982, Study 1; Saal, Johnson, & Weber, 1989, Study 1) suggests that the amount of flirting is small and potentially rare (i.e., near 2.0 on 7-point scales). However, no study to date has documented the percent of zero-acquaintance interactions where flirting occurs. When speed dating, females are chosen for dates by about ½ of males and males are chosen for dates by 1/3 of females and participants tend to reciprocate romantic interest (Kurzban, & Weeden, 2005). Even at events where flirting is common, mutual flirting occurs in less than half of interactions.

If romantic interest is somewhat rare and clear signals of romantic interest are unlikely (Grammer et al., 2000), the accuracy of detecting flirting when it is present may suffer. One reason that deception may be more difficult to detect than honesty is that in everyday interactions

there are far fewer incidences of deception than honesty (Serota, Levine, & Boster, 2010); the truth-bias is grounded in lived experience. Similarly, the frequency of expression of an emotion within a culture affects the accuracy with which that emotion is detected within that culture (Matsumoto, 2006). Two courtship relevant studies provide some evidence on base rate. When trained to look for particular nonverbal courtship displays in females, the accuracy of predicting subsequent approach behaviors by males was lower (i.e., 80%) than the accuracy of predicting non-approach (i.e., 100%) (Moore & Butler, 1989). Similarly, there is evidence that romantic interest is less easily judged than friendliness (Farris et al., 2008). Taken together, the absence of flirting may be more easily detected than the presence of flirting:

H5: Flirting detection accuracy will be higher when targets are not flirting than when targets are flirting.

Participant versus observer

No study to date has explored whether participants in an interaction are more accurate judges than outside observers. It is possible that the conversational and self-presentational demands of the interaction hinder participants' ability to make accurate judgments. Alternatively, being involved in the interaction may provide individuals unique and more precise insight into conversational partners' feelings and thoughts. To consider this issue, we offer:

RQ1: How will flirting detection accuracy of participants in an interaction compare to the accuracy of third party participant observers of that same interaction?

STUDY 1

Method Study 1

Procedure and Instrumentation

Participants were 52 male and 52 female college students randomly matched into 52 opposite-sex pairs. Participants were self-reported to be single (i.e., not in a committed romantic relationship or a "serious dating" relationship) and heterosexual. Participants received partial class credit for participating in the study worth less than .5% of their final grade. Participants were primarily white (78%), and other races/ethnicities were represented: 7% Asian-American, 7% Hispanic/Latino, 6% African-American, 2% Native-American. Participants' mean age was 19.2 years (SD = 2.1, range 18-30; mode = 19).

Pre-interaction questionnaire. Approximately one week before the in-person portion of the study, participants completed the online questionnaire to determine their eligibility to

participate (i.e., heterosexual and single were inclusion criteria), and completed other measures reported elsewhere.

Interaction procedure. Upon arrival at the dyadic interaction lab, participants were led to separate rooms and gave written consent to be audio and video recorded. Once both participants had arrived and had consented, participants were introduced. The interaction lab had two chairs facing one another at a distance of approximately three feet, and a small adjacent side table. Two digital, wall-mounted cameras were placed above each participant to record interactions – each camera recording one participant. Two wireless microphones placed on the arm of each chair captured audio feed. The microphones and cameras were connected to a password-protected computer in the control room next door where videos were streamed, recorded, and stored.

After being introduced, the participants were read study instructions. They were told the purpose of the study was to "better understand how people form first impressions," and that they would be interacting for about 10 min. To help facilitate and standardize the conversation, a set of cards with conversation starting questions was placed on the side table. Each participant was asked to choose five of the ten cards, and to take turns asking each other questions. Participants were instructed that the goal of the interaction was to have a conversation. They were encouraged to go on tangents and/or ask their own questions. Participants were asked to keep talking until the researchers returned.

Post-interaction procedure. After at least 10 min (but no more than 12 min) had passed, participants were interrupted by the researcher and put in separate rooms. Without consulting each other, they completed a post-interaction questionnaire. Participants reported their own flirting during the interaction ("I flirted" or "I did not flirt"), and perceived flirting in conversation partners ("he/she flirted" or "he/she did not flirt"), and items measuring social, task, and physical attraction (McCroskey & McCain, 1974) to the conversational partner. A logistic regression demonstrated that participants who self-reported being physically attracted to their conversational partner were 2.24 times (224%) more likely to report that they flirted, B = .81, SE = .25, t = 9.49, p <. 001. Social attraction was unrelated to reporting flirting. Task attraction was negatively related to reporting flirting, B = -.57, SE = .28, t = 4.12, p = .04.

Table 1. *Descriptive Statistics Study 1* (*N* = 104)

Variable	Count
Participants' flirting	I flirted (25)
	Female (14)
	Male (11)
	I did not flirt (79)
	Female (38)
	Male (41)
Perceived flirting	He/she flirted (20)
	He flirted (9)
	She flirted (11)
	He/she did not flirt (84)
	He did not flirt (43)
	She did not flirt (41)

Results Study 1

Table 1 reports the descriptive statistics for perceived partner flirting (my partner flirted = 1, did not flirt = 0) and partner self-reported flirting (I flirted = 1, I did not flirt = 0). Table 2 cross-classified participant sex, partners' flirting, and participants' flirting detection accuracy. If a participant's perception of whether his/her partner flirted agreed with the partner's self-reported flirting then detection accuracy was coded as 1 = correct. If a participant's perception of whether his/her partner flirted disagreed with the partner's self-report then detection accuracy was coded as 0 = incorrect.

The hypothesized interaction effects were explored. Overall, female detection accuracy was 69% while overall male detection accuracy was 71%, $n = 104$, $\chi^2(1) = .05$, $p = ns$.[1] Drawn from EMT, H1 predicted that when females flirt, males' flirting detection accuracy should be higher than when females do not flirt. There was evidence to the contrary of H1: males were more accurate when their female partners were not flirting (84%) than when they were flirting (36%), $n = 52$, $\chi^2(1) = 9.19$, $p < .01$. According to the two hypotheses derived from the TSS, it was predicted that when male targets were flirting, they would be more accurately judged than

Table 2. *Recognition Accuracy Comparison between .Interactant (Study 1) and Observer Simulation (Study 2) by Participant Sex and Partner's Flirting (N = 104)*

		Recognition accuracy	
Interactant/observer sex	Partner/target flirting	Study 1	Simulation (95% CIs)
Overall			
Female		69% (*n* = 36)	53% (40 – 65%)
Male		71% (*n* = 37)	63% (50 – 75%)
By Partner Flirting			
Female	Flirted	18% (*n* = 2)	23% (7 – 43%)
Female	Did not flirt	83% (*n* = 34)	65% (47 – 79%)
Male	Flirted	36% (*n* = 5)	43% (18 – 73%)
Male	Did not flirt	84% (*n* = 32)	68% (54 – 81%)
Combined Accuracy	Flirted	28% (*n* =7)	32% (16 – 48%)
Combined Accuracy	Did not flirt	84% (*n* = 66)	68% (56 – 76%)

Note: CIs = confidence intervals.

when male targets were not flirting (H2a), and when female targets were not flirting, they would be more accurately judged than when female targets were flirting (H2b). The results offer evidence to the contrary of H2a: when male targets were more accurately judged when they were not flirting (83%) than when they were flirting (18%), $n = 52$, $\chi^2(1) = 17.07$, $p < .001$. There was support for H2b: female participants who were not flirting were more accurately judged (84%) than female participants who were flirting (36%), $n = 52$, $\chi^2(1) = 9.19$, $p < .01$. These frequencies implied a potential interaction effect between sex and flirting condition (Figure 1). Because participants were paired in cross-sex interactions observer sex and target sex were confounded (i.e., female observers always judged male targets, male observers always judged female targets). Therefore, the effects of observer sex could not be disentangled from the effects of target sex, which excluded tests of H3 and H4a-b.

H5 predicted that targets who were not flirting would be more accurately perceived than targets who were flirting. The results indicated strong support for H5: accuracy (84%) was higher when participants' partners were not flirting compared to when partners were flirting

(28%), $n = 104$, $\chi^2(1) = 28.00$, $p < .001$.

To test for the potential interaction effect observed in Figure 1 and to account for the nesting of observations by dyad, generalized linear mixed modeling estimated the effects of flirting condition and sex on accuracy. The flirting condition was significant, log odds = -4.22, odds ratio = .015. The odds of accurately detecting flirting decreased by 99% when conversational partners were flirting. Neither the sex nor the target flirting condition by sex interaction effect was significant.

Discussion Study 1

The results showed no support for hypotheses derived from EMT, which suggests that males overestimate female interest so as not to miss an opportunity to mate, thereby rending their judgment more accurate when females are actually flirting, but impairing judgment when they are not flirting. Although males were more accurate in detecting female flirting (36%) than females were in detecting male flirting (18%), neither rate was above chance levels. Instead, males were more accurate when judging females' lack of flirting than females' flirting.

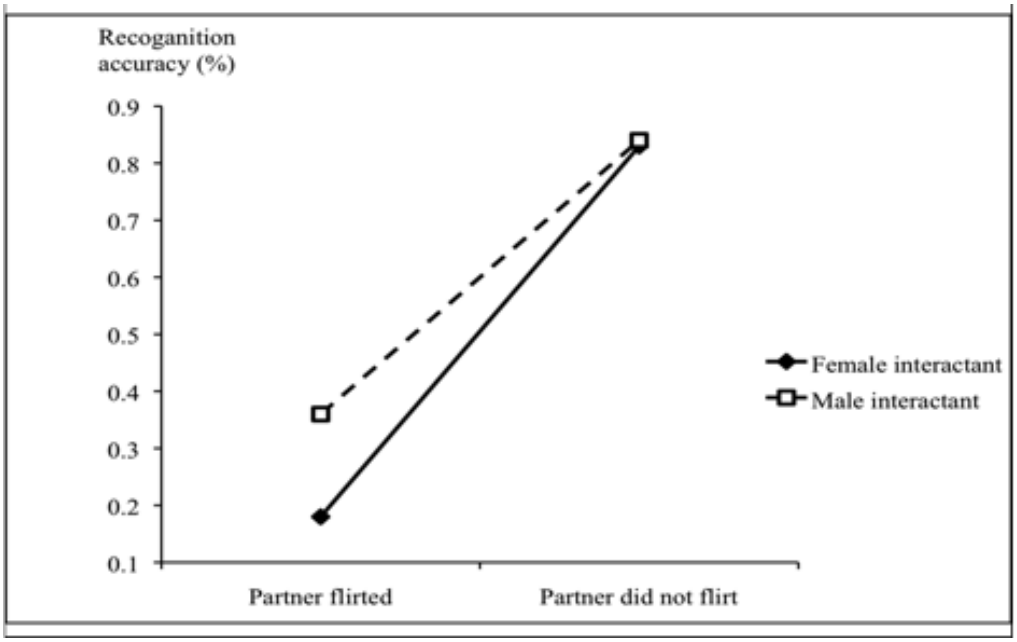

Figure 1: *Study I recognition accuracy of female and male in-person participants by partner flirting versus non-flirting.*

The results showed mixed support for the predictions derived from the TSS (i.e., males' flirting would be more accurately detected; females' lack of flirting would be more accurately detected). Female participants neither perceived that their partners flirted more often (males

perceived flirting in 21% of interactions and females perceived flirting in 17% of interactions), nor did they detect flirting more accurately when it did occur, showing a lack of support for H2a. When considering the results from the perspective of male participants, males were significantly accurate above chance when judging females who were not flirting (84%) compared to females who were flirting (36%), which supports predictions derived from the TSS (H2b).

The results provided strong support for H5: detection accuracy was significantly higher when participants' partners were not flirting than when they were flirting. Male and female participants were equally likely to accurately detect flirting overall (71% vs. 69% respectively). However, male participants were more likely to detect flirting when it was present (36%) than were females (18%), but equally accurate when a partner was not flirting (84% vs. 83% respectively). Logistic analysis revealed that this did not translate into a significant interaction effect.

Because participants were paired in cross-sex interactions, observer sex and target sex were always confounded. The effects due to observer sex could not be disentangled from the effects due to target sex. This is particularly important when interpreting whether the support for H2b was due to target sex and/or observer sex. Study 2 was conducted to disentangle moderations of accuracy, and to experimentally test the effects of base rate and the TSS on accuracy.

STUDY 2

Method Study 2

Procedure and Instrumentation

Participants (N = 261) were recruited from introductory public speaking courses at a large Midwestern public university in exchange for partial course credit worth less than .5% of their final grade. There were more female participants (n = 185) than male participants (n = 76). The sample was predominantly white or Caucasian (86.8%) and other races and ethnicities were represented: 5.6% Hispanic or Latino, 3.2% African-American or Black, 2.4% mixed race, and 2.0% Asian or Asian-American. Participants were on average 19.6 years of age (SD = 3.29; range was 18 to 47; *mdn* = 19; *mode* = 19). Participants' self-reported sexual orientation was heterosexual (94.8%), with 2.4% identifying as gay or lesbian and 2.8% identifying as bisexual.

Procedure and instrumentation. Participants signed up to attend a 20 min lab session in groups of 10-20 individuals. Upon arriving to the lab, participants were told that purpose of the

study was to study first impressions in initial interactions. Participants then consented to participate in the study and completed demographic measures. Groups of participants were randomly assigned to one of eight experimental conditions. In all experimental conditions participants watched 6 one min audio/video clips projected on a screen, viewing only one person from the dyad on the screen. Each experimental condition had 3 male and 3 female targets in the set of 6. After watching each video, participants answered seven yes/no questions about the person in the video. To avoid priming participants to look for flirting or romantic attraction, the questionnaire had five distracter items and an *acquaintance check* item (i.e., "Have you ever met the person in the video before?"). Distracter items included, "Did this person genuinely enjoy the conversation?" and "Do you think this person would make a good boss?" Very few participants had previously met the person in the video (1.9%). The dependent measure was "Do you think this person was flirting?" To check for suspicion, after watching all 6 videos, participants were asked to write at the bottom of the questionnaire what they thought the study was about. Employing a broad criterion of suspicion, 11.2% of participants reported suspicion that the study was about attraction, flirting, or dating, but no one identified the true nature of the study (i.e., flirting detection accuracy). Finally, participants were debriefed and dismissed.

Accuracy criteria. Participants were coded as accurate when they perceived the person in the video was flirting and the person in the video had self-reported flirting (true positive) or when they perceived the person in the video was not flirting and the person in the video had self-reported not flirting (true negative). Participants were not accurate when they either perceived flirting in the video when it was not present (false positive) or failed to perceive flirting when it was present (false negative).

Video clip selection and experimental conditions

Observation targets. Twenty-six video clips used in Study 2 were selected from Study 1 interactions. The selected videos included 13 men and 13 women. Because participants in Study 2 could only see one person in the video, 50% of interactions and 25% of participants in Study 1 were included in Study 2. The one min clips were extracted within minutes six through 11 of the interaction because this is the most accurate period for perceiving interest (Place, Todd, Penke, & Asendorpf, 2009) and the period associated with interactants' showing nonverbal behaviors indicative of interest (Grammer et al., 2000). The clips captured a conversational exchange, a story, or an answer to a question in its entirety. The person in the video talked the majority of the

time, but the video may have included some off-screen verbal comments/questions.

Experimental conditions: There were two predicted moderators of flirting accuracy: portion of videos were targets were flirting and portion of videos that followed the TSS (i.e., men flirting, women not flirting). There were 5 potential conditions for each moderator: 0%, 33%, 50%, 66%, 100%. A complete 5 (traditional script) x 5 (amount of flirting) experimental design could not be created because all traditional conditions required manipulation of the flirting behaviors. For example, 100% traditional condition must have all males flirting and all females not flirting, while a 0% traditional condition must have all males not flirting and all females flirting. This restricted manipulation of the experimental conditions. The 8 conditions employed were the only possible combinations of the two independent variables given these restrictions: 1 condition 0% flirting/ 50% traditional; 1 condition 33% flirting/ 50% traditional; 4 conditions 50% flirting with 0%, 33%, 66%, and 100% traditional, 1 condition 66% flirting/ 50% traditional, 1 condition 100% flirting/ 50% traditional.

Assignment of targets to experimental conditions. The 26 video clips were randomly assigned to experimental conditions. The order of video presentation (flirting/not; male/female) was randomized. Each video clip was watched a similar number of times by participant observers, and each clip was randomly assigned to conditions and was surrounded by different videos presented in a random order. Because assignment was random and the number of participants in viewing groups varied, each video was watched on average 59.88 times (range 38-87 times). Finally, because base rate was expected to affect accuracy, the total proportion of videos that contained flirting watched in Study 2 (32%) was controlled to approximate the proportion of videos that contained flirting in Study 1 (24%). To achieve this proportion, experimental conditions that had less or no flirting were over-assigned.

Results Study 2

Including the results from all conditions and all participants, there were 1,557 observations of targets in videos (see Table 3 for cell frequencies and observer accuracy percentages).[2] Participants' overall accuracy was 57%, which approximates the accuracy reported in Place et al. (2009). To test study hypotheses, a multilevel model was constructed wherein observer judgments (Level 1) were nested within observer characteristics (Level 2), which were nested within experimental condition (Level 3). MLM allows researchers to model and control for within individual differences and non-independence of conditions and observers

(Snijders & Bosker, 2012). Target characteristics (i.e., target sex, target flirting/not, target known by observer/not) were modeled at Level 1, participant observer characteristics (i.e., observer suspicion of study purpose, observer sex, race/ethnicity, age, sexual orientation) were modeled at Level 2 variables, and the experimental conditions (i.e., base rate, TSS) were Level 3 variables. Given the mixed factorial design, intraclass correlations (ICCs) were estimated. The Level 3 ICC was extremely small, < .001. The Level 2 ICC was .03, which suggested that 3% of variance in accurately judging target flirting existed between participants. These ICC estimates suggested that logistic regression was the appropriate test for further analyses (Snijders & Bosker, 2012).

Table 4. *Study 2 Cross-Classification of Observer Sex, Target Sex, Flirting Conditions, Perceived Target Flirting, and Recognition Accuracy* (*N* = 1557)

Observer	Target	Flirting conditions	Perceived flirting (*n*)	Accuracy (*n*)
Female	Female	Flirting		51% (187)
Female	Female	Non-flirting		67% (368)
Female	Male	Flirting		22% (170)
Female	Male	Non-flirting		64% (385)
Male	Female	Flirting		43% (76)
Male	Female	Non-flirting		68% (149)
Male	Male	Flirting		33% (64)
Male	Male	Non-flirting		62% (158)
			Total	57%
Female	Female		39% (555)	61% (344)
Female	Male		32% (553)	51% (283)
Male	Female		36% (222)	60% (134)
Male	Male		36% (222)	51% (114)
		Flirting	38% (203)	38% (203)
		Non-flirting	34% (349)	66% (672)

The main effects of participant and target sex were explored, controlling for participant observer demographic characteristics, observer suspicion, and observer familiarity with the

target. None of the control variables significantly affected accuracy. The results suggested that male and female observers were equally accurate, B = .02, SE = .12, p = .87. The results suggested that female targets' flirting was more accurately detected than males targets' flirting, B = .45, SE = .11, p < .001, R^2 = .02.

To test hypotheses derived by theory, specific interaction effects were probed. As predicted by EMT, H1 stated that male perceptual biases would result in greater flirting accuracy detection when females were actually flirting. Controlling for demographic covariates, study suspicion, and target familiarity, the results suggested that males' accuracy was significantly higher when female targets were not flirting than when they were flirting, n = 225, B = -1.01, SE = .29, p = .001, R^2 = .07. In contrast to H1, log odds showed that accuracy was 63% higher when male observers were judging female targets who were not flirting compared to female targets who were flirting. The two hypotheses drawn from the TSS predicted that target sex would moderate observer accuracy. The results revealed that flirting in male targets was more accurately detected when targets were not flirting than when they were flirting, n = 777, B = -1.65, SE = .18, p < .001, R^2 = .16. Log odds shows that participant observers were 81% more likely to accurately decode a male observer who was not flirting than one who was flirting, which does not support H2a. Finally, the results indicated that observers were significantly more likely to accurately detect female targets who were not flirting than female targets who were flirting, n = 780, B = -.77, SE = .16, p < .001, R^2 = .04. According to log odds, participants were 54% more likely to accurately perceive female targets who were not flirting than to detect flirting in female targets who were flirting, which supports H2b.

Perception frequency: The interaction effects predicted in H3 and H4a-b were tested controlling for participant demographic characteristics, observer suspicion, and observer target familiarity. None of the control variables significantly affected perceptions. Female observers perceived female targets to flirt 39% of the time and male observers perceived female targets to flirt in 36% of time (Table 3), but this difference was not significant, n = 780, B = .12, SE = .16, p = .48. Male observers perceived male targets to flirt 36% of the time and female observers perceived male targets to flirt 32% of the time, but this difference was not significant, n = 777, B = -.22, SE = .17, p = .19. Therefore, the results did not support H4a or H4b. However, observers perceived female targets to flirt 38% of the time and male targets to flirt 33% of the time overall, but this difference was not significant at p = .001, n = 1557, B = .25, SE = .11, p =

.041, R^2 = .01.

H5 predicted that targets who were not flirting would be more accurately judged than targets who were flirting. The results indicated strong support for H5, B = -1.17, SE = .12, p < .001, R^2 = .09. Logistic regression demonstrated that accuracy improved by 59% when observers were judging a target that was not flirting compared to a target that was flirting. This supported H5 and was consistent with Study 1.

Experimental effects: As a final test of the TSS (H2a-b) and flirting base rate (H5) the effects experimental condition on accuracy was tested. Participant observers were randomly assigned to experimental groups that systematically varied flirting base rate and the degree to which the videos matched the TSS. Controlling for demographic covariates as well as study suspicion and familiarity with target, the results indicated that both experimental effects significantly influenced accuracy. Contrary to predictions drawn from the TSS, as experimental conditions increased the degree to which males were flirting and females were not, accuracy of observers decreased, yet this effect was not significant at p = .001, n = 1557, B = -.14, SE = .06, p = .019, R^2 = .004. In support of H5, base rate was a significant predictor of accuracy, wherein videos with more flirting were less accurately judged, n = 1557, B = -.16, SE = .04, p < .001, R^2 = .01.

Participant versus observer

RQ1 queried whether accuracy would differ between in-person participants and third party participant observers. This RQ was answered in two ways: (a) directly exploring Study 1 results, and (b) conducting a computer simulation matching Study 1 parameters with Study 2 data. For the 26 videos selected from Study 1 and used in Study 2, the in-person interaction partners from Study 1 accurately judged the presence or absence of flirting 67% of the time. Accuracy for participants in Study 1 when flirting occurred was 44% and accuracy when flirting did not occur was 71%. For third-party observers in Study 2 watching videos of the same interactions, overall detection accuracy was 57%; accuracy when flirting occurred was 38% and accuracy when flirting did not occur was 66%. Therefore, participants in the interaction were more accurate detectors of flirting than third party observers. Because the overall base rate of flirting in Study 2 was higher than Study 1 and base rate influences accuracy, a data simulation was conducted.

This simulation compared Study 1 and Study 2 findings, while accounting for the cell

constrains of Study 1. Conditional bootstrapping with replacement was used for resampling the observations of Study 2 (N = 1557). The observed cell frequencies of Study 1 were used as parameter specifications for random resampling of the observations in Study 2.[3] Resampled datasets were combined to form estimates of detection accuracy with 95% confidence intervals. The simulation results demonstrate that the averaged recognition accuracy percentage for female observers observing male targets was 53% across all flirting conditions (95% interval = 40% to 65%). This estimate was lower than accuracy reported in Study 1 (69%), and suggested that in some conditions in-person participants were more accurate than third party observers (Table 2). Although the accuracy pattern was similar between the simulation and Study 1, there were two conditions where in-person participants were accurate at rates beyond the 95% confidence interval. In-person female participants were more accurate than third party female participants, and in-person participants were more accurate than third party participants when the target was not flirting.

Discussion Study 2

Using an experimental design to test flirting perceptions and flirting detection accuracy of third party observers, Study 2 was able to explore the hypothesized interaction effects not tested in Study 1. Perceptual differences in accuracy predicted by EMT were not supported: males were not able to detect flirting in females more frequently when it occurred than when it did not occur (H1). Similar to the results of Study 1, the results of Study 2 supported only half of the hypotheses drawn from the TSS. Specifically, both males and females were more accurately judged when they were not flirting than when they were flirting. For females, this is in accord with the traditional script (H2b), but for males this was not (H2a). Contrary to H3, observers did not perceive male targets to flirt significantly more frequently than female targets. Contrary to H4a and H4b, male and female observers perceived the frequency of flirting in targets similarly.

In contrast, the effects of base rate were robust. The experimental results further supported the importance of base rate in explaining variation in accuracy: conditions with more flirting by targets diminished accuracy compared to videos with conditions with less flirting by targets. The experimental results showed a lack of support for a TSS bias. Specifically, conditions where male targets were flirting (more traditional) showed decreased accuracy compared to conditions where female targets were flirting (less traditional), yet the effect size

was small.

In response to RQ1, in-person observers were more accurate than third-party observers. The accuracy of the conversational partners of the targets in the 26 videos selected from Study 1 and used in Study 2 was higher than the accuracy of third party observers watching the same videos. The simulation results showed that in-person accuracy was higher than third party accuracy in two conditions: when the observer is female (and the target is male) and when the target is not flirting. Accuracy was very similar when the target was flirting.

GENERAL DISCUSSION

The present multi-study investigation advanced hypotheses drawn from EMT and the TSS, and extended prior research by exploring the influence of flirting base rate, or the presence/absence of flirting in the target(s), on accuracy. The results show very clear support for the effect of flirting base rate on observer flirting detection accuracy in zero-acquaintance interactions, and mixed or no support for other explanations of accuracy.

Error management theory

Hypotheses drawn from EMT (Haselton & Buss, 2000) explored the perceptual biases of male observers compared to female observers. While all observers tended to see female targets as more flirtatious (38%) than male targets (33%), this difference was not significant and was not the result of males' increased perception of female flirting. Rather, when female observers judged whether female targets were flirting, flirting was perceived most frequently (39%). These results are inconsistent with EMT, but not without precedent. La France et al. (2009) found that when estimating female targets' flirtatiousness, there were no differences between male and female observers' perceptions. Rather, La France et al. found that larger sex differences consistent with EMT existed when observers judged female targets' seductiveness and promiscuousness (see also Lindgren et al., 2008). This suggests that the measure employed in the present investigation may have been insufficiently valenced in terms of sexual intent to capture expected socio-sexual perceptual biases.

It is important to point out that the present investigation may be the first to test whether the perceptual biases predicted by EMT (i.e., reduced Type 2 error) result in improved flirting detection when female targets are actually flirting. Neither Study 1 nor 2 found support that presumed perceptual biases led to greater accuracy. Study 1 did offer some evidence that in the context of an in-person interaction, males were more accurate than females in detecting flirting

when it was actually happening. However, this observer effect could not be disentangled from a significant target effect (i.e., females being more readable), which was shown to be a significant predictor of accuracy in Study 2. Overall, the present investigation shows nearly no support for EMT in the context of judging flirting in zero-acquaintance interactions, and suggests that increased perceptions of flirting do not translate to reduced Type 2 error for male observers.

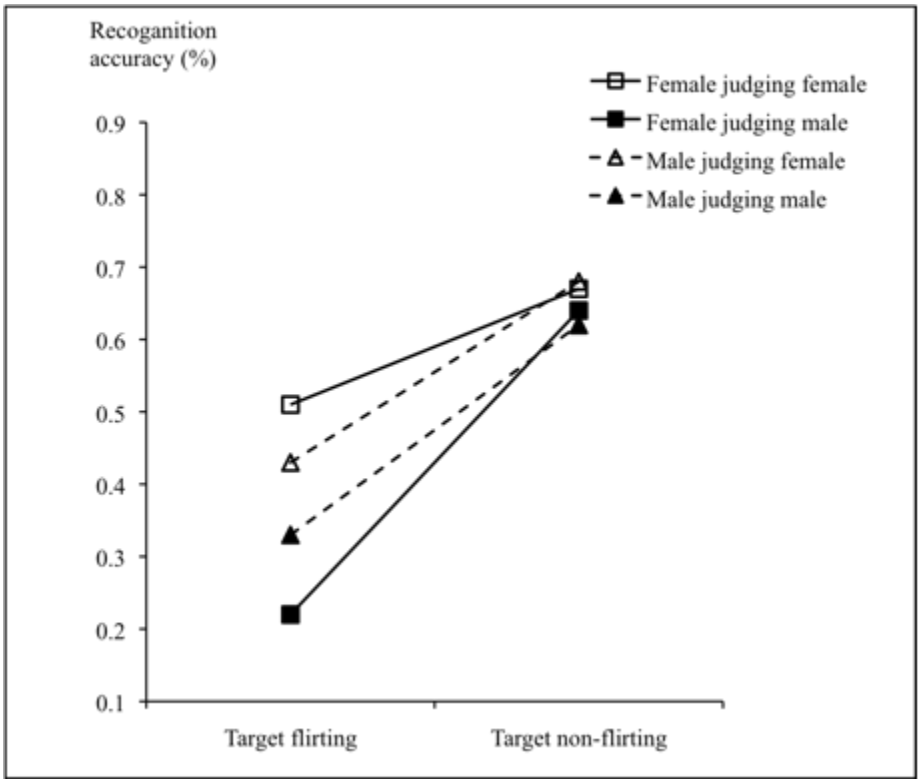

Figure 2: *Study 2 recognition accuracy of female and male third-party observers by target flirting versus non-flirting.*

Traditional sexual script

It was anticipated that traditional script consistent behaviors (i.e., males flirting and females not flirting) would be more accurately judged that script inconsistent behaviors. The results of both studies show mixed support for these predictions. The in-person participant observers in Study 1 and the third party observers in Study 2 were more accurate when the target was not flirting than when the target was flirting. This was found for both male and female targets (see Figures 1 and 2). This means that although males' flirting was not more accurately interpreted, females' lack of flirting *was* more accurately interpreted. This shows mixed support for the traditional script's role in detection accuracy. Perhaps the TSS does not inform accuracy judgments, or, at best, is only a useful heuristic for judging females' behavior but not males'

behavior. Overall, the results call into question the influence of the TSS in zero-acquaintance interactions. Men in Study 1 did not flirt more than females or perceive that their partners flirted at a greater rate, and male observers in Study 2 did not assume male targets were flirting at a greater rate than females or particularly often in an absolute sense. As a theoretical perspective, the TSS might overestimate the degree to which males' impressions and behaviors are sexually valenced.

Base rate

The most robust finding in the present investigation was the role of base rate on detection accuracy. The presence of a veracity effect renders the detection of deception much lower than the detection of truth (Levine et al., 2006). The present investigation offers strong support for a similar bias in detecting flirting. The results of Study 1, where participants in the interaction accurately decoded the absence of flirting 84% of the time, are consistent with the results of the experimental manipulation in Study 2: as the percent of videos with targets flirting increased from 0% to 100%, flirting detection accuracy rates fell from 61% to 42%. Both findings support prior research that suggests that a lack of interest is more easily interpreted than the presence of interest (Farris et al., 2008; Moore & Butler, 1989). Whether participating in the interaction or merely observing the interaction, observer accuracy rates improved by over 20 percentage points when individuals watched targets who were not flirting compared to targets who were flirting. This effect existed independent of participant sex, target sex, and observer demographic characteristics.

There are two possible reasons that the absence of flirting would be more easily detected than its presence. It is quite face threatening to flirt in an obvious manner with a potential partner one already knows (Kunkel et al., 2003). It is probably even more threatening to do so in a zero-acquaintance, video-recorded experimental setting. Therefore, any flirting that would transpire under such circumstances might be particularly ambiguous and protective of face, and, therefore, exceedingly difficult to decode. Secondly, the actual incidence of flirting in interactions with strangers or new acquaintances is quite low (Abbey, 1982; Saal et al., 1989). Therefore, individuals do not have much practice seeing it. In this context, individuals probably nearly no experience having their perceptions of target flirting explicitly confirmed by the target him/herself. By comparison, they have a great deal of experience interacting with others who have no romantic or sexual interest whatsoever.

Limitations and Directions for Future Research

Flirting was operationalized in the present study by a single binary item (i.e., "During the interaction, were you/ was this person flirting?"). This measure was used to facilitate the creation of the dependent measure of accuracy and simplify accuracy interpretations. Although this permitted comparisons to other studies using dichotomous dependent measures (i.e., Farris et al., 2008; Place et al., 2009), it limited the applicability of results to studies exploring interest in dating (i.e., Grammer et al., 2000) and using continuous measures of friendliness/flirtatiousness/seductiveness (i.e., Abbey, 1982). Socio-sexual communication can be broadly conceived in a variety of ways, including flirting, sexual intent, promiscuity/seductiveness, or romantic/dating interest (Lindgren et al., 2008). The lack of coordination and consistency in measuring these concepts is a limitation to this line of research as a whole (Lindgren et al., 2008), and a valid and reliable measurement of each of these concepts is an important and needed topic of future research.

The present research offers clear implications for the importance of context on flirting detection accuracy. Given the strong and consistent effect of base rate in accuracy, it appears that the normative amount of flirting in any given context may play a critical role in influencing the clarity of flirting and the effect of base rate on observer accuracy. In circumstances where courtship-relevant cues are not salient and behavioral norms steer toward appropriateness, flirtatious signals are not only infrequent, but they are probably very difficult to detect. Indeed, the appropriateness of pursuing a relationship within any given context influences perceptions of sexual interest for both men and women (Henningsen, Henningsen, & Valde, 2006). In contexts where courtship-relevant cues are salient, permissible or even encouraged, and potential partners are receptive to such signals (e.g., a bar, party, or speed-date), the detection of interest may be easier when it is present and the lack of interest less detectable (Place et al., 2009). Furthermore, because observers have no access to base rate information, observer accuracy is difficult to improve. If future research could establish reliable base rates of flirting in different contexts or identify others factors that predict base rate (i.e., drinking alcohol, prior acquaintanceship) it could practically improve flirting detection accuracy.

The perceptual biases related to the TSS might also be context dependent. According to the results of the present investigation, this bias is not a useful heuristic for accurately interpreting males and females' behavior in the context of a zero-acquaintance conversation.

Perhaps initial interactions are less scripted than courtship-salient contexts and present considerably less contextual pressure for males to quickly and obviously show their physical attraction. Simply, males are scripted to be the pursuer in courtship-relevant contexts, but are not similarly scripted in mundane get-to-know-you contexts, rendering males' flirting behavior less interpretable.

One final consideration for future research is a focus on individual differences in flirting behaviors and styles of communicating attraction (Hall et al., 2010). Deception research suggests that there are a few prolific liars (Serota et al., 2010) and that liars vary in terms of how often they are detected. Similarly, the present manuscript suggests that some individuals are more easily read when flirting than others, and some individuals are more accurate when detecting flirting than others. Factors beyond biological sex that influence the transparency of flirting, the ability to detect flirting, and the likelihood of flirting are all promising topics of future work.

Endnotes:

1. With a sample of $N = 104$ observations, $\chi^2(1) = .05$, and $p = .05$, a power analysis showed that the power to detect a small effect size ($d = .10$) was .175, and for a moderate effect size (d = .30) the power of .864.

2. With a sample of $N = 1557$ observations, $\alpha = .05$, and one-tailed significance tests, the power to detect the observed odds ratio of .02 exceeded .99. With a sample of $N = 780$ observations, $\alpha = .05$, and one-tailed significance tests, the power to detect observed odds ratios between .12 - .22 exceeded .99.

3. (a) Each resampled dataset included 104 observations (52 female, 52 male observers), (b) in each observation, target sex was opposite of the observer sex, (c) in the 52 female observer cases, 11 male targets flirted and 41 male targets did not flirt, and (d) in the 52 male observer cases, 14 female targets flirted and 38 female targets did not flirt (Table 2). With these four specifications, observations of Study 2 were randomly selected with replacement to create 1000 resampled datasets that matched the observed statistics in Study 1.

References

Abbey, A. (1982). Sex differences in attributions for friendly behavior: Do males misperceive females' friendliness? *Journal of Personality and Social Psychology, 42,* 830-838. doi:10.1037//0022-3514.42.5.830

de Weerth, C., & Kalma, A. (1995). Gender differences in awareness of courtship initiation tactics. *Sex Roles, 32,* 717-734. doi:10.1007/BF01560186

Eaton, A. A., & Rose, S. (2011). Has dating become more egalitarian? A 35 year review using *Sex Roles*. *Sex Roles, 64,* 843-862. doi:10.1007/s11199-011-9957-9

Farris, C., Treat, T. A., Viken, R. J., & McFall, R. M. (2008). Perceptual mechanisms that characterize gender differences in decoding women's sexual intent. *Psychological Science, 19,* 348-354. doi:10.1111/j.1467-9280.2008.02092.x

Gagnon, J. H. (1990). The explicit and implicit use of the scripting perspective in sex research. *Annual Review of Sex Research, 1,* 1–43.

Grammer, K., Kruck, K., Juette, A., & Fink, B. (2000). Nonverbal behavior as courtship signals: The role of control and choice in selecting partners. *Evolution and Human Behavior, 21,* 371-390. doi:10.1016/S1090-5138(00)00053-2

Hall, J. A., Carter, S., Cody, M. J., & Albright, J. M. (2010). Individual differences in the communication of romantic interest: Development of the flirting styles inventory. *Communication Quarterly, 58,* 365-393. doi:10.1080/01463373.2010.524874

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology, 78,* 81-91. doi:10.1037//0022-3514.78.1.81

Henningsen, D. D. (2004). Flirting with meaning: An examination of miscommunication in flirting interactions. *Sex Roles, 50,* 481-489. doi:10.1023/B:SERS.0000023068.49352.4

Henningsen, D. D., Henningsen, M. L. M., & Valde, K. S. (2006). Gender differences in perceptions of women's sexual interest during cross-sex interactions: An application and extension of cognitive valence theory. *Sex Roles, 54,* 821-829. doi:10.1007/s11199-006-9050-y

Impett, E. A. & Peplau, L. A. (2003). Sexual compliance: Gender, motivational, and relationship perspectives. *Journal of Sex Research, 40,* 87-100. doi:10.1080/00224490309552169

Koenig, B. L., Kirkpatrick, L. A., & Ketelaar, T. (2007). Misperception of sexual and romantic

interests in opposite-sex friendships: Four hypotheses, *Personal Relationships, 14,* 411-429. doi:10.1111/j.1475-6811.2007.00163.x.

Kunkel, A. D., Wilson, S. R., Olufowote, J., & Robson, S. (2003). Identifying implications of influence goals: Initiating, intensifying, and ending romantic relationships. *Western Journal of Communication, 67,* 382-412.

Kurzban, R., & Weeden, J. (2005). HurryDate: Mate preferences in action. *Evolution and Human Behavior, 26,* 227-244. doi:10.1016/j.evolhumbehav.2004.08.012

La France, B. H. (2010). What verbal and nonverbal communication cues lead to sex? An analysis of the traditional sexual script. *Communication Quarterly, 58,* 297-318. doi:10.1080/01463373.2010.503161

La France, B. H., Henningsen, D. D., Oates, A., & Shaw, C. M. (2009). Social-sexual interactions? Meta-analyses of sex differences in perceptions of flirtatiousness, seductiveness, and promiscuousness. *Communication Monographs, 76,* 263-285. doi:10.1080/03637750903074701

Laner, M. R., & Ventrone, N. A. (1998). Egalitarian daters/traditional dates. *Journal of Family Issues, 19,* 468-764. doi:10.1177/019251398019004005

Levine, T. R., Kim, R. K., Park, H. S., & Hughes, M. (2006). Deception detection accuracy is a predictable linear function of message veracity base-rate: A formal test of Park and Levine's Probability Model. *Communication Monographs, 73,* 243-260. doi:10.1080/03637750600873736

Lindgren, K. P., Parkhill, M. R., George, W. H., & Hendershot, C. S. (2008). Gender differences in perceptions of sexual intent: A qualitative review and integration. *Psychology of Women Quarterly, 32,* 423-439. doi:10.1111/j.1471-6402.2008.00456.x

Matsumoto, D. (2006). Culture and nonverbal behavior. In V. Manusov & M. L. Patterson (Eds.), *Handbook of nonverbal communication* (pp. 219-236). Thousands Oaks, CA: Sage Publications.

McCroskey, J. C., & McCain, T. A. (1974). The measurement of interpersonal attraction. *Speech Monographs, 41,* 261-266.

Moore, M., & Butler, D. (1989). Predictive aspects of nonverbal courtship behavior in women. *Semiotica, 3,* 205-215. doi:10.1515/semi.1989.76.3-4.205

Place, S., Todd, P. M., Penke, L., & Asendorpf, J. B. (2009). The ability to judge the romantic

interest of others. *Psychological Science, 20*, 22-26. doi:10.1111/j.1467-9280.2008.02248.x

Saal, F. E., Johnson, C. B., & Weber, N. (1989). Friendly or sexy? It may depend on who you ask. *Psychology of Women Quarterly, 13*, 263-276. doi:10.1111/j.1471-6402.1989.tb01001.x

Schmitt, D. P. (2005). Fundamentals of human mating strategies. In D. M. Buss (Ed.), *The evolutionary psychology handbook* (pp. 258-291). New York, NY: Wiley.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36*, 2-25. doi:10.1111/j.1468-2958.2009.01366.x

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Los Angles, CA: Sage.