Am J Community Psychol (2015) 56:46–56
DOI 10.1007/s10464-015-9740-0

ORIGINAL ARTICLE

Predicting Bystander Behavior to Prevent Sexual Assault on College Campuses: The Role of Self-Efficacy and Intent

Sarah McMahon[1] · N. Andrew Peterson[1] · Samantha C. Winter[1] · Jane E. Palmer[2] · Judy L. Postmus[1] · Ruth Anne Koenick[3]

Published online: 21 July 2015

Abstract Bystander intervention has been increasingly applied to prevent sexual violence on college campuses. Its underlying theory assumes unidirectional relationships between variables, predicting that bystander behaviors (i.e., actions taken to intervene in sexual violence situations) will be influenced by bystander intentions (BI; i.e., likelihood to intervene in the future), which in turn will be affected by bystander efficacy (BE; i.e., confidence to intervene). One question for theory is whether a reciprocal relationship exists between BI and BE. We used structural equation modeling (SEM) with longitudinal data to test unidirectional and reciprocal causal relations between BI and BE. Participants (n = 1390) were students at a northeastern US university. Four models were examined using SEM: (1) a baseline model with autoregressive paths; (2) a model with autoregressive effects and BI predicting future BE; (3) a model with autoregressive effects and BE predicting future BI; and, (4) a fully cross-lagged model. Results indicated that reciprocal causality was found to occur between BI and BE. In addition, a final model demonstrated indirect effects of a bystander intervention program on bystander behaviors through both BI and BE at different time points. Implications for theory and practice are described, and directions for future research discussed.

Keywords Campus sexual assault · Bystander intervention · Cross-lagged model

Introduction

Bystander intervention is an increasingly popular approach to sexual violence prevention, particularly on college campuses. The recent Campus SaVE Act and the White House Task Force Report to Protect Students from Campus Assault call for colleges and universities to augment efforts to address campus sexual assault and explicitly promote a bystander education approach (Department of Education 2014; White House 2014). Bystander intervention education frames sexual violence as a community problem, and one that all individuals have a responsibility to address (Banyard et al. 2004). This approach to sexual violence prevention offers an avenue to involve all members of the community and focuses on potential solutions (Casey and Lindhorst 2009), rather than more traditional means of addressing all women as potential victims and all men as potential perpetrators (Lonsway 1996). The appeal of a community-level approach, coupled with recent attention to the issue of campus sexual assault, has prompted an increased interest in bystander intervention education programs across the country.

A number of evaluations of bystander intervention programs have demonstrated positive impacts on a range of outcomes. Katz and Moore's (2013) recent meta-analysis found that bystander intervention education programs had moderate effects on both bystander self-efficacy and intentions to intervene, and smaller effects on actual bystander behavior. However, they found that most program evaluations do not include bystander behavior measures and there are very few longitudinal studies evaluating

✉ Sarah McMahon
smcmahon@ssw.rutgers.edu

1 Center on Violence Against Women and Children, School of Social Work, Rutgers University, New Brunswick, NJ, USA

2 Department of Public Administration and Policy, American University, Washington, DC, USA

3 Office for Violence Prevention and Victim Assistance, Rutgers University, New Brunswick, NJ, USA

bystander intervention programs. In fact, only 25 % of the studies included in their analysis included actual behavior as an outcome, and of those, only two studies examined behaviors over time (2 and 4 months post-training). Most existing studies measure behavioral intent as an outcome; however, only measuring one's intent to intervene as a bystander in sexual violence situations is not necessarily a good proxy for one's actual bystander behavior (Banyard 2014; Murphy et al. in press).

Additionally, it is unclear how these key outcome variables (self-efficacy, intentions, and behaviors) relate to one another. Past research has found that one's self-efficacy and intentions to be a bystander are positively associated with bystander behaviors (e.g., Banyard 2008; Banyard and Moynihan 2011). However, the assumption of unidirectional causal relations between these variables has not yet been tested in the context of sexual violence prevention. Lastly, since most of these evaluation designs have relied on convenience or quasi-experimental designs, more rigorous evaluations are needed. The current paper addresses these gaps by using longitudinal data from a randomized control trial of a bystander intervention education program to study the relationship between bystander efficacy, intentions, and behavior.

The program under study is *SCREAM* (Students Challenging Realities and Educating Against Myths) *Theater*, a peer education, interactive theater program that seeks to engage students in sexual assault prevention and bystander intervention. The literature demonstrates theater as a form of peer education is an effective method for delivering health promotion messages, including sexual violence (Ahrens et al. 2011; Black et al. 2000; Christensen 2013). The program is based on principles well established in the literature on sexual violence prevention and bystander intervention, including key concepts such as the bystander decision-making process (Banyard et al. 2007; Latane and Darley 1970), barriers to engaging in bystander behavior in situations involving sexual assault (Burn 2009), and addressing peer norms (Fabiano et al. 2003).

SCREAM Theater can be a one or three session program. In the one dose version, students view an interactive skit on sexual assault and bystander intervention and participate in a question and answer session with the characters. In the three dose version, two more doses are added which are comprised of small, skill-building workshops (for a detailed description of the one and three dose programs, see McMahon et al. in press). The program was cited by the White House Task Force to Protect Students from Sexual Assault as an example of interactive theater used to promote bystander intervention. Previous evaluations demonstrated that one dose of the program resulted in decreased rape myths and increased bystander intentions (McMahon et al. 2014) and behaviors (McMahon et al. in press), and that short term results indicated that the three dose program generally reported better outcomes (McMahon et al. in press), and the study was registered with ClinicalTrials.Gov.

Bystander Efficacy, Intentions, and Behavior

Increasing bystander efficacy, intentions (i.e., willingness to intervene), and behaviors are common goals in bystander education programs (Katz and Moore 2013). All of these constructs are supported by the Theory of Planned Behavior (TPB), which purports that an individual's behavior is predicted by his or her *intention* to perform the behavior (Ajzen 1988, 1991). Intentions reflect a person's readiness to perform a behavior and are shaped by three determinants: (1) the person's attitude towards the behavior; (2) the subjective norms around performing the behavior, and (3) the person's perception of control over performing the behavior (i.e., perceived behavioral control, or self-efficacy). The TPB indicates that behavioral intent and self-efficacy directly predict behavior.

Perceived behavioral control, or self-efficacy, is an especially important construct in the TPB model because it theoretically predicts both intentions and actual behavior. The larger literature on helping supports the salience of self-efficacy, ranging from helping in emergency situations (Latane and Nida 1981), to robberies (Huston et al. 1981), to bullying (Cappadocia 2012; Gini et al. 2008; Thornberg et al. 2012), to sexual violence (Burn 2009; Exner and Cummings 2011; Moynihan et al. 2011). Fischer et al.'s (2011) meta-analysis of bystander intervention in dangerous and non-dangerous emergencies found that the "bystander effect" (i.e., lack of intervention by bystanders when other people are around; Darley and Latane 1968), is less likely to occur if the bystander feels able, competent, and confident to intervene

Banyard's work on bystander intervention with college students further establishes the link between self-efficacy and bystander behavioral intent, self-efficacy and behaviors, and behavioral intent and behaviors in situations involving sexual and dating violence. Consistently, she and her colleagues found that greater perceptions of self-efficacy were significant predictors of positive behavioral intent or actual behaviors (Banyard 2008; Banyard and Moynihan 2011; Banyard et al. 2004). This is supported by other work in the field of bystander intervention as well (e.g., Coker et al. 2011).

Although the significance of bystander efficacy and its relationship to intentions to intervene is supported theoretically and empirically, the direction of causality between these variables has not yet been empirically tested. It may be that the association is not as straightforward as earlier research supposed. More specifically, it may be that reciprocal causality exists between efficacy and intentions or perhaps increases in intentions precede increases in individuals' self-efficacy. This is an important question for the field, as it can help influence the design of bystander intervention programs. For example, for those individuals designing and implementing bystander intervention programs, it is useful to know more about the relationship between these key constructs. Should prevention programs be addressing efficacy and intentions sequentially, if one affects the other, or should they be addressed simultaneously? Additionally, with the ultimate goal of impacting behavior, how do these constructs relate to actual behavior? Hence, the hypotheses proposed for the study include:

H1 Both bystander intentions and efficacy will demonstrate strong autoregressive effects over time, indicating stability across time periods

H2 A fully cross lagged model in which there is reciprocal causality between bystander intentions and efficacy will provide a significantly better fit than other models which hypothesize unidirectionality

H3 The intervention will affect bystander behavior at Time 5 indirectly through its effect on bystander intentions and efficacy at Time 2

Discerning the direction of these relationships and how they ultimately relate to bystander behavior will help researchers and prevention program facilitators better understand the nuances of models of change over time.

Methods

The data from this study are from a larger, randomized control trial of a sexual violence prevention program on a large, Northeastern university campus. All students viewed the one-dose version of SCREAM Theater, and then were randomized, with the experimental group receiving two additional doses. All methods were approved by the Institutional Review Board.

Participants

Incoming first-year students were recruited at new student orientation sessions between June 2010 and August 2010. The research team attended each orientation session to describe the study, provided informed consent, and invited students to complete an anonymous paper and pencil survey prior to viewing a performance by *SCREAM Theater* (the first dose of the intervention). University records indicated that 4362 of 6033 (72 %) entering first year students attended orientation.

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.

Design

In total, the research team collected pretest surveys (T1) and contact information from 4311 individuals. These students were then contacted in early September 2010 to complete the first follow-up survey online (T2). A total of 2021 students completed T2 and, at that time, agreed to participate in the longitudinal study. Stratified random assignment procedures (based on gender and ethnicity) were then used to randomly assign students to the experimental or comparison group. Students assigned to the experimental group were invited to attend two additional sessions of *SCREAM Theater* (three dose group); the comparison group received no further doses (one dose group). Because the three dose students were asked to attend additional sessions, the level of attrition was expected to be greater for this group. Hence, the groups were weighted, with the three dose (experimental group) ($n = 1224$) representing 60 % of randomized sample and the one dose or comparison group ($n = 797$) representing 40 % of the randomized sample.

All participants received monetary compensation for each completed survey and, if assigned to the three dose group, for each additional dose they attended. In December 2010, February/March 2011, and September 2011, all participants regardless of group assignment were invited to take a third (T3), fourth (T4), and fifth (T5) online survey, respectively. The current study focuses on this T2–T5 panel data to examine the long-term impact of the program on bystander efficacy, intentions, and behavior.

Sample

This study uses an intention-to-treat analysis strategy, which requires inclusion of all individuals who were randomized in an experimental study design, including those who did not attend follow up (Hollis and Campbell 1999; Armijo-Olivo et al. 2009). Therefore, the analytic sample for this study includes (a) all participants that were randomly assigned to either the one dose group (comparison) or the three dose group (experimental) after the first follow-up survey (T2) and (b) had at least one follow up survey with an identification code that could be matched to the pretest survey (T1). The latter criterion was critical because only the follow-up surveys contained information regarding each participant's intervention assignment. Of the 2021 students who agreed to participate in the longitudinal study by completing the first follow-up survey, a total of 1390 students (837 in the experimental group and 553 in the comparison group) had at least one follow up survey that matched to T1 (69 %). The sample was 62.4 % female, 45.2 % White, 29.8 % Asian, 8.6 % Black, 7.5 % Latino/a and 8.8 % "Other". 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 males students in the analytic sample ($p < 0.01$).

Measures

Table 1 presents the descriptive statistics for all study variables.

Bystander Intentions

Items from the *Bystander Attitude Scale, Revised (BAS-R)* (McMahon et al. 2014) were used in this study. A total of 11 items were listed, each stating a different bystander behavior in a 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, with 1 = "Unlikely" to 5 = "Very likely." For the purposes of this study, three subscales were adapted, including Primary Prevention Bystander attitudes (n = 3 items), Secondary Prevention Bystander Attitudes (n = 6), and Proactive Bystander attitudes (n = 2). Example items from each of the subscales, respectively, are, "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?" "In the future, how likely are you to confront a friend if I heard rumors that they had forced someone to have sex", and "In the future, how likely are you to join or volunteer with an organization that works to stop rape and abuse"? A latent variable was computed using a composite score of the three subscales ($\alpha = .88$ at T1).

Table 1 Descriptive statistics

Variable	Mean	SD
Bystander efficacy (time 2)	73.1	16.06
Bystander intentions (time 2)		
Factor 1	4.0	0.85
Factor 2	4.1	0.71
Factor 3	2.6	0.97
Bystander efficacy (time 3)	73.3	15.69
Bystander intentions (time 3)		
Factor 1	4.1	0.87
Factor 2	4.1	0.73
Factor 3	2.6	1.05
Bystander efficacy (time 4)	72.8	16.5
Bystander intentions (time 4)		
Factor 1	4.0	0.88
Factor 2	4.1	0.78
Factor 3	2.6	1.02
Bystander behaviors (time 5)	45.2	17.32

Bystander Efficacy

To assess confidence in one's ability to intervene, the *Bystander Efficacy Scale* was used (Banyard et al. 2007). 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 mean across all items (Banyard 2008). At T1, this 18-item scale had good reliability ($\alpha = .91$).

Bystander Behaviors

To measure bystander behaviors, a modified version of the *Bystander Behavior Scale—Revised*, was used (McMahon et al. 2014). Participants were asked to indicate whether they had intervened (since the last survey) in any of 16 items that were identical to the items in the bystander intentions scales. Participants could indicate "yes," "no" or "wasn't in situation" for each item. In order to create the bystander behavior composite variable, the procedures developed by Murphy (Murphy 2013; Murphy et al. in press) were followed. That is, a ratio of the number of times a person intervened was divided by the number of times the person was in the situation. Because students had various opportunities to intervene throughout the study, the ratio reflects a cumulative, bystander intervention score. Specifically, the ratio includes all reported bystander situations and interventions leading up to the present survey time period.

Intervention

Individuals who were in the comparison group were assigned a value of 0. Individuals who were assigned to the experimental group, whether they attended the additional doses or not, were given a value of one.

Missing Data

An analysis of missing data indicated that 759 students had complete data on all variables of interest for this study (e.g., intervention, bystander intentions, bystander efficacy, and bystander behaviors for T2–T5). Attrition and self-generated identification code matching issues accounted for about 11.4, 20.6, 28.7, 32.5 and 36.5 % of the missing data in T2, T3, T4, and T5, respectively. In more detail, 136 cases in T2 had missing information on one or more of the outcome variables of interest due to attrition, matching issues, and item non-response; 263 in T3; 364 in T4, and 423 in T5. Therefore, in the most extreme case in T5, approximately 45 % of the values had to be estimated using imputation.

AMOS uses a maximum likelihood (ML) method of imputation. This method of imputation provides better parameter estimates and standard errors than ad hoc procedures such as listwise deletion, pairwise deletion, and stochastic regression. According to Newman (2003), ML methods of imputation also provide similar parameter estimates and standard errors to multiple imputation methods for missing data that is over 25 %. The ML methods have also been suggested for estimating latent variables (e.g., bystander intentions in this study).

Data Analysis Strategies

First, we examined whether reciprocal causality existed between efficacy and intentions. Data from waves 2, 3, and 4 were used to test for reciprocal causality using AMOS 22 (Arbuckle 2007). Second, structural equation modeling was performed to test whether the intervention influenced behaviors through its effects on the hypothesized mediators of efficacy and intentions. Missing data for all waves were imputed using the maximum likelihood imputation procedures of AMOS 22.0. Data from waves 2, 3, and 4 were used to predict cumulative behavior in wave 5 in these analyses.

We compared four specific models: (1) a baseline model with autoregressive paths; (2) a model with autoregressive effects and efficacy predicting intentions at later time points; (3) a model with autoregressive effects and intentions predicting efficacy at later time points; and, (4) a fully cross-lagged model with the autoregressive effects and both intentions and efficacy predicting each other at later time points. These four models were tested using data from the full sample of respondents. In each of the models, the baseline variables were hypothesized as correlated. Direct paths were then hypothesized between the variables. The residual terms associated with the variables at follow-up were also hypothesized as correlated, and the error terms of the same items measured at baseline (Time 2) and follow-up periods were also hypothesized as correlated. These error terms were hypothesized as correlated because we assumed that factors contributing to measurement error in any particular item or latent construct would be consistent across the time periods. Based on the testing of these four models, we used the best fitting model to then test the impact of the intervention (one vs. three doses) on intentions and efficacy over time, as well as their effect on behaviors at Time 5.

The fit of the models was assessed by several fit indices that are considered to be acceptable measures of fit (Kline 2005): Chi-Square (χ^2), the Comparative Fit Index (CFI), the Normed Fit Index (NFI) and the Root Mean Square Error of Approximation (RMSEA). Non-significant χ^2 values and values of the CFI and NFI greater than .95 are indicative of good model-to-data fit, while lower RMSEA values indicate acceptable model-to-data fit (Hu and Bentler 1999). According to Browne and Cudeck (1992), guidelines for interpreting the RMSEA include: $<.05$ = good fit; .05–.08 = acceptable fit; .08–.10 = marginal fit; $>.10$ = poor fit. Maximum likelihood estimation was used to test the models which included only observed variables. Because we hypothesized a specific direction of effects, one-tailed significance statistics were utilized.

Results

As can be seen in Table 2, each of the four models provided a good fit to the data. The CFI and NFI values were greater than .90, indicating good model-to-data fit. In addition, the RMSEA values were generally below .08, also indicating that each of the models provided a good fit to the data in this analysis. The baseline model is of interest because it allowed us to evaluate the underlying measurement model as well as the stability of the constructs over time. The fit indices in Table 1 indicate that the measurement model provided a good fit to the data. In addition, the autoregressive paths for both efficacy and intentions in the baseline model were statistically significant, indicating that the constructs were stable over the time periods, thereby providing support for Hypothesis 1.

More importantly, however, are the differences in Chi-square values between the models that were compared in our analysis. The Chi-square difference test was of particular interest because it allowed us to compare models that were hierarchically nested. Results of the Chi-square difference test presented in Table 2 show that Model 2, which specified efficacy as predicting intentions at later time points, provided a significantly better fit to the data than the baseline model (Model 1), X^2_{diff} (2) = 121.462, $p < .001$. In addition, Model 3, which specified intentions as predicting efficacy at later time points, also provided a significantly better fit to the data than the baseline model (Model 1), X^2_{diff} (2) = 167.996, $p < .001$. Results of the Chi-square difference test in Table 1 also show the comparison between Model 4, the fully cross-lagged model, and Model 3.

Because Model 3 showed the best fit of the first three models tested, we compared the fit of the fully cross-lagged design to this model to determine if it provided a significantly better fit to the data. Results indicate that Model 4, the fully cross-lagged model, provided a significantly better

Table 2 Summary of model fit indices

Test	Model 1: baseline with autoregressive effects	Model 2: autoregressive effects and BE predicting BI	Model 3: autoregressive effects and BI predicting BE	Model 4: fully cross-lagged model
CFI	.950	.966	.972	.983
NFI	.946	.961	.967	.978
RMSEA	.082	.070	.063	.051
X^2 value	424.404	302.942	256.408	169.935
Df	41	39	39	37
p value	<.001	<.001	<.001	<.001
Difference in X^2 test				
Model compared	–	Model 1	Model 1	Model 3
Change in X^2	–	121.462	167.996	86.473
Change in *df*	–	2	2	2
$p <$	–	.001	.001	.001

fit to the data than Model 3, $X^2_{\text{diff}}\,(2) = 86.473$, $p > .001$. In Model 4, all of the paths were statistically significant. Taken together, these data suggest that Model 4, the fully cross-lagged model, fit the data better than the other competing models tested in this study. As can be seen in the results of the analysis for Model 4 that are shown in Fig. 1, this model included statistically significant factor loadings for intentions, ranging from .41 to .92, for each measure relative to its hypothesized latent factor. In addition, the autoregressive paths for the variables were strong, indicating that the constructs were stable over time. Importantly, the paths between variables at baseline and later time points were also statistically significant, indicating that reciprocal causality existed between efficacy and intentions. These models provided support for the second hypothesis.

The results of our preliminary analysis of reciprocal causality were used to inform our subsequent analysis in which we tested whether the intervention influenced behaviors through its effects on the hypothesized mediators of efficacy and intentions. Partial support was found for the third hypothesis, While the intervention showed a positive, indirect impact on behaviors, it was a weak effect. Results of this analysis are presented in Fig. 2 and Table 3. This model provided a good fit to the participant data (CFI = .976; NFI = .970; RMSEA = .051; $X^2(54) = 245.537$, $p < .001$), and explained 28 % of the variance in intentions at T3, 22 % of the variance in efficacy at T3, 28 % of the variance in intentions at T4, 30 % of the variance in efficacy at T4, and 13 % of the variance in behaviors at T5.

The model shown in Fig. 2 indicates that the intervention had positive, indirect effect on behaviors at Time 5 through its effects on efficacy at T2. Efficacy at T2 and intentions at T2 were also found to have direct and indirect effects on behaviors through their reciprocal effects on variables at later time points. These effects are represented by the bolded arrows in Fig. 2. As can be seen in Fig. 2 and Table 3, the intervention had a small, but positive, statistically significant effect on efficacy at T2 Efficacy at T2 was then found to have a positive, statistically significant direct effect on behaviors at T5 as well as an indirect effect through its impact on intentions. This is seen through the positive impact of efficacy at T2 on intentions at T3, which had a positive impact on intentions at T4, which had a direct impact on behaviors at T5. Together, these findings suggest a positive impact of the intervention on behaviors. This positive impact occurred through the intervention's effects on intentions and efficacy, variables which demonstrated reciprocal causality over time.

Discussion

The results from this study provide further support for the use of bystander programs on a college or university setting to increase pro-social behaviors to address situations involving sexual violence. Additionally, the results offer a rigorous, longitudinal view of how one particular program affected students' bystander efficacy, intentions, and behaviors. The first hypothesis of the study suggested that both bystander intentions and efficacy will demonstrate strong autoregressive effects over time. This hypothesis was fully supported, indicating that the constructs remained stable across time periods at the group level.

The second hypothesis of the study suggested that a fully cross lagged model in which there is a reciprocal causality between bystander intentions and efficacy will provide a significantly better fit than other models which hypothesize unidirectionality. The results of this study did not support a one-way linear relationship between these

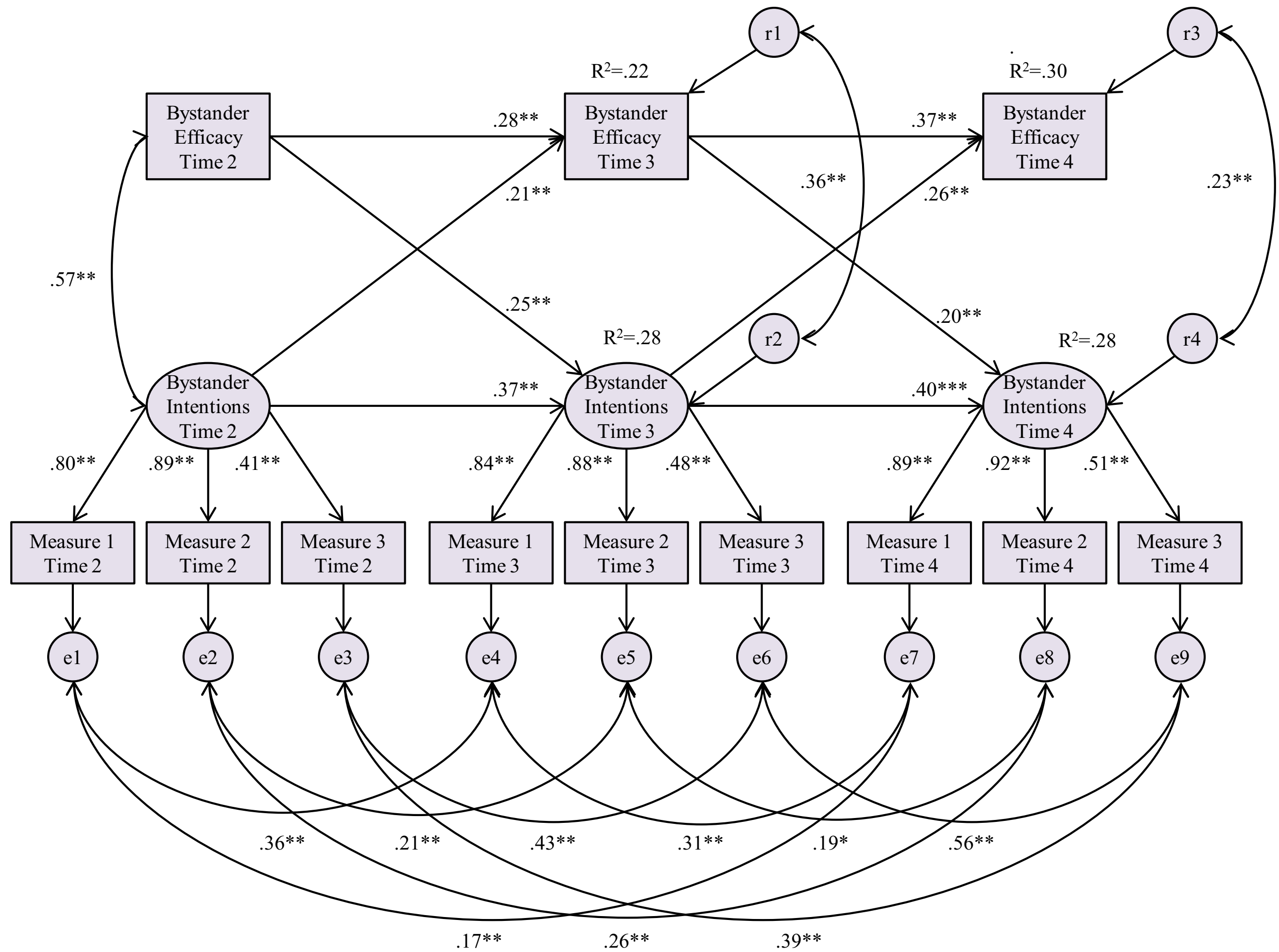

Fig. 1 Fully cross-lagged model with autoregressive effects and bystander efficacy and bystander intentions predicting each other at later time points. *Note* *$p < .01$; **$p < .001$

variables. Instead, the model suggests that there is a more complex, interactive relationship between these constructs, with a reciprocal relationship demonstrated over time. This suggests that an individual's self-efficacy or confidence to intervene in sexual assault situations as a bystander as well as his or her willingness to intervene affect one another and evolve over time. The interrelatedness of these constructs is an important contribution to the current bystander literature and represents a key area for future inquiry and for bystander program implementation.

The third and final hypothesis of the study was that the intervention will affect bystander behavior at Time 5 indirectly through its effect on bystander intentions and efficacy at Time 2. The results partially support this hypothesis, indicating that participation in *SCREAM Theater* had positive indirect effects on behavior through intentions and self-efficacy; these positive effects evolved over time and in a reciprocal fashion. However, the effect was quite weak.

The findings have implications for the development of bystander intervention education programs. According to the results, self-efficacy and bystander intentions both directly influence behavior at different times, but also act as mediators for influencing behavior. Therefore, changing student behavior to prevent sexual violence on campus requires more than just teaching them *how* to intervene. Providing suggestions about what students can do in various situations is important, but bystander intervention programs should also address how to strengthen students' confidence and intentions or willingness to act as bystanders to best prepare them to actually intervene as a bystander when the situation arises. Further research is needed that tests which strategies to improve students' self-efficacy and bystander intentions work best.

Receiving three doses of the intervention, as compared to receiving one, had a small but significant impact on efficacy, intentions, and ultimately on behaviors. Hence, additional doses of learning and practicing bystander

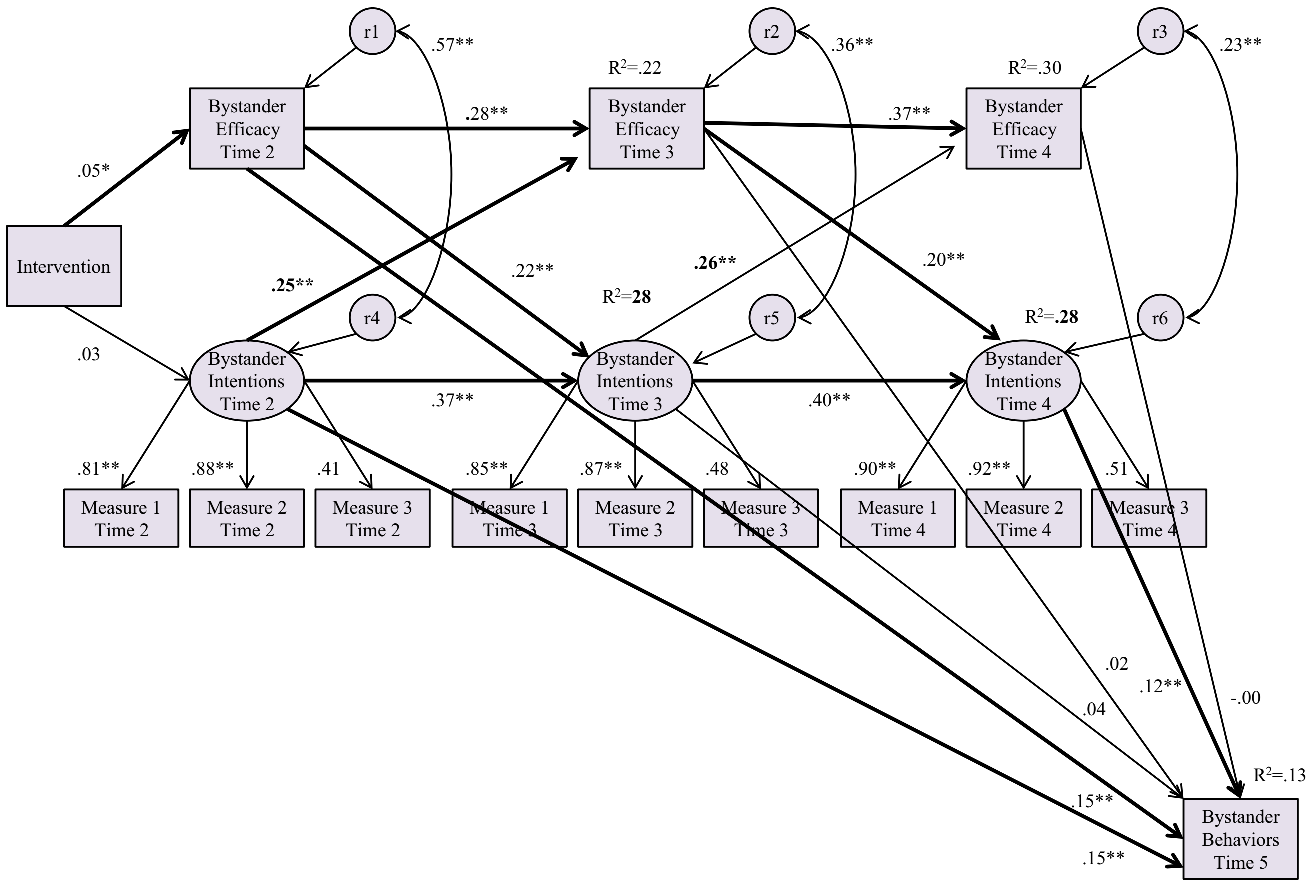

Fig. 2 Structural equation model predicting bystander behaviors. *$p < .05$; **$p < .001$

intervention skills in smaller group settings affected students' confidence and willingness to become active bystanders in preventing sexual violence on campus.

These findings have implications for both program planning and evaluation measurement. Bystander intervention programs would likely benefit from minimally addressing both efficacy and behavioral intent, based on the premise that these work together in encouraging behaviors. To build efficacy and behavioral intent, students should have the opportunity to participate in skill development (Coker et al. 2011), as demonstrated in the current study. Similar to the findings of Ahrens et al. (2011), the results of the current study indicates that small group opportunities for students to discuss and "practice" the application of bystander intervention strategies to real life situations may be an important component for educators to include in curricula. The current study also found a weak yet significant impact on bystander behaviors of those receiving three doses of the program as compared to those receiving only one dose. This supports previous research that suggests greater exposure to prevention messages results in better outcomes (Banyard 2011; Coker et al. 2011). Together, these findings indicate that providing a "one-time" program in large venues such as new student orientations may not be optimal for increasing students' bystander efficacy, behavioral intent, and bystander behaviors in the context of sexual violence. This finding has also been supported in a short term evaluation of the SCREAM program (i.e., McMahon et al. in press).

For measurement purposes, these findings indicate that evaluation efforts should include measure all three constructs of efficacy, intentions, and behavior. Although the goal of many intervention programs is to increase bystander behaviors, the results of this study indicate that this type of change may take time and therefore longitudinal measures are important. Additionally, the measurement of bystander behavior is quite complex (McMahon et al. in press). For example, it is necessary to first determine if students have the opportunity to intervene, if they actually do, and then how they intervene. If students are not in situations where they can act as a bystander, their behavior might not change as a result of participating in a prevention program. Therefore, measuring students levels of bystander efficacy and intentions could serve as intermediate "proxy" outcomes that could eventually impact behaviors when students have such opportunities to intervene.

Table 3 Associations between variables in the cross-lagged model

Path	Standardized regression estimate	Unstandardized regression estimate	95 % CIs	*p* value
INT → BE (T2)	.05	1.58	.02–3.14	.04
INT → BI (T2)	.03	.02	−.02 to .06	.18
BE (T2) → BI (T3)	.22	.01	.01–.01	.000
BI (T2) → BE (T3)	.25	10.18	7.72–12.63	.000
BE (T2) → BE (T3)	.28	.27	.22–.32	.000
BI (T2) → BI (T3)	.37	.47	.38–.56	.000
BI (T3) → BI (T4)	.40	.42	.35–.48	.000
BE (T3) → BI (T4)	.20	.01	.00–.01	.000
BI (T3) → BE (T4)	.26	9.05	7.26–10.84	.000
BE (T3) → BE (T4)	.37	.39	.34–.44	.000
BE (T2) → BB (T5)	.15	.15	.10–.21	.000
BI (T2) → BB (T5)	.15	6.26	3.57–8.95	.000
BE (T3) → BB (T5)	.02	.03	−.03 to .08	.23
BI (T3) → BB (T5)	.04	1.35	−.81 to 3.51	.15
BI (T4) → BB (T5)	.12	3.81	2.06–5.57	.000
BE (T4) → BB (T5)	−.00	.00	−.05	.48
	Total standardized indirect effects			
INT → BB (T5)	.01			
BE (T2) → BB (T5)	.03			
BI (T2) → BB (T5)	.04			
BE (T3) → BB (T5)	.02			
BI (T3) → BB (T5)	.05			

To aid interpretation, the table excludes coefficients for error terms, residual terms, and indicators for latent variables

INT intervention, *BE* bystander efficacy, *BI* bystander intentions, *BB* bystander behaviors, *T2* time 2, *T3* time 3, *T4* time 4, *T5* time 5, *CI* confidence intervals

Limitations and Future Directions

The results of this study should be interpreted within the context of a number of limitations. First, due to the inability to match some survey participants across time periods, a number of students were not represented in the longitudinal analyses. Self-generated identification codes are an excellent, albeit imperfect, way to keep participation in surveys on sensitive topics anonymous and to reduce social desirability bias (Schnell et al. 2010). Future research would benefit from identifying more effective anonymous tracking procedures for longitudinal studies.

Second, this study was not able to use a true "control" group (i.e., individuals who did not receive the intervention) because the university mandates the program for all incoming students. Future research on similar programs should include a true control when possible, in order to better isolate the impact of the intervention. Additionally, because of the length of the study, it is possible that maturation and exposure to other educational messages and experiences may have impacted students' bystander outcomes.

The measurement of bystander behaviors in sexual violence situations is still evolving within the field (McMahon et al. in press). While the current bystander behavior measure assessed whether students participated in certain opportunities to intervene, this measure may not have captured other possible bystander intervention behaviors. Beyond knowing whether students intervened or not, there is no information about how they intervened, or whether it resulted in a positive outcome. Hence, a more sophisticated bystander behavior measure is needed.

Additionally, the amount of behavior explained is small (13 %) which indicates that there are variables influencing bystander behavior other than self-efficacy and behavioral intent. The larger bystander literature indicates that a number of personal, situational, and environmental factors may also influence an individual's decision to intervene in a sexual violence situation (Banyard 2011). Further work is needed to test these various ecological factors and their interplay to determine how college campuses can best influence student's bystander behaviors.

Despite these limitations, this study contributes to our understanding of the key constructs of bystander intentions, bystander efficacy, and bystander behavior in the context of campus sexual violence. The results demonstrate that the relationship among these variables is complex and evolves over time. There is currently a dearth of research examining these variables longitudinally, and future research is needed to explore the relationships among these constructs and especially over time.

Acknowledgments Preparation of this manuscript was supported by a Grant from the Centers for Disease Control and Prevention (Grant Number: 1R01CE001855-01, PI: McMahon, CO-PI: Postmus). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC.

References

Ahrens, C., Rich, M., & Ullman, J. (2011). 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, 17*(6), 760–776.

Ajzen, I. (1988). *Attitudes, personality, and behavior*. Chicago: Dorsey Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211.

Arbuckle, J. L. (2007). *Amos* (Version 7.0) [Computer Program]. Chicago: SPSS.

Armijo-Olivo, S., Warren, S., & Magee, D. (2009). Intention to treat analysis, compliance, drop-outs and how to deal with missing data in clinical research: A review. *Physical Therapy Reviews, 14*(1), 36–49.

Banyard, V. (2008). Measurement and correlates of prosocial bystander behavior: The case of interpersonal violence. *Violence and Victims, 23*(1), 83–97.

Banyard, V. L. (2011). Who will help prevent sexual violence: Creating an ecological model of bystander intervention. *Psychology of Violence, 1*(3), 216–229.

Banyard, V. L. (2014). Improving college campus–based prevention of violence against women a strategic plan for research built on multipronged practices and policies. *Trauma, Violence, & Abuse, 15*(4), 339–351.

Banyard, V., & Moynihan, M. (2011). Variation in bystander behavior related to sexual and intimate partner violence prevention: Correlates in a sample of college students. *Psychology of Violence, 1*(4), 287–301.

Banyard, V. L., Moynihan, M. M., & Plante, E. G. (2007). Sexual violence prevention through bystander education: An experimental evaluation. *Journal of Community Psychology, 35*(4), 463–481.

Banyard, V. L., Plante, E., & Moynihan, M. M. (2004). Bystander education: Bringing a broader community perspective to sexual violence prevention. *Journal of Community Psychology, 32*(1), 61–79.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*, 230–258. doi:10.1177/0049124192021002005.

Black, B., Weisz, A., Coats, S., & Patterson, D. (2000). Evaluating a psychoeducational sexual assault prevention program incorporating theatrical presentation, peer education, and social work. *Research on Social Work Practice, 10*(5), 589–606.

Burn, S. (2009). A situational model of sexual assault prevention through bystander intervention. *Sex Roles, 60*(11–12), 779–792.

Cappadocia, M. C. (2012). Individual motivations and characteristics associated with bystander intervention during bullying episodes among children and youth. *Canadian Journal of School Psychology, 27*(3), 201–216.

Casey, E., & Lindhorst, T. (2009). Toward a multi-level, ecological approach to the primary prevention of sexual assault: Prevention in peer and community contexts. *Trauma Violence & Abuse, 10*(2), 91–114.

Coker, A., Cook-Craig, P. G., Williams, C. M., Fisher, B. S., Clear, E. R., Garcia, L. S., et al. (2011). Evaluation or green dot: An active bystander intervention to reduce sexual violence on college campuses. *Violence Against Women, 17*, 777–796.

Christensen, M. C. (2013). Using theater of the oppressed to prevent sexual violence on college campuses. *Trauma, Violence, & Abuse, 14*(4), 282–294.

Darley, J. M., & Latane, B. (1968). Group inhibition of bystander intervention. *Journal of Personality and Social Psychology, 10*, 215–221.

Department of Education. (2014). Violence Against Women Act; Final Rule. *Federal Register, 79*(202), 62752–62790.

Exner, D., & Cummings, N. (2011). Implications for sexual assault prevention: College students as prosocial bystanders. *Journal of American College Health, 59*(7), 655–657.

Fabiano, P. M., Perkins, H. W., Berkowitz, A., Linkenbach, J., & Stark, C. (2003). Engaging men as social justice allies in endingviolence against women: Evidence for a social norms approach. *Journal of American College Health, 52*(3), 105–112.

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., et al. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*(4), 517.

Gini, G., Albiero, P., Benelli, B., & Altoe, G. (2008). Determinants of adolescents' active defending and passive bystanding behaviour in bullying. *Journal of Adolescence, 31*(1), 93–105.

Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ, 319*(7211), 670–674.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. doi:10.1080/10705519909540118.

Huston, T. L., Ruggiero, M., Conner, R., & Geis, G. (1981). Bystander intervention into crime: A study based on naturally-occurring episodes. *Social Psychology Quarterly, 44*(1), 14–23.

Katz, J., & Moore, J. (2013). Bystander education training for campus sexual assault prevention: An initial meta-analysis. *Violence and Victims, 28*(6), 1054–1067.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Latane, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Latane, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin, 89*(2), 308–324.

Lonsway, K. A. (1996). Preventing acquaintance rape through education: What do we know. *Psychology of Women Quarterly, 20*(2), 229–265.

McMahon, S., Allen, C. T., Postmus, J. L., McMahon, S. M., Peterson, N. A., & Lowe Hoffman, M. (2014). Measuring bystander attitudes and behavior to prevent sexual violence. *Journal of American College Health, 62*(1), 58–66.

McMahon, S., Banyard, V. L., Palmer, J. E., Murphy, M., & Gidcyz, C. A. (in press). Measuring bystander behavior in the context of sexual violence prevention: Lessons learned and new directions. *Journal of Interpersonal Violence*. doi:10.1177/0886260515591979.

McMahon, S., Winter, S., Palmer, J., Postmus, J. L., Peterson, N. A., Zucker, S., et al. (in press). A randomized control trial of a peer education theater bystander program to prevent sexual violence. *Health Education Research*. doi:10.1093/her/cyv022.

Moynihan, M., Banyard, V. J., Arnold, J. S., Eckstein, R. P., & Stapleton, J. G. (2011). Sisterhood may be powerful for reducing sexual and intimate partner violence: An evaluation of the bringing in the bystander in-person program with sorority members. *Violence Against Women, 17*(6), 703–719.

Murphy, M. (2014). *Sexual assault-specific bystander behavior: Accounting for opportunity in a prospective analysis of the effects of individual, social norms, and situational variables* (Doctoral dissertation). Retrieved from OhioLink Dissertation and Theses Center. (Document No. ohiou1389008840).

Murphy, M. J., Dardis, C. M., Wilson, M. S., Gidycz, C. A., & Berkowitz, A. D. (in press). Predictors of sexual assault-specific prosocial bystander behavior and intentions: A prospective analysis. *Violence Against Women*.

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods, 6*(3), 328–362.

Schnell, R., Bachteler, T., & Reiher, J. (2010). Improving the use of self-generated identification codes. *Evaluation Review, 34*, 391–418.

The White House. (2014). *The first report of the White House task force to protect students from sexual assault.* Retrieved May 9, 2014 from http://www.whitehouse.gov/sites/default/files/docs/report_0.pdf.

Thornberg, R., Tenenbaum, L., Varjas, K., Meyers, J., Jungert, T., & Vanegas, G. (2012). Bystander motivation in bullying incidents: To intervene or not to intervene? *Western Journal of Emergency Medicine*, *13*(3), 247–252.