

ACASI Gender-of-Interviewer Voice Effects on Reports to Questions about Sensitive Behaviors among Young Adults

Jennifer Dykema¹, Jessica Price¹, Kerryann DiLoreto¹, Eric White¹,
Nora Cate Schaeffer^{1,2}

¹University of Wisconsin Survey Center, 1800 University Ave., Madison, WI 53726

²Department of Sociology, University of Wisconsin-Madison, 8128 William H. Sewell
Social Science Building, 1180 Observatory Drive, Madison, WI 53706

Abstract

While previous research indicates that audio computer-assisted self-interviewing (ACASI) yields higher reports of threatening behaviors, very few studies have examined the potential effect of the gender of the ACASI voice on survey reports. In this study we examined gender-of-voice effects for a set of sensitive questions administered via ACASI as part of the first wave of data collection of the Midwest Young Adult Study. Respondents included 766 foster care youth who were asked to participate in an in-person study about their experiences while wards of several states in the Midwest. Respondents were randomly assigned to listen to (while simultaneously viewing) questions read by a female or male voice. Gender of the ACASI voice was crossed with the respondent's gender to test for interaction effects. Questions asked about delinquent, violent, and illegal behaviors. The analysis explored the effects of the voice's gender on disclosure of the sensitive behaviors. Results showed higher reporting among male respondents when responding to a female voice. Male respondents were also more likely to report consistently about fighting when the voice was female. Taken together, these findings indicate male respondents were potentially more accurate when reporting to the female voice. Reports by female respondents were not influenced by the gender of the ACASI voice. Our analysis adds to the relatively small body of research on gender-of-voice effects in surveys, with important findings on measuring sensitive behaviors using ACASI with young adults.

Key Words: ACASI, voice, gender, interviewers, threatening or sensitive questions, youth, interviewer effects

1. Introduction

Audio computer-assisted self-interviewing (or ACASI) – in which respondents listen and respond to pre-recorded questions during an in-person interview – has become the preferred method for administering sensitive questions in face-to-face interviews. Several studies demonstrate that compared to paper-and-pencil questionnaires, or other forms of computerized interviewing such as CAPI or CASI, ACASI yields higher reporting of sensitive behaviors (Dykema, Basson, & Schaeffer, 2007; Tourangeau & Smith, 1996; Turner, Ku, Rogers, Lindberg, Pleck, & Sonenstein, 1998a). In spite of the wide-spread use of ACASI, little research has explored how operational characteristics, including the gender of the ACASI voice, affect a respondent's propensity to disclose personal or

potentially threatening information. Issues of concern in interviewer-administered surveys, such as whether to use same- or opposite-gender interviewers, or allow respondents to choose, are also relevant for ACASI. The purpose of the present study was to evaluate the impact of the gender of the ACASI voice on responses to questions about threatening behaviors among young adults at high-risk for engaging in these behaviors.

1.1 Studies Evaluating Gender Effects with Live Interviewers

Interviewer effects occur when interviewers influence survey responses through their behaviors (e.g., how they read the survey questions) or external characteristics (e.g., race, gender, and age). Interviewer effects may vary based on the topic of a question or the population under study. Schaeffer, Dykema, & Maynard (2010:451) note, “An interviewer characteristic is most likely to affect responses to questions that make the characteristic salient or relevant in the interaction, activate stereotypes, or evoke the respondent’s concerns with affiliation, relative status, or deference.” In contrast to race or age, an interviewer’s gender is likely to be readily identifiable and accurately inferred based on the voice. Depending on a question’s topic, respondents may refer to gender-based stereotypes or conversational norms when providing a response (Tannen, 1996). For example, respondents may disclose or report higher levels of symptoms or sensitive behaviors to a female interviewer based on a stereotype that women are more sympathetic (Pollner, 1998), or nonjudgmental (Nass, Robles, Heenan, Bienstock, & Treinen, 2003).

Overall, however, the gender of the interviewer has had only limited effects in the handful of studies in which it has been examined (Davis, Couper, Janz, Caldwell, & Resnicow, 2010; Kane & Macaulay, 1993; Schaeffer et al., 2010). For questions that specifically ask about sensitive or threatening behaviors, findings are mixed with regard to whether respondents report more to female or male interviewers, or whether it depends on the pairing by gender of interviewers and respondents (Davis et al., 2010). Of the studies reviewed, most reported that disclosure among female respondents was greater when the interviewer was female. This was evidenced in reports about sexual abuse (Dailey & Claus, 2001); sexual behaviors (Catania, Binson, Canchola, Pollack, & Hauck, 1996); alcohol use (Johnson & Parsons, 1994); and symptoms of depression, substance abuse, and conduct disorders (Pollner, 1998). The only exception was that Johnson and Parsons (1994) reported that homeless female respondents revealed more about substance abuse to male interviewers.

The behavior of male respondents has followed a slightly different pattern. Although male respondents disclosed more threatening behaviors when reporting to female interviewers about having been sexually abused (Dailey & Claus, 2001), or about symptoms of depression, substance abuse, and conduct disorders (Pollner, 1998), they disclosed more to male interviewers when reporting about sensitive and often risky behaviors including sexual behaviors (e.g., number of lifetime partners, extramarital and casual sex, sexual violence, condom disuse) (Catania et al., 1996; Wilson, Brown, Mejia, & Lavori, 2002), substance use (Fendrich, Johnson, Shaligram, & Wislar, 1999; Johnson & Parsons, 1994), and alcohol use (Johnson & Parsons, 1994). Several studies also demonstrated that gender-of-interviewer effects were more pronounced among men than women (e.g., Catania et al., 1996; Pollner, 1998) such that male respondents may be more influenced by an interviewer’s gender.

1.2 Studies Evaluating Gender Effects for Recorded Voices

Very few studies have examined gender effects for recorded voices. Two early studies evaluated gender-of-voice effects using ACASI. Rogers, Miller, Forsyth, Smith, & Turner (1996) recorded questions about sexual behaviors and drug and alcohol use using a male or female voice, and randomly assigned subjects to one voice. They reported that the gender of the voice had no effect on data quality, which was measured by the proportion of respondents who listened to questions read in their entirety before recording a response. In developmental work to test the feasibility of ACASI for the National Survey of Adolescent Males (Turner, Forsyth, O'Reilly, Cooley, Smith, Rogers, & Miller, 1998b), researchers varied whether questions about sexual behaviors were read by a male or one of two female voices. After completing the survey, respondents were queried about their reactions to the ACASI voice, and most could not remember the gender of the voice. Based on these studies, Turner et al. (1998b:470) concluded that "even in sex surveys, the gender of the voice is unimportant." However, neither study directly tested the effect of the gender of the voice on the answers respondents gave to the questions.

Only one study, conducted exclusively with male respondents, directly tested the effect of the ACASI voice's gender on survey reports. Fahrney, Uhrig, & Kuo (2010) explored the impact of a male versus female voice on reports of sexual activity among a sample of men classified as MSM (men who have sex with men). The male respondents provided higher reports of unprotected receptive anal sex with HIV-status-unknown partners but slightly lower reports of HIV-negative partners in the past month to a female voice. Having anal sex with partners whose HIV status is unknown is considered a risky behaviour, which respondents would be expected to underreport (Catania et al., 1996). In contrast, because men tend to over-report sexual partners (Laumann, Gagnon, Michael, & Michaels, 1994), lower reports of sexual partners who are known to be HIV-negative by MSM could be considered more accurate. Thus, the findings from this study offer evidence that is consistent with more accurate reporting among the male respondents who heard questions read by a female voice.

More recent studies have evaluated the effect of the gender of voice on disclosure of sensitive behaviors using interactive voice response (IVR) methodology in which respondents in telephone interviews listen to questions read by a recorded voice, and enter responses using the phone's keypad. Within the context of an RDD telephone survey, Tourangeau, Couper, & Steiger (2003) tested voice effects for a male, female, and male-female combined condition in which questions were read using a male voice but response categories were read by a female voice. They found no effects of the gender of the voice on reports about sensitive topics including drugs, sexual behaviors, voting, and church attendance. Similarly, Couper, Singer, & Tourangeau (2004) reported that the voice's gender had no impact on reports about several sensitive behaviors in their sample of adults drawn from voting and bankruptcy records or among respondents drawn from a general list sample. In contrast, Nass et al. (2003) reported that male and female college students were more likely to report engaging in sensitive behaviors when the recorded IRV voice was female.

1.3 The Current Study

The purpose of the current study was to examine the effect of the gender of the ACASI voice on reports of several threatening behaviors among young adults. We examined levels of disclosure by looking at the proportion of respondents who reported engaging in

various threatening behaviors. The threatening behaviors in the survey included items about fighting and criminal behavior. Criminal behavior included both violent and non-violent offenses, and questions asked about incidents that the respondent may have witnessed, committed, or been victim to. Based on previous research concerning voice effects when questions were read by a live interviewer or a recorded voice, we generated three predictions. First, gender of voice has not demonstrated strong or consistent effects using either a live interviewer or a recorded voice, and we did not expect to find large or consistent effects here either. Second, if we did find effects, we expected that female respondents would be more likely to disclose threatening behaviors to a female interviewer, but that male respondents would be more likely to report threatening behaviors to a male interviewer. Third, we expected that male respondents would be more affected by the voice manipulation than female respondents such that any gender-of-voice effects would be larger for males than females. To the best of our knowledge, this is the first study to directly test what effect voice gender has on survey reports among both male and female respondents, and one of only two studies that directly examines the effect of the gender of voice on survey reports using ACASI.

2. Methods

2.1 Data

Data were provided by the Midwest Evaluation of the Adult Functioning of Former Foster Youth (referred to as the Midwest Young Adult Study), a longitudinal study in which computer-assisted personal-interviews (CAPI) have been conducted with eligible youth since 2002 (Courtney, Terao, & Bost, 2004). The goals of the study were to collect data on adult self-sufficiency among youth exposed to foster care, and to measure the kinds of services these youth receive. Questions asked about living situations, relationships with foster care families, educational histories, preparation for independent living, engagement in risky behaviors, and future plans. Youth in Illinois, Iowa, and Wisconsin were identified as eligible for the study if they were in out-of-home care supervised by a public child welfare agency and between the ages of 17 and 17 ½, and had been in the care of the state for one or more years prior to their 17th birthday. All youth who met the eligibility criteria in Iowa and Wisconsin were included in the sample. In Illinois, a sample was drawn because the state contained more eligible youth than funds permitted interviewing. Of the 957 adolescents deemed eligible for the study, interviews were conducted with 776 youth, for an overall response rate of 81% (AAPOR RR1) (AAPOR 2009). The study was sponsored by the public child welfare agencies in Illinois, Iowa, and Wisconsin, and the Chapin Hall Center for Children at the University of Chicago. Interviews were conducted by the University of Wisconsin Survey Center (UWSC). Our analysis focused on reports from 766 respondents – 384 males and 382 females – collected during Wave 1 in 2002-2003.¹

2.2 Experiment

In the CAPI instrument a section of potentially sensitive questions was administered using ACASI. The questions and response categories were pre-recorded and loaded into the instrument as sound (.wav) files. For each question, respondents were presented with

¹ Ten respondents were dropped from the analysis. Two of these were coded as partially completed interviews who did not answer any of the questions in the ACASI module. The remaining eight were missing values for the experimental variable.

a screen containing both the question text and response categories. Each of these was also read in their entirety through headphones.² Questions in the ACASI module were recorded using a female and a male interviewer in their early twenties. At the time of the interview, respondents were randomly and automatically assigned to hear either the female or male voice. Once the random selection was made, the respondent heard either the male or the female voice for the duration of the module.

2.3 Measures

Questions in the ACASI module either came directly from or were adapted from The National Longitudinal Study of Adolescent Health (Add Health). Questions asked about respondents' levels of engagement in various delinquent, violent, and illegal behaviors and their sexual orientation; female respondents also received questions about menstruation and pregnancy. We examined the effect of the gender of the voice on reports to the 28 behavioral frequency questions asked of both male and female respondents (see Figure 1).³

For the analysis of the individual questions, we examined the effect of the gender of voice on the proportion of respondents who reported engaging in the behavior. Because questions used various response categories, we dichotomized responses and coded answers as "1" to indicate the respondent had engaged in the behavior (versus "0" if they reported never engaging in the behaviour or engaging zero times).⁴ Gender of the ACASI voice was coded "1" for female versus "0" for male. Using the indicators from the individual questions, we also created an index that tallied the number of times the respondent engaged in each of the behaviors in order to test the overall effect of voice across questions (Nass et al., 2003; Tourangeau et al., 2003).

Although higher reports of threatening behaviors are typically interpreted as more accurate, we desired other techniques for determining whether the female or male voice was associated with better reporting. Four questions asked explicitly about fighting including: taking part in a group fight, got into a physical fight, getting into a serious physical fight, and times in a physical fight treated by a doctor or nurse. We examined the pattern of reporting among the four items in order to make assessments about how consistently respondents reported about these behaviors. For example, the questions

² Respondents were not able to adjust the volume of the audio without assistance from the interviewer. Default settings on the laptops prevented interviewers from completely muting the volume of the ACASI module. Further, respondents were not able to enter an answer until the audio file for a given question was completed. Therefore, respondents listened to the full audio file for each question and response category set.

³ All 766 respondents provided complete data to the 28 questions examined here.

⁴ Respondents were coded as having engaged in the behavior if they answered: "1 or 2 times" to "5 or more times" for the questions that included the grouped response categories; "yes" for yes-no questions; 1 or more days for the question about carrying a weapon; and 1 or more times for the question about physical fights that had to be treated by a doctor or nurse. Respondents who reported that they "never" fought in response to a question asking with whom they fought the last time they were in a physical fight were skipped past the question about physical fights that had to be treated by a doctor or nurse. They are grouped with respondents who answered "0" times to the question about physical fights that had to be treated by a doctor or nurse.

Figure 1: Wording of questions included in the analysis, Midwest Young Adult Study, 2002-2003

Questions using the response categories: never, 1 or 2 times, 3 or 4 times, and 5 or more times

In the past 12 months, how often did you paint graffiti or signs on someone else's property or in a public place?

In the past 12 months, how often did you deliberately damage property that didn't belong to you?

In the past 12 months, how often did you lie to your parents or guardians about where you had been or whom you were with?

How often did you take something from a store without paying for it?

How often did you get into a serious physical fight?

How often did you hurt someone badly enough to need bandages or care from a doctor or nurse?

How often did you run away from home?

How often did you drive a car without its owner's permission?

In the past 12 months, how often did you steal something worth more than \$50?

How often did you go into a house or building to steal something?

How often did you use or threaten to use a weapon to get something from someone?

How often did you sell marijuana or other drugs?

How often did you steal something worth less than \$50?

In the past 12 months, how often did you take part in a fight where a group of your friends was against another group?

How often were you loud, rowdy, or unruly in a public place?

During the past 12 months, how often did each of the following things happen? You saw someone shoot or stab another person.

Someone pulled a knife or gun on you.

Someone shot you.

Someone cut or stabbed you.

You got into a physical fight.

You were jumped.

You pulled a knife or gun on someone.

You shot or stabbed someone.

Questions using the response categories: yes and no

Have you ever been arrested?

Have you ever been convicted of a crime?

Have you ever spent a night in jail, prison, juvenile hall, or other correctional facility?

Question using the response categories: none, 1 day, 2 or 3 days, 4 or 5 days, 6 or more days

During the past 30 days, on how many days did you carry a weapon-such as a gun, knife, or club-to school?

Question using an open-ended response format

During the past 12 months, how many times were you in a physical fight in which you were injured and had to be treated by a doctor or nurse?

imply different levels of severity such that a respondent's answer would be inconsistent if he reported not engaging in a less serious incident of the behavior (e.g., got in a physical fight) but then reported affirmatively to a more serious incident of the behaviour (e.g., got into a serious physical fight). We coded responses as consistent (with a value equal to "1" in the analysis) versus inconsistent (with a value equal to "0") as shown in Table 1.

Table 1. Coding for consistency in reporting about fighting, Midwest Young Adult Study, 2002-2003

Take part in group fight	Got in physical fight	Get into serious physical fight	Fights treated by doctor/nurse	Consistency of responses	Sample size (n)
Never	Never	Never	0 times	Consistent	168
Never	Never	Never	1 or more	Inconsistent	1
Never	Never	1 or more	0 times	Inconsistent	47
Never	Never	1 or more	1 or more	Inconsistent	4
Never	1 or more	Never	0 times	Consistent	40
Never	1 or more	Never	1 or more	Inconsistent	3
Never	1 or more	1 or more	0 times	Consistent	201
Never	1 or more	1 or more	1 or more	Consistent	28
1 or more	Never	Never	0 times	Inconsistent	10
1 or more	Never	Never	1 or more	Inconsistent	1
1 or more	Never	1 or more	0 times	Inconsistent	8
1 or more	Never	1 or more	1 or more	Inconsistent	1
1 or more	1 or more	Never	0 times	Consistent	13
1 or more	1 or more	1 or more	0 times	Consistent	175
1 or more	1 or more	1 or more	1 or more	Consistent	66

2.4 Analysis

For the analysis of the individual items, we regressed the indicator for whether the respondent reported engaging in the behavior on the gender of the voice. Because our dependent variables are dichotomous, we used logistic regression and report the odds ratios to test for significant differences. Models were estimated separately for males and females because levels of engaging in the behaviors varied considerably by gender, and we expected the effects of gender of voice to vary by the respondent's gender (Catania et al., 1996; Mensch, Hewett, & Erulkar, 2003). To assess the effect of the voice's gender on a count of the number of behaviors summed into an index, we performed a two-way ANOVA with gender of the voice and gender of the respondent as between-subject factors. For the analysis of consistency in reporting about fighting, we regressed consistency in reporting about fighting on the gender of the voice. We examined effects among the full sample consisting of all respondents, and also for the subsample of respondents who reported engaging in at least one of the fighting behaviors.

3. Results

3.1 Analysis of Survey Responses

Results for the analysis of the individual items are presented in Table 2. As shown in the table, the rate at which the behaviors were reported varied dramatically between male and female respondents and across the behaviors. Among male respondents, reporting to a female voice versus a male voice increased the odds of reporting about getting into a serious fight (OR = 1.66, $p < .05$), going into a building to steal (OR = 1.59, $p < .10$), threatening to use a weapon (OR = 1.64, $p < .10$), selling drugs (OR = 1.66, $p < .05$), spending a night in jail (OR = 1.50, $p < .05$), and getting into a physical fight (OR = 1.61,

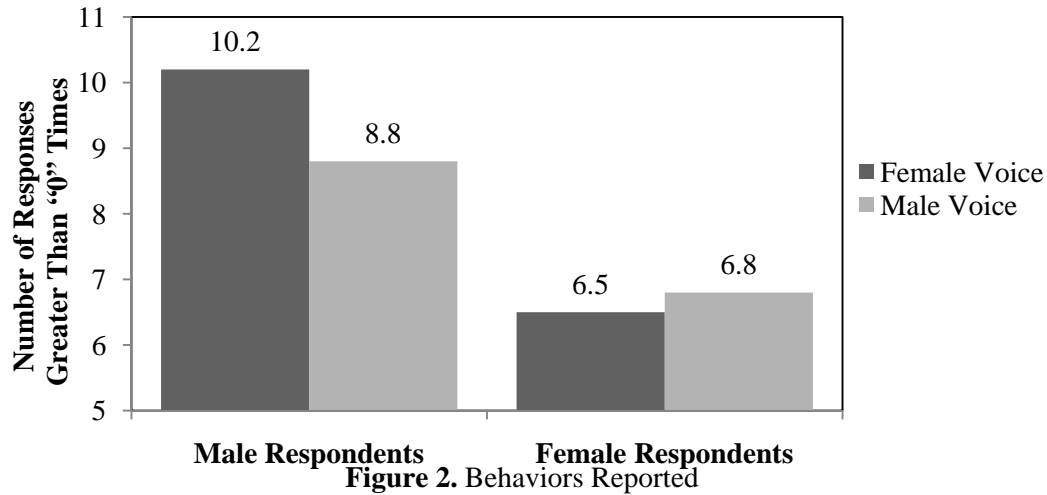
$p < .05$). Male respondents were consistently more likely to report engaging in behaviors to the female voice; this occurred for 27 of the 28 behaviors. (The exception was painting graffiti). Among female respondents, only reports about shooting or stabbing someone differed significantly between the female and male voices such that disclosure was higher when females reported to a female voice ($OR = 2.48, p < .10$). There was no discernable trend in reporting among the female respondents.

Table 2. Percent reporting ever engaging in the behavior and odds ratios by gender of voice for male and female respondents, Midwest Young Adult Study, 2002-2003

Question Label	Male Respondents				Female Respondents			
	Female Voice	Male Voice	Odds Ratio	95% CI	Female Voice	Male Voice	Odds Ratio	95% CI
Paint graffiti	11.1	11.4	0.97	0.51-1.83	4.6	6.7	0.67	0.27-1.63
Damage property	30.7	24.3	1.38	0.88-2.16	13.2	19.2	0.64	0.37-1.12
Lie to parents/guardians	66.8	63.8	1.14	0.75-1.74	64.9	61.5	1.16	0.76-1.76
Take something from store	51.8	44.9	1.32	0.88-1.97	36.8	38.5	0.93	0.61-1.41
Get into serious fight	78.4	68.6	1.66*	1.05-2.62	62.6	66.3	0.85	0.56-1.30
Hurt someone	54.3	45.9	1.40	0.93-2.09	28.7	27.4	1.07	0.68-1.67
Run away	43.2	43.2	1.00	0.67-1.50	44.8	47.1	0.91	0.61-1.37
Drive without permission	20.6	20.0	1.04	0.63-1.71	11.5	15.4	0.71	0.39-1.30
Steal something GT \$50	18.1	12.4	1.56	0.88-2.74	5.7	5.3	1.09	0.45-2.64
Go into building to steal	22.1	15.1	1.59+	0.94-2.69	5.2	7.7	0.65	0.28-1.52
Threaten to use weapon	19.6	13.0	1.64+	0.94-2.84	4.0	7.7	0.50	0.20-1.25
Sell marijuana/drugs	36.2	25.4	1.66*	1.07-2.58	11.5	16.8	0.64	0.36-1.16
Steal something LT \$50	46.2	42.7	1.15	0.77-1.73	31.6	33.2	0.93	0.61-1.43
Take part in group fight	46.2	38.9	1.35	0.90-2.03	31.6	26.4	1.29	0.82-2.00
Be in a loud public place	49.7	47.0	1.12	0.75-1.66	52.3	47.1	1.23	0.82-1.84
Been arrested	65.3	60.5	1.23	0.81-1.86	42.5	42.3	1.01	0.67-1.52
Convicted of a crime	35.2	28.6	1.35	0.88-2.08	15.5	16.3	0.94	0.54-1.63
Night in jail	52.3	42.2	1.50*	1.00-2.25	24.1	27.4	0.84	0.53-1.34
Saw someone shoot/stab	28.1	22.2	1.38	0.86-2.19	16.1	17.3	0.92	0.53-1.57
Knife/gun pulled on R	37.7	36.8	1.04	0.69-1.57	18.4	18.8	0.98	0.58-1.64
R shot by someone	8.0	4.3	1.93	0.81-4.63	0.6	0.5	1.20	0.07-19.27
R stabbed by someone	21.6	20.0	1.10	0.67-1.81	10.3	9.6	1.08	0.55-2.12
Got in physical fight	77.9	68.6	1.61*	1.02-2.54	59.8	67.3	0.72	0.47-1.10
R was jumped	45.2	42.2	1.13	0.76-1.70	20.7	23.1	0.87	0.53-1.42
R pulled knife/gun	19.6	15.7	1.31	0.77-2.22	8.6	8.7	1.00	0.49-2.04
R shot/stabbed someone	9.5	5.4	1.85	0.84-4.08	5.7	2.4	2.48+	0.83-7.39
Carried weapon to school	7.5	4.9	1.59	0.68-3.74	5.2	4.8	1.08	0.43-2.72
Fight treated by doctor	15.6	12.4	1.30	0.73-2.32	13.2	13.0	1.02	0.56-1.85

Note: * $p < .05$, + $p < .10$.

Comparing the overall number of behaviors reported using the index, there was a significant interaction between the gender of the voice and the gender of the participant, $F(1, 762) = 4.82, p < .05$ (see Figure 2). Post-hoc analyses (ANOVA) indicated that male respondents reported engaging in significantly more behaviors when reporting to a female voice, $F(1, 382) = 5.28, p < .05$, but the gender of the voice did not affect disclosure among female respondents, $F(1, 380) = 0.37, p > .3$. There was also a main effect for gender, $F(1, 762) = 58.08, p < .001$, with male respondents reporting engaging in significantly more behaviors than female respondents.



3.2 Analysis of Consistency in Reporting about Fighting

Table 3 presents results for the analysis of consistency in reporting about fighting. The first row of the table presents results for the full sample of respondents. Among male respondents, reporting to the female voice increased the odds of reporting consistently but the effect was not significant; the reverse pattern was shown for female respondents but was not significant. The second row in the table shows results when the sample is limited to only those respondents who reported engaging in at least one of the fighting behaviors. Among male respondents, the odds of reporting consistently were higher (OR = 1.82, $p < .10$) when reporting to the female voice. The reverse pattern was again demonstrated by female respondents reporting to the female voice but the effect was not significant

Table 3. Percent reporting consistently about fighting and odds ratios by gender of voice for male and female respondents, Midwest Young Adult Study, 2002-2003

	Male Respondents				Female Respondents			
	Female Voice	Male Voice	Odds Ratio	95% CI	Female Voice	Male Voice	Odds Ratio	95% CI
All respondents (N=776)	92.5	88.1	1.66	0.83-3.30	87.9	91.8	0.65	0.33-1.27
Only those who reported fighting 1 or more times (N=598)	91.2	85.0	1.82+	0.91-3.65	83.1	89.2	0.60	0.30-1.19

Note: + $p < .10$.

4. Discussion

Findings from our study contradicted our expectation that we would find few and inconsistent effects for reports of threatening behaviors based on the gender of the ACASI voice. Among male respondents, of the 28 behaviors examined, 27 showed a consistent trend of higher reports when male respondents provided answers to the female voice. The effects were significant ($p < .05$) or marginally significant ($p < .10$) for 6 behaviors (i.e., getting into a serious fight, going into a building to steal, threatening to use a weapon, selling marijuana or other drugs, spending a night in jail, and getting into a physical fight). Further, a test of the overall effect of the voice's gender across questions

indicated a significant interaction between the gender of the respondent and the gender of the voice such that male respondents reported engaging in more behaviors, on average, to the female voice than to the male voice. Thus, while our expectation that male respondents would be more likely to report to a male interviewer received no support in our experiment, we found support for the expectation that male respondents would be more affected by the voice manipulation than female respondents. In contrast, our expectation that female respondents would be more likely to disclose threatening behaviors to a female interviewer was unfounded. Among female respondents, the gender of the ACASI voice had virtually no effect on reports of engaging in the sensitive behaviors. Like Catania et al. (1996), we only found gender-of-voice effects for male respondents. These authors argued that among women there may be ceiling effects for reporting about sensitive behaviors such that because women tend to disclose more in general, it may be hard to detect differential reporting among female respondents based on gender of the interviewer's voice.

We offer several explanations to explain the gender-of-voice effects we found among male respondents. These are speculative as we do not have data to test the explanations directly. Self-disclosure theory posits that people will be more honest and disclose more information to individuals with whom they trust and feel comfortable (e.g., Catania et al., 1996; Jourard, 1971). Research in law enforcement and criminal justice finds that female interrogators may be more effective in eliciting information about criminal activity from males (Hunt, 1984; Weisel, 2002). At the time of the study, many male respondents already had encounters with law enforcement personnel or spent at least one night in jail. Extrapolating from these circumstances, it is conceivable that the young, at-risk male respondents in our experiment may have ascribed greater authority to the male ACASI voice which in turn caused them to be wary about reporting engaging in the delinquent, violent, and criminal behaviors for fear of punishment.

In contrast to self-disclosure theory which holds that when people disclose they are providing more honest and accurate information, other researchers have offered what we collectively call "explanations for exaggeration." Exaggeration hypotheses hold that the tendency for males to report higher levels of engaging in sensitive behaviors may be associated with more dishonest and less valid reporting. For example, to explain higher reports by male respondents to male interviewers about some sexual behaviors, (Catania et al., 1996:371) described the "macho hypothesis." The macho hypothesis states that men report more to men in order to seem more virile and manly. We find no support for the hypothesis in our study. Weisel (2002:102) has argued that "a female interviewer may inadvertently encourage male interviewees to put on a macho bravado and exaggerate some points" in her study of contemporary gangs. However, consistent with other research in this area (e.g., Turner et al., 1998a:871), we have taken the approach that higher reports of the behaviors indicate better data quality. This is based on the belief that because sensitive and threatening behaviors are underreported, higher reporting produced under one experimental condition is likely to be more accurate. Consistent with the assumption that "more is better," Kreuter, Presser, & Tourangeau (2008) reported that higher levels of reporting for socially undesirable behaviours and characteristics were associated with more accurate reports when survey responses were compared to validation data. In our data we examined the impact of the gender of the voice on consistency in reporting about fighting. Among male respondents the odds of reporting consistently were 82% higher when responding to a female voice. We take this as evidence of more accurate reporting to the female voice. Thus, not only did the male respondents report more to the female voice, their responses are also more consistent.

In contrast to self-disclosure theory or exaggeration hypotheses, the effects we report here could also be explained by a more cognitive and less psychological model. It may be that the female voice somehow “feminized” the behaviors under study such that respondents were lead to consider less extreme instances of the behavior when reporting. This could have decreased the threshold for reporting about the behaviors and increased the likelihood that respondents would disclose to the female voice. (So, for example, when asked about the frequency of engaging in serious fights, respondents might have considered less extreme occurrences of serious fighting.) Evidence for this is provided by the current study in that while the male respondents who heard a female voice were more likely to report engaging in the behaviors, the gender of the voice did not affect the mean number of times they reported engaging in any of the behaviors (results not shown). This explanation does not explain, however, why the gender of the voice had no effect on female respondents.

This study was limited in several regards. First, our respondents were youth in foster care. These respondents are more likely to have engaged in the high-risk behaviors under study here than other respondents. The high incidence of engaging in the behaviors, however, probably increased our ability to detect statistically significant differences. Our results may not be generalizable to older respondents or respondents who are less likely to engage in delinquent, violent, and illegal behaviors. Second, we only isolated one characteristic of the ACASI voice to test – the gender. We have no way of knowing what impact other social characteristics, such as age, race, or socioeconomic status, or other voice characteristics such as tone, pitch, or pace might have had. More research is needed to disentangle the effects of gender from other characteristics. Third, consistent with other research in this area, we have taken the approach that higher reports of the behaviors indicate better data quality but we do not have external data to use to validate reports.

The vast majority of studies employing ACASI methods use either a female or synthetic voice (Couper et al., 2004). Results from our study offer further support for this convention. As noted by Fahrney et al. (2010) however, studies of MSM tend to match male respondents with male ACASI voices under the belief that this will promote more accurate reporting (see Wolitski, Parsons, Gomez, Purcell, Hoff, & Halkitis, 2005). Both their findings and our results suggest this may not be effective and may lead to underreporting of behaviors. More experimental research is needed to explore how gender-of-ACASI-voice effects might vary in different populations and across different types of questions and topics. Coupling methodological inquiries with data that can be used to validate responses to determine if higher reports are more accurate is critical.

Acknowledgements

Research support was provided by the University of Wisconsin Survey Center (UWSC) at the University of Wisconsin-Madison, which is supported by the College and the Graduate School. We wish to thank Mick Couper for his input during the early development of the experimental design, Mark Courtney for his permission to implement the gender-of-voice experiment, and John Stevenson for providing comments on an earlier draft. A preliminary version of this paper was presented at the 2010 meeting of the

American Association for Public Opinion Research. Opinions expressed here are those of the authors and do not necessarily reflect those of the sponsors or related organizations.

References

- The American Association for Public Opinion Research. 2009. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 6th edition. AAPOR.
- Catania, J. A., D. Binson, J. Canchola, L. M. Pollack, and W. Hauck. 1996. Effects of interviewer gender, interviewer choice, and item wording on responses to questions concerning sexual behaviour. *Public Opinion Quarterly*, 60: 345-375.
- Couper, M., E. Singer, and R. Tourangeau. 2004. Does voice matter? An interactive voice response (IVR) experiment. *Journal of Official Statistics*, 20: 551-570.
- Courtney, M. E., S. Terao, and N. Bost. 2004. Midwest Evaluation of the Adult Functioning of Former Foster Youth: conditions of youth preparing to leave state care. Chapin Hall Center for Children. University of Chicago.
- Dailey, R. M., and R. E. Claus. 2001. The relationship between interviewer characteristics and physical and sexual abuse disclosures among substance users: a multilevel analysis. *Journal of Drug Issues*, 31: 867-888.
- Davis, R. E., M. P. Couper, N. K. Janz, C. H. Caldwell, and K. Resnicow. 2010. Interviewer effects in public health surveys. *Health Education Research*, 25: 14-26.
- Dykema, J., D. Basson, and N. C. Schaeffer. 2007. Face-to-face surveys. In *The SAGE Handbook of Public Opinion Research*, W. Donsbach and M. W. Traugott (Eds). London: Sage Publications Ltd. 240-248.
- Fahrney, K., J. Uhrig, and T. Kuo. 2010. Gender-of-voice effects in an ACASI study of same-sex behavior. RTI Press publication No. MR-0017-1004. Research Triangle Park, NC: RTI International.
- Fendrich, M., T. Johnson, C. Shaligram, and J. S. Wislar. 1999. The impact of interviewer characteristics on drug use reporting by male juvenile arrestees. *Journal of Drug Issues*, 29: 37-58.
- Hunt, J. (1984). The development of rapport through the negotiation of gender on field work among police. *Human Organization*, 43(4), 283-296.
- Johnson, T. P., and J. A. Parsons. 1994. Interviewer effects on self-reported substance use among homeless persons. *Addictive Behaviors*, 19: 83-93.
- Jourard, S. 1971. *Self-Disclosure: An Experimental Analysis of the Transparent Self*. New York: Wiley.
- Kane, E. W., and L. J. Macaulay. 1993. Interviewer gender and gender attitudes. *Public Opinion Quarterly*, 57: 1-28.
- Kreuter, F., S. Presser, and R. Tourangeau. 2008. Social desirability bias in CATI, IVR, and Web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, 72: 847-865.
- Laumann, E., J. Gagnon, R. Michael, and S. Michaels. 1994. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.
- Mensch, B. S., P. C. Hewett, A. S. Erulkar. 2003. The reporting of sensitive behavior by adolescents: a methodological experiment in Kenya. *Demography*, 40: 247-268.
- Nass, C., E. Robles, C. Heenan, H. Bienstock, and M. Treinen. 2003. Speech-based disclosure systems: effects of modality, gender of prompt, and gender of user. *International Journal of Speech Technology*, 6: 113-121.

- Pollner, M. 1998. The effects of interviewer gender in mental health interviews. *The Journal of Nervous & Mental Disease*, 186: 369-373.
- Rogers, S. M., H. G. Miller, B. H. Forsyth, T. K. Smith, and C. F. Turner. 1996. Audio-CASI: the impact of operational characteristics on data quality. In *JSM Proceedings, Survey Research Methods Section*. Alexandria, VA: American Statistical Association. 1042-1047.
- Schaeffer, N. C., J. Dykema, and D. W. Maynard. 2010. Interviewers and interviewing. In *Handbook of Survey Research, Second Edition* J. D. Wright and P. V. Marsden (Eds). United Kingdom, Emerald Group Publishing Limited. 437-470.
- Tannen, D. 1996. *Gender and Discourse*. New York, Oxford: Oxford University Press.
- Tourangeau, R., M. Couper, and D. Steiger. 2003. Humanizing self-administered surveys: experiments on social presence in Web and IVR surveys. *Computers in Human Behavior*, 19: 1-24.
- Tourangeau, R., and T. W. Smith. 1996. Asking sensitive questions: the impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60: 275-304.
- Turner, C. F., L. Ku, S. M. Rogers, L. D. Lindberg, J. H. Pleck, and F. L. Sonenstein. 1998a. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science*, 280: 867-873.
- Turner, C. F., B. H. Forsyth, et al. 1998b. Automated self-interviewing and the survey measurement of sensitive behaviors. In *Computer Assisted Survey Information Collection* M. P. Couper, R. P. Baker, J. Bethlehem et al (Eds). New York, John Wiley and Sons. 455-474.
- Weisel, D. L. (2002). *Contemporary gangs: An organizational analysis*. New York: LFB Scholarly Publishing.
- Wilson, S., N. Brown, C. Mejia, and P. Lavori. 2002. Effects of interviewer characteristics on reported sexual behaviour of California Latino couples. *Hispanic Journal of Behavioral Sciences*, 24: 38-62.
- Wolitski, R., J. Parsons, C. Gomez, D. Purcell, C. Hoff, and P. Halkitis. 2005. Prevention with gay and bisexual men living with HIV: rationale and methods of the Seropositive Urban Men's Intervention (SUMIT). *AIDS*, 19(Supplement 1): S1-S11.