

---

# A randomized controlled trial of a multi-dose bystander intervention program using peer education theater

Sarah McMahon\*, Samantha C. Winter, Jane E. Palmer, Judy L. Postmus, N. Andrew Peterson, Sharon Zucker and RuthAnne Koenick

\*Correspondence to: S. McMahon. E-mail: smcmahon@ssw.rutgers.edu

Received on May 27, 2014; accepted on May 26, 2015

---

## Abstract

This article reports findings from a longitudinal, experimental evaluation of a peer education theater program, Students Challenging Realities and Educating Against Myths (SCREAM) Theater. This study examines the impact of SCREAM Theater on a range of bystander-related outcomes (i.e. bystander intentions, bystander efficacy, perception of friend norms and bystander behaviors) in situations involving sexual violence and whether there was a differential impact of the program by participant sex. First-year college students completed three waves of surveys (pretest, first post-test and second post-test). All participants received one dose of the intervention during summer orientation after the pretest. After the first post-test, participants were randomly assigned to receive two additional doses, or to a control condition, in which they received no additional doses. Students in both one- and three-dose groups reported a number of positive increases. Overall, an intent-to-treat analysis ( $n = 1390$ ) indicated three doses of the intervention during the first semester of college resulted in better outcomes than the one-time intervention during summer orientation alone. Although both male and female students' scores increased during the study period, female students consistently scored higher than male students on each outcome. The findings suggest that peer education theater holds promise for bystander intervention education on college campuses.

## Introduction

Campus sexual assault is a major public health problem, with an estimated 19–25% of women and 6% of men reporting attempted or completed sexual assaults since entering college [1, 2]. The recent report from the White House Task Force to Protect Students from Sexual Assault explicitly encourages college campuses to employ bystander intervention methods [3]. Such an approach challenges community members to change attitudes about sexual violence and teaches them to intervene when faced with situations involving sexual violence [4, 5].

Students Challenging Realities and Educating Against Myths (SCREAM) Theater is a peer education, interactive theater program that seeks to engage students in sexual assault prevention and bystander intervention. It is recognized as an evidence-informed strategy within the state of New Jersey, has been identified as a sample program by the National Sexual Violence Resource Center and was cited by the White House Task Force as an example of interactive theater used to promote bystander intervention [6, 7]. SCREAM Theater combines a number of evidence-based areas including sexual violence prevention, bystander intervention, peer education and theater. The program is based on principles firmly established in the literature on sexual violence prevention and bystander intervention [8, 9]. As outlined in Fig. 1, these include key concepts such as bystander decision-making process [10], barriers specific to sexual assault bystander situations [9] and addressing social norms [11]. Peer education

*SCREAM Theater Sessions*

Session	Purpose	Format	Participants
1 (June – August, 2010)	To engage audience members through theater while demonstrating positive peer norms about intervening and sexual violence.	A three part program: a skit acted out by peer educators that depicts a sexual assault; an in-character question and answer session; and an out-of-character information session.	All students
2 (Oct., 2010)	To provide an introduction to the bystander intervention process (Latane & Darley, 1970); address common barriers to intervention (Burn, 2009); show audience members different options for intervening (McMahon et al, 2013), and address the notion of “pluralistic ignorance” by reinforcing that the students’ peers are often willing to act as bystanders (Fabiano, Perkins, Berkowitz, Linkenbach & Stark, 2003).	Facilitator presents introductory information about bystander intervention. Scenes from Session 1 skit are re-created for small groups of students, but interrupted to explore different bystander intervention possibilities. Each time a scene is interrupted, students are given options of how the bystander could have acted differently, and, are able to anonymously vote on what action, the bystander should take.	Students in the three dose group only
3 (Nov., 2010)	To address the barrier that students are willing to intervene but do not know what to do (Burn, 2009; McMahon et al, 2013); provide concrete skill building for bystander situations and increase self-efficacy (Coker et al, 2011), and positively shift perceptions of peer norms about intervening (Brown & Messman-Moore, 2010).	The six scenes outlined in Sessions 1 & 2 are assigned to small groups of student participants; specifically outlining potential points of bystander intervention. <i>SCREAM Theater</i> peer educators work with the small groups to brainstorm concrete, helpful ways to intervene. Student participant groups select one of their suggested interventions, and present it to the rest of the participants by acting it out in their own way.	Students in the three dose group only

Note. Pretest (T1) occurred prior to Session 1. The first posttest (T2) occurred in early September, 2010. The second posttest (T3) occurred in December, 2010.

**Fig. 1.** SCREAM Theater sessions.

is used in many health education settings to change knowledge, attitudes and behaviors [12], and the use of peers has been utilized in evidence-based bystander programs such as Bringing in the Bystander [8, 13]. Additionally, SCREAM draws

on the literature that demonstrates that theater as a form of peer education is an effective method for conveying health-related messages, including information about sexual violence [14–18]. SCREAM Theater is unique in its use of peer education theater

to convey messages about sexual violence as well as bystander intervention.

SCREAM Theater can be a one- or three-session program (as described in Fig. 1). The one-session program is a 75-min skit that is followed by a question-and-answer session and is traditionally used as a stand-alone program at orientation. SCREAM Theater has performed for all incoming students at this university for over 20 years. Both men and women attend the program together, which is a deliberate part of the program design so that they can hear one another's perspectives during discussion segments [17]. The three-session program includes the skit and two follow-up small group skill-building sessions lasting 60 min each. Rather than a facilitator prescribing how students could intervene, these sessions allow students to generate their own range of possible bystander intervention actions, tailored to specific situations and settings on their campus [19, 20]. An initial evaluation of the one-session program indicated that participants positively increased their bystander intentions (willingness to intervene) and decreased levels of rape myths [18].

Because bystander intervention programs are relatively new, there remains a need for rigorous studies to evaluate the impact of these programs on college students. In particular, questions remain about the best ways to deliver bystander education and what dosage is sufficient to produce both attitudinal and behavioral change. Banyard *et al.*'s [8] and Coker *et al.*'s [21] studies suggested that longer exposure to programs may result in better outcomes. Given the resources needed to implement multisession programs, the question of dosage is especially important for colleges and universities [22]. Additionally, it is important to determine whether prevention programs provide the same impact for both men and women [5, 23, 24]. Hence, this study extends previous research on SCREAM Theater by first replicating the pre-post design to determine if one dose of the program has a positive impact, and second, by employing a randomized control design that included measures of both attitudes and behaviors to assess the short-term impact of the program. Given the importance of the public health problem of sexual violence, the potential of

the bystander approach and the limited research to date, this study is needed to address the following questions: (i) does exposure to one session of SCREAM Theater positively impact students' bystander intentions, efficacy, friend norms and behaviors? (ii) does exposure to three sessions of SCREAM Theater result in significantly better outcomes than exposure to one session? Additionally, both research questions assessed the impact of SCREAM Theater on bystander outcomes for male versus female participants.

---

## Methods

---

### Participants

The population under study was all incoming, first-year students who attended summer orientation at a large public university in the northeast in 2010 (June–August). University records indicated that 4362 of 6033 entering first-year students attended orientation (72.3% participation). Because SCREAM Theater is an institutionalized requirement for all incoming students who attended orientation, it was not feasible to exclude students from seeing the performance (to have a control group that received no doses). Instead, randomization occurred after all students participated in this first dose of intervention.

The research team attended each orientation session to provide informed consent and invited the 4362 students to complete an anonymous survey by pencil and paper, prior to viewing a performance by SCREAM Theater. To make participation anonymous, participants created a self-generated identification code that included the month and date of their birth and the first three letters of their mother's first name. Participant contact information was collected from individuals who opted to participate in a raffle to win a television or iPad on a form separate from the survey. In total, the research team collected contact information from 4311 individuals who completed the pretest surveys, for a response rate of 99% for the pretest. All methods were approved by the Institutional Review Board and the study was registered with ClinicalTrials.Gov.

## Randomization procedures

Students who provided their contact information for the raffle were contacted via email in early September 2010 and invited to complete the first post-test. A total of 2021 students completed the first post-test and agreed to participate in the longitudinal study. Stratified random assignment by sex and ethnicity was conducted based on the random sampling algorithm of SPSS to generate a random allocation sequence. This was used to randomly assign students to the experimental group ( $n = 1224$ ) and the one-dose group ( $n = 797$ ). A 60/40 split was utilized because the experimental group was required to physically go to two additional sessions, which would likely yield higher attrition than the control group. Students assigned to the experimental group were invited to attend two additional sessions of SCREAM Theater (henceforth, the three-dose group); the control group received no further doses (henceforth, the one-dose group). In December 2010, all participants ( $n = 2021$ ) regardless of group assignment were invited to take the second post-test (see Fig. 2 for consolidated standards of reporting trials (CONSORT) participant flow chart).

## Retention strategies

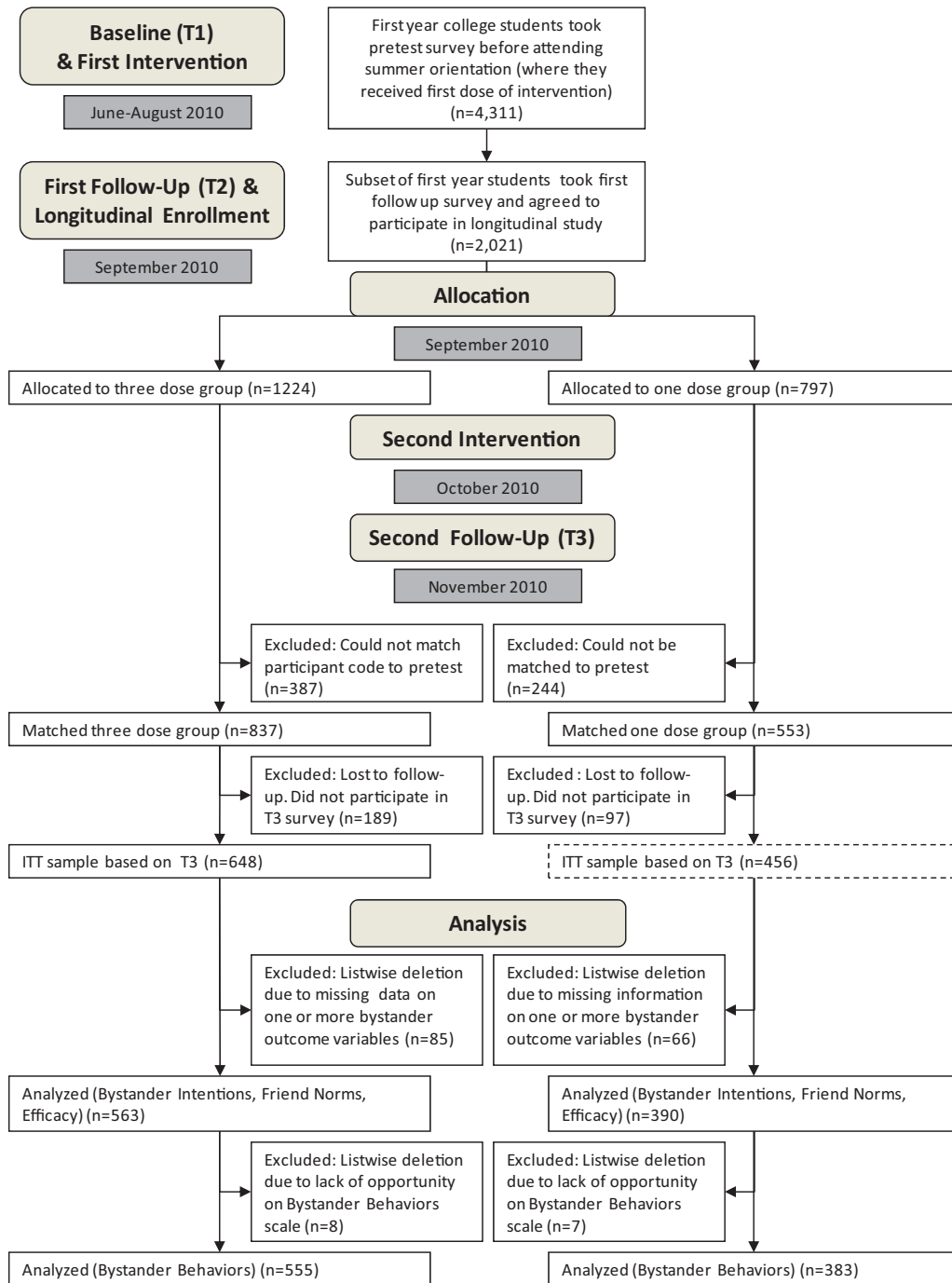
A variety of evidence-informed retention strategies were utilized in the study. First, the incentives for participation were carefully considered. At pretest, a raffle for iPads and televisions, was used based on findings from past research highlighting the positive impact of raffle-based incentives for college students [25]. For subsequent participation in surveys and programs, participants received progressive cash incentives—a strategy also supported by previous research [26, 27]. Each time students participated they received a larger incentive, starting with \$10 and increasing to \$30.

Second, a rigorous tracking and communication system was implemented to keep in monthly contact with study participants [27, 28]. Students received regular reminders that they were enrolled in a study and information about what next to expect through Twitter, Facebook, on-campus posters, emails and text messages. Periodically, participants were asked

to confirm or update their contact information with the research team. A brand identity, called ‘iSCREAM’, was created for the project so that students would recognize when correspondence originated from the research team [29]. Finally, although the pretest was conducted during new student orientation using paper and pencil, the mode of data collection was changed to internet-based for the follow-up surveys because electronic surveys are more convenient for study participants (and, admittedly, the research team) [26, 29, 30].

## Intention-to-treat sample

Intention-to-treat (ITT) analyses, also referred to as ‘analyse as randomized’ [31], require inclusion of all individuals that were randomized in an experimental study design, including those that do not attend follow-up [32]. Therefore, the ITT analytic sample includes all participants who (i) were randomly assigned to either the one-dose group or the three-dose group after the first post-test and (ii) had at least one follow-up survey with an identification code that could be matched to the pretest. The latter criterion was critical because only the follow-up surveys contained information regarding each participant’s intervention assignment. As Figure 2 displays, 2021 students (47% of the pretest sample) took the first post-test and agreed to participate in the longitudinal study. Of these 2021 participants, only 1390 (68.8%: 837 in the two-dose group, 553 in the one-dose group) could be matched to a baseline survey due to inconsistent self-generated identification codes. Of these matched students in the ITT sample, 1104 (79.4%) completed the second post-test (648 in the three-dose group and 456 in the one-dose group). An additional 151 cases were removed because they had incomplete data on one or more of the bystander intentions, friend norms or efficacy variables. Therefore, the analytic sample for analyses involving the bystander intentions, friend norms and efficacy variables included 953 participants (563 in the three-dose group and 390 in the one-dose group). The analytic sample for analyses involving the bystander behavior variable included 938 participants (555 in the



**Fig. 2.** Flow diagram of participation and analytic sample.

three-dose group and 383 in the one-dose group). The lower sample size for the bystander behavior variable is because participants were excluded if they indicated they did not have the opportunity to intervene as a bystander during the study period.

## Measures

### *Dependent variables*

*Bystander intentions.* Items from the Bystander Attitude Scale, Revised (BAS-R) [33], based on Banyard *et al.*'s work [34], were used in the study. A total of 16 items were listed, each stating a different bystander behavior in a potential sexual violence situation. Participants indicated how likely they were to engage in the behavior in the future on a Likert scale from 1 to 5, 'Unlikely' to 'Very likely'. An example item is 'In the future, how likely are you to check in with a friend who looks drunk when she goes to a room with someone else at a party?' A composite score was created by averaging across the 16 items ( $\alpha = 0.88$  at pretest).

*Bystander friend norms.* Participants were asked to indicate how likely they think their friends would be to engage in the BAS-R behaviors in the future on a Likert scale from 1 to 5, 'Unlikely' to 'Very likely'. The items on the scale were the same as for bystander intentions, except 'how likely are YOU to...' was replaced with 'how likely do you think YOUR FRIENDS are to...' A composite score included the average score for the 16 items ( $\alpha = 0.87$  at pretest).

*Bystander efficacy.* To assess confidence in one's ability to intervene, the Bystander Efficacy Scale was used [8, 34]. Respondents were asked to rate their confidence in performing certain bystander behaviors on a scale of 0 ('can't do') to 100 ('very certain can do'). For example, 'express my discomfort if someone says that rape victims are to blame for being raped' or 'ask a friend if they need to be walked home from a party'. Each individual received a score based on the average score across all items [24]. At pretest, the 18-item scale had good reliability ( $\alpha = 0.91$  at pretest).

*Bystander behaviors.* To measure bystander behaviors, a modified version of the Bystander

Behavior Scale—Revised, was used [33], which is based on the work by Banyard *et al.* [34]. Participants were asked to indicate whether they had intervened in any of the 16 items that were identical to the items in the bystander intentions and bystander friend norm scales. Participants could indicate 'yes,' 'no' or 'wasn't in situation' for each item.

A bystander behavior variable was created to measure the extent of bystander behaviors participants engaged in since the last survey. The procedures developed by Murphy [35] were followed to create a measure of bystander behaviors that accounts for whether or not the participant had an opportunity to intervene. Therefore, the ratio of the number of times a person was in the situation to intervene was divided by the number of times the person intervened. As mentioned earlier, individuals who answered 'wasn't in situation' for all 16 items were not included in the analytic sample ( $n = 15$ ) for the analyses involving the bystander behavior variable. This study only includes the change in participants' bystander behaviors after randomization (i.e. from the first post-test to the second post-test) due to a change in reference period from pretest to the first post-test. In the pretest, participants were asked whether they intervened as a bystander in the past 12 months, but in the first post-test, the reference period for past bystander behavior was 'since attending new student orientation'.

### *Independent variables*

Independent variables included treatment (whether the participant was assigned to the one-dose or three-dose group) as well as sex (Male or Female). (Social desirability was also assessed as a part of the survey instrument, but analyses demonstrated low reliability and in follow-up focus groups with students for a subsequent study, they indicated that the measure was too outdated to provide any meaningful information. Therefore, it was not included in these analyses.)

### **Analytic strategy**

All data analyses were conducted using IBM SPSS Version 21. An item non-response analysis per wave

indicated minimal missing data (<5%), therefore listwise deletion was used. A number of preliminary screening *t*-tests were then conducted to compare baseline equivalence for demographics and outcome variables. A series of chi-square tests were performed to compare the full ITT sample with the analytic sample. Lastly, binary logistic regression was used to test whether certain groups were more likely to have matched participant codes across the three-survey periods, and Poisson regression was conducted to examine whether certain groups of students were more likely to stay in the study from the pretest until the second post-test. In addition, correlations between bystander outcome variables were examined for evidence of multicollinearity, and all variables were within the acceptable range of moderate correlation [36]. Kurtosis and skewness values were also calculated and all items fell within the acceptable range [37].

#### *Research Question 1 analysis*

To examine the first research question (the impact of one dose of SCREAM Theater between the pretest and first post-test), paired-samples *t*-tests were performed for bystander intentions, bystander friend norms and bystander efficacy (bystander behavior was not tested due to the reference period). A multivariate analysis of covariance (MANCOVA) was performed to test whether the impact of the first dose of the intervention between the pretest and first post-test was consistent for male and female students. To isolate the effects of sex between the pretest and first post-test, the MANCOVA controlled for bystander intentions, bystander friend norms and bystander efficacy pretest scores.

#### *Research Question 2 analysis*

To address the second research question (the impact of one versus three doses of SCREAM Theater between the first and second post-tests), two analyses were conducted because bystander behaviors needed to be examined separate from the other outcomes.

First, to look at the impact on bystander intentions, bystander friend norms and bystander

efficacy, MANCOVA was used. This tested the impact of sex and the interaction between sex and one versus three doses of SCREAM Theater. To isolate the effects of the additional doses of SCREAM Theater between the first and second post-tests, the MANCOVA controlled for bystander intentions, bystander friend norms and bystander efficacy scores at the first post-test. Both multivariate findings (measuring the effect of the one- versus three-dose program on all bystander outcomes, in general) and univariate findings (measuring the effect of the one- versus three-dose program on bystander intentions, bystander friend norms and bystander efficacy variables, separately) are presented in the results. Second, to address the impact of one versus three doses of the program on bystander behaviors, a simple analysis of covariance (ANCOVA) was conducted to examine the main effects of SCREAM Theater, the main effects of sex, and the interaction of sex and SCREAM Theater on bystander behaviors between the first and second post-tests. To isolate the effects of the intervention, the ANCOVA controls for bystander scores at the first post-test. The results of this ANCOVA are presented separately from the impact on intentions, friend norms and efficacy.

#### *Per-protocol analysis*

In addition to the ITT analyses, which are the focus of this article, identical analyses were also carried out on the per-protocol sample ( $n = 662$  for analyses related to bystander intentions, friend norms and efficacy;  $n = 739$  for bystander behavior analyses). The per-protocol sample included only students who (i) passed reliability checks built into the survey (i.e. they answered two questions correctly that were in the survey to gauge that the participant was still reading) and (ii) attended all doses of the intervention as assigned (i.e. they did not attend additional doses if they were assigned to the one-dose group and they attended both additional doses if they were assigned to the three-dose group).

## Results

### Descriptive statistics

Table I presents the descriptive statistics of the demographic characteristics of the samples and the university's incoming student population. The final ITT analytic sample testing the impact on bystander intentions, bystander friend norms and bystander efficacy included 953 students, with 563 individuals (59.1%) in the three-dose group and 390 individuals (40.9%) in the one-dose group. Almost two-thirds (63.5%) of the sample identified as female and 44.5% identified as Caucasian/white. In general, the descriptive statistics for the analytic sample closely mirrored the racial and ethnic distribution for the population (i.e. the entire incoming class), although there were significantly more female than male students in the analytic sample ( $P < 0.01$ ).

### Tests for baseline equivalence

#### Pretest versus ITT samples

The chi-square comparison of three samples (pretest, ITT full and ITT analytic) found that there were significantly more female students than

male students but no significant differences by ethnicity between the pretest sample and the ITT full sample. Between the ITT full sample and the ITT analytic sample, there were no significant differences by ethnicity or sex.

#### One-dose versus three-dose samples

The results of a chi-square test indicated that there were no statistically significant differences between the one- and three-dose groups by sex at the first post-test. Results from a separate chi-square test suggest that ethnicity also did not differ significantly between the one- and three-dose groups. *t*-Test results indicated that there were no significant differences between the one-dose and three-dose groups on all outcome variables at the first post-test (after randomization).

#### Male versus female participants

The means were significantly different between male and female participants on the bystander intentions, friend norms and efficacy outcome variables at the pretest and for the bystander behavior score at the first post-test, with female participants scoring

**Table I.** Demographic characteristics of the one-dose, three-dose, analytic sample and the population of incoming university students

	Received one-dose sample (%) ( <i>n</i> = 390)	Received three-dose sample (%) ( <i>n</i> = 563)	Analytic sample (%) ( <i>n</i> = 953)	All incoming students (Population) ( <i>n</i> = 6033)
Sex				
Female	63.1 ( <i>n</i> = 246)	63.8 ( <i>n</i> = 359)	63.5 ( <i>n</i> = 605)	49.6
Male	36.9 ( <i>n</i> = 144)	36.2 ( <i>n</i> = 204)	36.5 ( <i>n</i> = 348)	50.4
Race/Ethnicity				
Asian, South Asian or Pacific Islander	32.8 ( <i>n</i> = 128)	31.6 ( <i>n</i> = 178)	32.1 ( <i>n</i> = 306)	30.1
Caucasian/White	42.3 ( <i>n</i> = 165)	46.0 ( <i>n</i> = 259)	44.5 ( <i>n</i> = 424)	46.1
Other	24.9 ( <i>n</i> = 97)	22.4 ( <i>n</i> = 126)	23.4 ( <i>n</i> = 223)	24.6



higher than male participants on each outcome variable. This sex difference is consistent with previous research [8].

### Attrition and matching across time

An attrition analysis revealed that there were no significant differences by sex, but for ethnicity, Asian participants had a higher probability of staying in the study compared with White participants ( $P < 0.05$ ), but students in the ‘Other’ category did not have significantly different odds of staying in the study. In addition, an analysis of which participants’ self-generated identification codes were matched to a pretest survey indicated there were no significant differences by sex. However, results indicated that ethnicity was significant, in that Asian participants were more likely to be matched than White students ( $P < 0.01$ ), but participants in the ‘Other’ category did not have significantly different odds of being matched compared to White students.

### Research Question 1

#### *The impact of one dose*

To answer the first research question, the first step in this analysis was to examine the impact of one dose prior to randomization. Table II presents the results from paired-sample *t*-tests that were used to assess the impact of one dose of SCREAM Theater on bystander intentions, friend norms and efficacy. Results indicated that there was a significant difference in participants’ mean scores on bystander intentions [ $t(952) = -2.816, P < 0.01$ ] and bystander friend norms [ $t(952) = -2.922, P < 0.01$ ] between the pretest and the first post-test (i.e. participants’ bystander intentions and bystander friend norms increased significantly after the first dose of intervention). There was no significant difference in students’ bystander efficacy scores from pretest to the first post-test.

#### *The impact of one dose on male versus female participants*

Table II summarizes the univariate results from the MANCOVA that demonstrated the impact of one

dose of SCREAM Theater on each outcome separately (bystander intentions, friend norms and efficacy) by sex. From the pretest to the first post-test (i.e. after one dose of the intervention), there were significant univariate effects by sex for bystander intentions [ $F(1948) = 17.709, P < 0.001, \eta^2 = 0.018$ ], friend norms [ $F(1948) = 10.241, P < 0.01, \eta^2 = 0.011$ ] and efficacy [ $F(1948) = 6.029, P < 0.05, \eta^2 = 0.006$ ]. The positive change in female student’s scores on each of these three bystander outcomes was significantly larger than male student’s change in score. There was also a significant multivariate effect by sex between the pretest and the first post-test [Wilks’  $\lambda = 0.981, F(3, 946) = 6.019, P < 0.001, \eta^2 = 0.019$ ]—suggesting that the positive change in female students’ score on the group of bystander variables together was greater than male students’ change.

### Research Question 2

#### *The impact of one dose versus three doses of intervention*

To answer the second research question, the next step was to analyse whether, after randomization, participation in three doses of SCREAM Theater resulted in better outcomes than one dose. Table II presents the univariate results of the impact of one versus three doses of SCREAM Theater on each outcome (bystander intentions, friend norms and efficacy) by sex between the pretest and the second post-test. Results indicated that there was a significant multivariate effect for treatment [Wilks’  $\lambda = 0.991, F(3, 944) = 2.816, P < 0.05, \eta^2 = 0.009$ ] between the first and second post-tests (i.e. participants in the three-dose group scored significantly higher on the set of these three bystander outcomes between the first and second post-tests, as compared to the one-dose group).

A single ANCOVA was used to analyse the impact of the intervention on bystander behaviors between the first and second post-tests. We could not analyse the impact of one dose of the intervention, separately, due to a change in reference period for the variable. As Table III displays, there was no

**Table II.** The impact of one versus three doses of SCREAM Theater on bystander intentions, friend norms and efficacy by gender

	Male		Female		Total		Gender, treatment effects	
	Mean	SD	Mean	SD	Mean	SD	F or T	P value
<b>Bystander intentions</b>								
Pretest	3.39	0.725	3.72	0.666	3.60	0.705	-2.816 <sup>a</sup>	=0.005**
Post-test 1	3.46	0.608	3.76	0.587	3.65	0.611	17.709 <sup>b</sup>	<0.001***
Post-test 2	3.51	0.626	3.83	0.599	3.71	0.629	-3.833 <sup>c</sup>	<0.001***
One dose	3.46	0.657	3.77	0.637	3.66	0.661	15.343 <sup>d</sup>	<0.001***
Three dose	3.53	0.603	3.87	0.569	3.75	0.603	3.680 <sup>e</sup>	=0.055+
<b>Bystander friend norms</b>								
Pretest	3.10	0.702	3.41	0.662	3.30	0.693	-2.922	=0.004**
Post-test 1	3.20	0.621	3.45	0.569	3.36	0.600	10.241	=0.001**
Post-test 2	3.27	0.638	3.56	0.557	3.45	0.604	-6.043	<0.001***
One dose	3.25	0.659	3.54	0.600	3.43	0.637	13.355	<0.001***
Three dose	3.28	0.625	3.58	0.526	3.47	0.581	0.008	=0.927
<b>Bystander efficacy</b>								
Pretest	70.68	15.599	73.87	14.672	72.71	15.088	-0.008	=0.994
Post-test 1	69.61	16.521	74.49	15.469	72.71	16.026	6.029	=0.014*
Post-test 2	70.11	16.697	75.11	15.027	73.29	15.833	-1.348	=0.178
One dose	68.38	16.926	74.68	14.648	72.35	15.802	1.190	=0.276
Three dose	71.33	16.466	75.41	15.294	73.93	15.836	0.712	=0.399
<b>n</b>								
Pretest	348		605		953		6.019 <sup>f</sup>	<0.001***
One dose	144		246		390		5.855 <sup>g</sup>	=0.001**
Three doses	204		359		563		2.816 <sup>h</sup>	=0.038**

Note. Significant levels are as follows: + $P < 0.1$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Means and standard deviations at pretest, post-test 1 and post-test 2 are presented. <sup>a</sup>Paired-Samples  $t$ -tests from (before randomization) pretest to post-test 1. <sup>b</sup>MANCOVA (univariate) main effect for gender (before randomization) from pretest to post-test 1. <sup>c</sup>Paired-Samples  $t$ -tests (after randomization) from post-test 1 to 2. <sup>d</sup>MANCOVA (univariate) main effect for gender (after randomization) from post-test 1 to 2. <sup>e</sup>MANCOVA (univariate) main effect for treatment (after randomization) from post-test 1 to 2. <sup>f</sup>MANCOVA (multivariate) main effect for gender (before randomization) from pretest to post-test 1. <sup>g</sup>MANCOVA (multivariate) main effect for gender (after randomization) from post-test 1 to 2. <sup>h</sup>MANCOVA (multivariate) main effect for treatment (after randomization) from post-test 1 to 2.

**Table III.** The impact of one versus three doses of SCREAM Theater on bystander behaviors by gender

	Male		Female		Total		Gender, Treatment Effects	
	Mean	SD	Mean	SD	Mean	SD	F	P value
<b>Bystander behaviors</b>								
Post-test 1	28.30	25.729	39.65	27.548	35.52	27.435	-5.636 <sup>a</sup>	<0.001***
Post-test 2	30.10	21.884	42.07	23.681	37.71	23.740	17.935 <sup>b</sup>	<0.001***
One dose	28.67	19.872	41.41	23.817	36.76	23.255	0.003 <sup>c</sup>	=0.959
Three dose	31.09	23.172	42.53	23.609	38.37	24.068		
<b>N</b>								
Pretest	342		596		938			
One dose	140		243		383			
Three doses	202		353		555			

Note. + $P < 0.1$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Means and standard deviations along with the univariate within-subject ANCOVA results are presented. <sup>a</sup>Paired-samples  $t$ -tests (after randomization) from post-test1 to 2. <sup>b</sup>ANCOVA main effect for gender (after randomization) from post-test1 to 2. <sup>c</sup>ANCOVA main effect for treatment (after randomization) from post-test1 to 2.

significant difference in bystander behaviors for the one-dose versus three-dose groups.

### *The impact of one versus three doses on male versus female participants*

Finally, we assessed whether there was a differential impact by sex. First, we looked at the bystander intentions, bystander efficacy and bystander friend norms outcomes. The MANCOVA revealed no significant interaction effects between sex and treatment on the bystander outcomes, as a whole, or on the individual bystander outcomes. Univariate results of MANCOVA (shown in Table II), however, indicated significant main effects for sex on bystander intentions [ $F(1946) = 15.343$ ,  $P < 0.001$ ,  $\eta P^2 = 0.016$ ] and bystander friend norms [ $F(1946) = 13.355$ ,  $P < 0.001$ ,  $\eta P^2 = 0.014$ ]. Results also indicated that there was a significant multivariate effect for sex [Wilks'  $\lambda = 0.982$ ,  $F(3, 944) = 5.855$ ,  $P < 0.01$ ,  $\eta P^2 = 0.018$ ] between the first and second post-tests, i.e. female students' scores showed consistently larger positive changes than the scores of male students between the first and second post-tests on overall bystander outcomes. Additionally, female students' scores were significantly higher than males' scores on bystander intentions and bystander friend norms individually between the first and second post-test.

To test the role of sex in the bystander behavior outcomes, the ANCOVA revealed a significant effect by sex [ $F(1, 933) = 17.935$ ,  $P < 0.001$ ,  $\eta P^2 = 0.019$ ], suggesting that although both female students and male students showed an increase in bystander behaviors between the first and second post-tests, female students showed a larger, positive change in bystander behaviors than male students between the first and second post-tests.

### **Per-protocol results**

A comparison of the results from analyses based on the ITT versus the per-protocol sample ( $n = 662$  for analyses related to bystander intentions, friend norms and efficacy;  $n = 739$  for bystander behavior analyses) revealed few overall differences in the

significant intervention and effects by sex on bystander intentions, friend norms, efficacy or behaviors. The notable differences were that the per-protocol sample showed significant univariate effects for the one- versus three-dose group comparison on bystander intentions [ $F(1655) = 16.724$ ,  $P < 0.001$ ,  $\eta P^2 = 0.025$ ] and bystander efficacy [ $F(1655) = 6.387$ ,  $P < 0.050$ ,  $\eta P^2 = 0.010$ ] between the first and second post-tests. Additionally, there was no longer a significant main effect on bystander efficacy by sex in the per-protocol analyses between the pretest and first post-test.

---

## **Discussion**

---

The results of this study indicated that participation in a peer education theater program resulted in a positive increase in bystander-related outcomes in the context of campus sexual assault prevention. This suggests that peer education theater may be an effective strategy for educating college students about their role in addressing sexual violence on campus. The results indicated that participation in one or three doses of the program yielded positive effects. This research expands on previous findings that one dose of the program is beneficial [5], resulting in significant increases in bystander intentions and perceptions of friend norms, although bystander efficacy did not improve significantly. Overall, the amount of bystander behaviors increased for all students during the study period; however, this increase was not associated with the number of doses they received.

When comparing participation in one versus three doses of SCREAM Theater, those receiving three doses demonstrated significantly better scores on the multivariate combination of outcomes (i.e. bystander intentions, friend norms and efficacy). These results support previous findings that exposure to multiple doses of bystander education result in stronger bystander outcomes [8, 21]. Therefore, these results suggest that it may be valuable to move away from the 'one-shot' types of bystander intervention education programs to implement ongoing opportunities for students to

develop the attitudes and skills needed for bystander intervention.

The findings regarding bystander behavior, however, do not paint a clear picture about involvement in additional doses. Although both groups reported an increase in the number of bystander behaviors, the three-dose group did not report engaging in significantly more bystander behaviors. Participants' intentions to intervene as a bystander significantly increased during the study period, which may be an important step prior to actually intervening as a bystander. Longer-term research is needed to better examine the impact of participation in the intervention to determine its impact on actual bystander behaviors.

The comparison of the ITT results from this study with those from an identical set of analyses run on the per-protocol sample is consistent with literature that suggests that results from ITT analyses are often more conservative. The per-protocol results suggest that additional doses of SCREAM Theater may have significant, positive effects for students who adhere to the assigned treatment protocol, i.e. by actually attending the intervention [38].

Across the board, sex played a significant role in the outcomes, which is consistent with a growing body of research [24, 39]. As the analysis revealed, sex differences began prior to entering college with female students reporting higher scores on all bystander outcomes. Because the issue of sex is consistently salient in this area of research, it may be important for programs in high schools and colleges to consider including information that specifically addresses gender norms around sexual violence and bystander intervention.

The lack of a significant impact of the intervention by sex between the first and second post-tests may suggest that, although there appear to be overall differences in students' bystander outcomes by sex, additional doses of SCREAM Theater did not seem to have differential impacts on male and female students. Further research should explore whether bystander intervention programs are more effective when presented to the entire community (including both men and women) or to single sex groups.

## Limitations and future research

The results of the study must be interpreted in the context of a number of limitations. First, the study was unable to utilize an untreated control group. The control group in this study was the existing or standard treatment [40]. This is due to the natural setting of the research, rather than conducting it in an experimental setting, which is the reality faced by many social science settings. Future research on sexual violence prevention that can obtain a control group without any exposure to intervention would be beneficial, perhaps by staggering when students receive programming. However, given past research that indicated the effectiveness of one dose of SCREAM Theater [18], there were also ethical considerations associated with withholding a treatment that is known to be effective. Second, the study was conducted on one public university campus and therefore findings may not generalize to students from other university settings. Discourse surrounding sexual violence is often influenced by community norms, and there may be important variations on other campuses. Additionally, events other than SCREAM Theater that occurred on campus may have impacted students' attitudes about bystander intervention. Third, the study utilized measures that required self-reporting, which may be subject to social desirability bias or inaccurate reporting. The social desirability measure we used demonstrated low reliability so it was not included, however, future research should consider the inclusion of a stronger social desirability instrument. Fourth, participants created a self-generated identification code as a means of protecting their identities and therefore reducing bias. However, the research team faced some barriers in matching these identification codes across time, which meant that a number of students had to be excluded from the analyses. There were also some incoming first-year students that did not attend the initial orientation and were therefore excluded from the study. Additionally, some of the measures utilized in the study are relatively new; further testing is underway to establish reliability and validity.

Attrition was also a challenge in this study. The initial pretest was given when students were all present in a room during orientation with pencil and paper; therefore, we had a 99% response rate. Moving forward, 47% of those students ( $n = 2021$ ) agreed to participate in the longitudinal study. This amount of dropout between the pretest and first post-test is not only consistent with other findings from the literature, [30] but also due to the design of this particular study. That is, in this study, students did not agree to participate in the longitudinal study until the first post-test. Therefore, only first post-test participants were randomly assigned to the one-dose or three-dose group. Also, the first post-test and second post-test involved online surveys, rather than an in-person survey with a captive audience. In addition, the failure to match students across time periods limited the size of our analytic sample [41]. Although 2021 students agreed to participate in the longitudinal study, only 1390 students (837 in the three-dose group and 553 in the one-dose group) had at least one follow-up survey that matched a pretest survey. Of these, 1104 completed the second post-test. While it is important for researchers to ensure that participants can remain anonymous in surveys about sensitive topics like sexual violence, there are a number of steps that researchers can implement to minimize inaccurate and inconsistent information. For example, participants should be asked to provide responses to a number of questions that will comprise the code [42]. In addition, these questions should be pilot-tested to assess their reliability.

The complications with matching and attrition may have introduced bias into the sample. Analyses indicated that there were significantly more females who participated in the study, which limits generalizability to the larger student body. Additionally, the results revealed that Asian students had better odds of staying in the study and having their surveys matched over time compared with white students, which suggests further research is needed to assess the role of ethnicity and race in the impact of bystander intervention programs as well as longitudinal studies more generally. While this research is underway as part of the SCREAM

Theater project [41], further research is needed to assess the role of ethnicity and race in the impact of bystander intervention programs, as well as participation in longitudinal studies. There may be other differences between the analytic sample and the original baseline group on key outcome variables or other descriptive variables. Further work is needed to determine how these variables may impact the final analytic sample in studies such as this, so that researchers can accurately report the representativeness of their findings [42].

Further work is also needed to explore the relationship among these different bystander outcomes. While the multivariate outcomes for bystander intentions, peer norms and efficacy were significant when comparing the three-dose and one-dose groups, not all outcomes variables were significant individually. These results suggest that there may be an important relationship among these various bystander outcomes, possibly representing a latent construct where these variables work together. The measurement of bystander behavior presents a complex challenge due to the need to incorporate both the individual's opportunities for intervention and their actual intervention [43]. Additionally, the outcomes measured for participants in the study were short-term (1 month after completion of the additional sessions). This study is part of a larger evaluation of SCREAM Theater that involved 18 months of data collection. Therefore, an important next step is to determine whether these changes were sustained over a greater length of time. Assessing the impact of the program on participants' perpetration is also an important area of future research, along with analysing the correlates of both perpetration and victimization. Lastly, because the effects were small, and more so for men, there is a need for the program to find ways to continuously increase its impact and effectiveness.

Regardless, the impact of the program by the end of the first semester for first-year students was significant, indicating that SCREAM Theater had a positive impact on participating students' attitudes and behaviors on engaging in sexual violence prevention in the short term. Given the widespread nature of the public health problem of sexual

violence, this is a critical step in determining how to prevent it in the context of college campuses.

---

### Funding

---

This work was supported by the Centers for Disease Control and Prevention. Preparation of this manuscript was supported by a grant from the Centers for Disease Control and Prevention (grant number: 1R01CE001855-01, PI: McMahon).

---

### Conflict of interest statement

---

None declared.

---

### References

---

- Krebs CP, Lindquist CH, Warner TD *et al.* *The Campus Sexual Assault (CSA) Study: Final Report*. Washington DC: US Department of Justice, National Institute of Justice, 2007. Available at: <https://www.ncjrs.gov/pdffiles1/nij/grants/221153.pdf>. Accessed: 26 May 2014.
- Kilpatrick D, Resnick H, Ruggiero K *et al.* *Drug Facilitated, Incapacitated and Forcible Rape: A National Study*. Washington DC: US Department of Justice, National Institute of Justice, 2007. Available at: <https://www.ncjrs.gov/pdffiles1/nij/grants/219181.pdf>. Accessed: 26 May 2014.
- White House. *Not Alone: The first Report of the White House Task Force to Protect Students from Sexual Assault*. The White House. Available at: [http://www.whitehouse.gov/sites/default/files/docs/report\\_0.pdf](http://www.whitehouse.gov/sites/default/files/docs/report_0.pdf). Accessed: 29 April 2014.
- Banyard VL, Plante EG, Moynihan MM. Bystander education: bringing a broader community perspective to sexual violence prevention. *J Community Psychol* 2004; **32**: 61–79.
- McMahon S, Postmus JL, Koenick RA. Conceptualizing the engaging bystander approach to sexual violence prevention on college campuses. *J Coll Stud Dev* 2011; **15**: 115–30.
- Prevention Strategies: SCREAM Theater. *New Jersey Coalition of Sexual Assault*. Available at: <http://njcasa.org/prevention/prevention-strategies>. Accessed: 6 February 2014.
- White House. *Bystander-Focused Prevention of Sexual Violence: Tip Sheet*. The White House. Available at: <https://www.notalone.gov/assets/bystander-summary.pdf>. Accessed: 29 April 2014.
- Banyard VL, Moynihan MM, Plante EG. Sexual violence prevention through bystander education: an experimental evaluation. *J Community Psychol* 2007; **35**: 463–81.
- Burn SM. A situational model of sexual assault prevention through bystander intervention. *Sex Roles* 2009; **60**: 779–92.
- Latane B, Darley JM. *The Unresponsive Bystander: Why Doesn't He Help?* Englewood Cliffs, NJ: Prentice-Hall, 1970.
- Fabiano P, Perkins H, Berkowitz A *et al.* Engaging men as social justice allies in ending violence against women: evidence for a social norms approach. *J Am Coll Health* 2003; **52**: 105–12.
- Jennings JM, Howard S, Perotte CL. Effects of a school-based sexuality education program on peer educators: the teen PEP model. *Health Educ Res*. 2014; **29**: 319–29.
- Moynihan MM, Banyard VL, Cares AC *et al.* Encouraging responses in sexual and relationship violence prevention: what program effects remain one year later? *J Interpers Violence* 2014; **30**: 110–32.
- Ahrens CE, Rich MD, Ullman JB. Rehearsing for real life: the impact of the InterACT Sexual Assault Prevention Program on self-reported likelihood of engaging in bystander interventions. *Violence Against Women* 2011; **17**: 760–76.
- Black B, Weisz A, Coats S *et al.* Evaluating a psychoeducational sexual assault prevention program incorporating theatrical presentation, peer education, and social work. *Res Soc Work Pract* 2000; **10**: 589–606.
- Christensen MC. Using theater of the oppressed to prevent sexual violence on college campuses. *Trauma Violence Abuse* 2013; **14**: 282–94.
- Kress VE, Shepherd JB, Anderson RI *et al.* Evaluation of the impact of a coeducational sexual assault prevention program on college students' rape myth attitudes. *J Coll Couns* 2006; **9**: 148–57.
- McMahon S, Postmus JL, Warrenner C *et al.* Utilizing peer education theater for the primary prevention of sexual violence on college campuses. *J Coll Stud Dev* 2014; **55**: 78–85.
- Casey E, Lindhorst T. Toward a multi-level, ecological approach to the primary prevention of sexual assault: prevention in peer and community contexts. *Trauma Violence Abuse* 2009; **10**: 91–114.
- McMahon S, Hoffman ML, McMahon SM *et al.* What would you do? Sexual violence bystander intervention strategies for college students. *J College and Character* 2013; **14**: 141–51.
- Coker AL, Cook-Craig PG, Williams CM *et al.* Evaluation of green dot: an active bystander intervention to reduce sexual violence on college campuses. *Violence Against Women* 2011; **17**: 777–96.
- Vladutiu CJ, Martin SL, Macy RJ. College- or university-based sexual assault prevention programs: a review of program outcomes, characteristics, and recommendations. *Trauma Violence Abuse* 2011; **1**: 67–86.
- Banyard VL. Measurement and correlates of pro-social bystander behavior: the case of interpersonal violence. *Violence Vict* 2008; **23**: 85–99.
- Brown AL, Messman-Moore TL. Personal and perceived peer attitudes supporting sexual aggression as predictors of male college students' willingness to intervene against sexual aggression. *J Interpers Violence* 2010; **25**: 503–17.
- Laguilles J, Williams E, Saunders D. Can lottery incentives boost web survey response rates? Findings from four experiments. *Res High Educ* 2011; **52**: 537–53.
- Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol* 2007; **17**: 643–53.

27. McFarlane J. Strategies for successful recruitment and retention of abused women for longitudinal studies. *Issues Ment Health Nurs* 2007; **28**: 883–97.
28. Robinson K, Dennison C, Wayman D *et al.* Systematic review identifies number of strategies important for retaining study participants. *J Clin Epidemiol* 2007; **60**: 757.
29. Ribisl, KM, Walton MA, Mowbray CT *et al.* Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: review and recommendations. *Eval Prog Plann* 1996; **19**: 1–25.
30. Dillman DA. *Mail and Internet Surveys: The Tailored Design Method—2007 Update With New Internet, Visual, and Mixed-Mode Guide*. Hoboken, NJ: Wiley, 2011.
31. Fergusson D, Aaron SD, Guyatt G *et al.* Post randomization exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 2002; **325**: 652–4.
32. Gupta S. Intention-to-treat concept: a review. *Perspect Clin Res* 2011; **2**: 109–12.
33. McMahon S, Allen CT, Postmus JL *et al.* Measuring bystander attitudes and behavior to prevent sexual violence. *J Am Coll Health* 2014; **62**: 58–66.
34. Banyard VL, Moynihan MM, Cares AC *et al.* How do we know if it works? Measuring outcomes in bystander-focused abuse prevention on campuses. *Psychol Violence* 2014; **4**: 101–15.
35. Murphy MJ. *Sexual Assault-Specific Bystander Behavior: Accounting for Opportunity in a Prospective Analysis of the Effects of Individual, Social Norms and Situational Variables [dissertation]*. Athens, OH: Ohio University, 2013.
36. Mayers A. *Introduction to Statistics and SPSS in Psychology*. Edinburgh, UK: Pearson Education, 2013.
37. Lewis-Beck M, Bryman AE, Liao TF (eds). *The SAGE Encyclopedia of Social Science Research Methods*, vol. 1. Thousand Oaks, CA: SAGE, 2004.
38. Armijo-Olivo S, Warren S, Magee D. Intention to treat analysis, compliance, drop-outs and how to deal with missing data in clinical research: a review. *Phys Ther Rev* 2009; **14**: 36–49.
39. Banyard VL. Who will help prevent sexual violence: creating an ecological model of bystander intervention. *Psychol Violence* 2011; **1**: 216–29.
40. Braver SL, Smith MC. Maximizing both external and internal validity in longitudinal true experiments with voluntary treatments: the “combined modified” design. *Eval Program Plann* 1996; **19**: 287–300.
41. Winter SC, Palmer JE. The differential impacts of a bystander intervention program for students of color: Results from a randomized controlled trial. 28th Annual Conference of the American Evaluation Association; Denver, CO, October 2014.
42. Schnell R, Bachteler T, Reiher J. Improving the use of self-generated identification codes. *Eval Rev* 2010; **34**: 391–418.
43. McMahon S, Palmer J, Banyard V *et al.* Measuring bystander behavior in the context of sexual assault prevention: Lessons learned and new directions. *J Interpers Violence* forthcoming.