Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications

Psychology, Department of

11-2019

What Happened Next? New Measures of Consequences of Bystander Actions to Prevent Interpersonal Violence

Victoria Banyard
Rutgers University

Elizabeth Moschella
University of New Hampshire, Durham

John Grych
Marquette University, john.grych@marquette.edu

Ernest N. Jouriles
Southern Methodist University

Follow this and additional works at: https://epublications.marquette.edu/psych_fac

Part of the [Psychology Commons](#)

Recommended Citation

Banyard, Victoria; Moschella, Elizabeth; Grych, John; and Jouriles, Ernest N., "What Happened Next? New Measures of Consequences of Bystander Actions to Prevent Interpersonal Violence" (2019). *Psychology Faculty Research and Publications*. 472.

https://epublications.marquette.edu/psych_fac/472

Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications/College of Arts and Sciences

This paper is NOT THE PUBLISHED VERSION; but the author's final, peer-reviewed manuscript. The published version may be accessed by following the link in the citation below.

Psychology of Violence, Vol. 9, No. 6 (November 2019): 664-674. [DOI](#). This article is © American Psychological Association and permission has been granted for this version to appear in [e-Publications@Marquette](#). American Psychological Association does not grant permission for this article to be further copied/distributed or hosted elsewhere without express permission from American Psychological Association.

What Happened Next? New Measures of Consequences of Bystander Actions to Prevent Interpersonal Violence

Victoria Banyard

School of Social Work, Rutgers The State University of New Jersey;

Elizabeth Moschella

Department of Psychology, University of New Hampshire

John Grych

Department of Psychology, Marquette University

Ernest Jouriles

Department of Psychology, Southern Methodist University

Interpersonal violence includes experiences of unwanted sexual contact, rape, physical and psychological dating violence, and controlling behavior such as stalking, and it disproportionately affects late adolescents and young adults (Black et al., 2011; Coker, Williams, Follingstad, & Jordan, 2011; Gidycz, Orchowski, King, & Rich, 2008).

Many incidents of sexual assault and dating violence are witnessed by one or more third parties, who are referred to as bystanders (Hamby, Weber, Grych, & Banyard, 2016; Planty, 2002). These individuals have the opportunity to take action to model violence-suppressing norms, prevent a risky situation from escalating, interrupt an act of violence in progress, and support a victim or confront a perpetrator after an incident. To date, research on bystander action has focused mainly on factors that lead to someone either intervening or failing to respond in a threatening situation. This information was used to develop programs to promote bystander behavior. However, much less is known about what happens after bystanders respond—that is, the consequences of bystander behavior. As bystander training programs are widely implemented at colleges and high schools, it is imperative to better understand the consequences people may face for intervening. The aim of the current study was to develop and evaluate measures that assess what happens after bystanders act, including the responses of victims, perpetrators, and other bystanders, and their feelings about their actions.

Key theories of bystander behavior propose that individuals consider the pros and cons of taking action when they decide whether or not to intervene (Banyard, 2015; Dovidio, Piliavin, Gaertner, Schroeder, & Clark, 1991; Latané & Darley, 1970), and research shows that one barrier to intervening is the fear of adverse consequences, which can include physical injury and harming their relationship with the victim or perpetrator (Bennett & Banyard, 2016; Bennett, Banyard, & Garnhart, 2014; Burn, 2009; Hamby et al., 2016; Yule & Grych, 2017). Gender differences have been reported in bystander attitudes, perceived barriers, and behavior (Banyard & Moynihan, 2011; Yule & Grych, 2017), and men and women may also face different consequences when they intervene in potentially dangerous situations. Burn (2009) hypothesized that sexual assault in particular may be more salient for women because of their greater risk for victimization, which may enhance their awareness of situations when action is needed and increase their motivation to intervene. At the same time, the salience of their risk for victimization also may make women more concerned for their safety if they intervene. Men, in contrast, endorse greater levels of rape myths than women, which may impede their sense of responsibility to act on behalf of a potential victim (Burn, 2009; McMahon, 2010). However, they may be more likely to help when other bystanders are women due to the potential activation of traditional gender norms related to chivalrous male behavior (Levine & Crowther, 2008). Little research has examined whether there are gender differences in the consequences experienced when bystanders intervene. Because the consequences of intervening may impact future helping, we examined whether men and women report different reactions when they intervened in potentially dangerous situations.

Understanding the consequences of bystander intervention requires reliable and valid measurement of what happens after bystanders respond. Researchers are now beginning to investigate consequences of bystander actions (Krauss et al., 2017; Witte, Casper, Hackman, & Mulla, 2017) and have noted both positive (e.g., feelings of pride) and negative outcomes (e.g., verbally threatened). However, the absence of published, psychometrically sound, measures of bystander consequences that considers both negative and positive consequences hinders research on this topic. Therefore, the purpose of the current study was to evaluate the preliminary psychometric properties of five newly created self-report scales of consequences to bystander action. One scale was designed to provide a general overview of consequences experienced by bystanders. The other four scales were incident-specific and designed to capture different aspects of bystander outcomes from the vantage point of the bystander: outcomes experienced by the bystander themselves, including their feelings about having taken action; reactions of the victim whom the bystander intended to help; reactions of the potential perpetrator; and responses from other bystanders. These five scales are collectively given a preliminary title of the Compendium of Bystander Consequences (CBC) to facilitate presentation of our findings here.

Hypothesis 1: Each of the CBC scales were expected to have two subscales—one focused on positive reactions or outcomes and one focused on negative repercussions of actions taken.

Hypothesis 2: Each of the scales was expected to demonstrate adequate reliability (DeVellis, 1991; Nunnally & Bernstein, 1994), and scores on consequences of a specific incident were hypothesized to correlate significantly with a shorter, broad measure of general bystander consequences.

Hypothesis 3: Previous work on gender differences in attitudes, barriers, and bystander behaviors (Yule & Grych, 2017) suggested that women would report higher levels of positive consequences (given more positive attitudes toward action and higher levels of action reported in other studies) and greater negative consequences given concerns about risk for victimization. Therefore, we hypothesized that women would report higher levels of positive and negative consequences than men.

Method

Initial Stages of Measurement Development

In the early stages of measurement development, we generated and reviewed items pertaining to outcomes experienced by bystanders as a result of helping in situations of sexual and dating violence. To generate the initial list of items with face and content validity, we reviewed previous research conducted on this topic. We also attended bystander prevention trainings and listened to examples and stories of bystanders' experiences. After an initial list of items was generated, focus groups (four focus groups with a total of 31 participants, five men and 26 women of traditional college age) were conducted to get feedback about whether items were clearly written and captured the scope of experiences of young adults. We also collected pilot data ($n = 139$) to gauge how long the survey would take participants to complete and how well items clustered together. A full list of the retained items for each scale can be found in Tables 1–5, which describe the factor analyses performed as part of the study described below.

Table 1. Factor Loadings for Exploratory Factor Analyses (EFA) and Confirmatory Factor Analyses (CFA) for Items in the General Outcomes of Bystander Action Scale

	EFA loading ^a		CFA 1 loading ^b	
Factor	1	2	1	2
1. People praised me for what I did.	.23	.85	.84	
2. A friendship was strengthened because of what I did.	.07	.81	.71	
3. <i>I saw positive posts on social media about what I did.</i>	.61	.60		
4. People said positive things about me to others because of what I did	.27	.87	.90	
5. I was harassed because of what I did.	.86	.14		.84
6. People said negative things about me to others because of what I did.	.86	.09		
7. I saw negative posts on social media about what I did.	.89	.17		.94
8. I was physically hurt because of what I did.	.91	.19		.91
9. I was threatened physically because of what I did.	.87	.25		
10. I was threatened verbally because of what I did.	.78	.35		
11. I was interviewed as a witness because of what I did.	.76	.33		
12. I got in trouble as a result of my action (e.g., charged with underage drinking).	.89	.25		.90
13. My involvement ended up costing me a lot of time (e.g., time it took to				

intervene, being interviewed later, talking with friends about what I did).	.77	.33		.85
Cronbach's α .			.85	.95

Note. Items in bold in EFA and CFA indicate final item for each subscale. Italicized items were dropped from CFA due to factorial complexity in EFA.

^a N = 162, Kaiser–Meyer–Olkin = .92. ^b N = 212.

Table 2. Factor Loadings for Exploratory Factor Analyses (EFA) and Confirmatory Factor Analyses (CFA) of Items on the Bystander Feelings Measure

	EFA ^a			CFA 1 ^b		
Factor	1	2	3	1	2	3
1. Felt unsure what I did was helpful.	.27	–.27	.72	.82		
2. Felt confused about what to do.	.26	–.22	.77	.78		
3. <i>Worried I made the situation worse.</i>	.42	–.09	.74			
4. Wished I hadn't stepped in.	.79	–.19	.30			.86
5. Felt bad for intruding on the situation.	.78	–.14	.29			.86
6. Worried about what others thought about me.	.82	–.13	.19			.83
7. Felt embarrassed about what I did.	.86	–.03	.23			.86
8. <i>Felt scared for my own safety.</i>	.58	.12	.40			
9. Regretted what I did.	.88	–.07	.16			.87
10. Felt I would never do that again.	.86	–.07	.22			
11. <i>Felt like others would not have intervened if I did not.</i>	.18	.13	.43			
12. Felt like I helped.	.04	.86	–.08		.77	
13. Felt good about doing something.	–.04	.84	–.06		.87	
14. Felt relieved that it went okay.	–.09	.85	.03		.69	
15. Felt proud of doing something.	–.03	.87	–.06		.78	
16. Felt I did the right thing.	–.08	.84	–.11		.76	
17. Felt like would do the same thing, if it happened again.	–.27	.75	–.06			
Cronbach's α .				.75	.90	.92

Note. Items in bold in EFA and CFA indicate final item for each subscale. Italicized items were dropped from CFA due to factorial complexity in EFA.

^a N = 206, Kaiser–Meyer–Olkin = .90. ^b N = 291.

Table 3. Factor Loadings for Exploratory Factor Analyses (EFA) and Confirmatory Factor Analyses (CFA) of Items on Victim Reactions to Bystander Action Scale

	EFA ^a		CFA 1 ^b	
	1	2	1	2
Negative reaction				
1. Person I tried to help got mad or upset at me.	–.15	.75		.70
2. Person I tried to help got in trouble because of my actions. ^c	–.05	.56		—
3. The action harmed my friendship with the person I tried to help. ^c	–.13	.77		.80
4. Person I tried to help told me to mind my own business. ^c	–.21	.71		.76
Positive reaction				
5. Person I tried to help was safe because of my actions.	.62	–.12	.62	
6. The action strengthened my friendship with the person I tried to help	.74	–.01	.73	
7. Person I tried to help thanked me for stepping in.	.78	–.16	.79	
8. Person I tried to help was relieved or felt better.	.79	–.27	.82	
9. Person I tried to help indicated I was helpful.	.85	–.15	.81	
Cronbach's α .			.84	.74

Note. Items in bold in EFA and CFA indicate final item for each subscale.

^a N = 187, Kaiser–Meyer–Olkin = .79. ^b N = 288. ^c Indicates endorsement by less than 20% of participants.

Table 4. Factor Loadings for Exploratory Factor Analyses (EFA) and Confirmatory Factor Analyses (CFA) of Items on Perpetrator Reactions to Bystander Action Scale

	EFA ^a		CFA 1 ^b	
Factor	1	2	1	2
Items				
Negative reaction				
1. Person who was harming got mad or upset at me.	–.13	.81	.70	
2. Person who was harming got in trouble because of my actions.	.20	.54		
3. The action harmed my friendship with the person who was harming.	.01	.58	.54	
4. Person who was harming told me to mind my own business.	–.04	.81	.73	
Positive reaction				
5. Person who was harming was safe because of my actions.	.47	.06		
6. The action strengthened my friendship with the person who was harming. ^c	.75	– .06		.57
7. Person who was harming thanked me for stepping in. ^c	.79	– .02		.71
8. Person who was harming was relieved or felt better.	.81	.03		.85
9. Person who was harming indicated I was helpful.	.84	.04		.92
Cronbach's α s			.67	.84

Note. Items in bold for EFA and CFA indicate final item for each subscale.

^a N = 165, Kaiser–Meyer–Olkin = .74. ^b N = 268. ^c Indicates endorsement by less than 20% of participants.

Table 5. Factor Loadings for Exploratory Factor Analyses (EFA) and Confirmatory Factor Analyses (CFA) of Items on Others' Response to Bystander Action Scale

	EFA ^a			CFA 1 ^b		
Factor	1	2	3	1	2	3
Negative feedback						
1. Others made fun of me for what I did. ^c	.79	.10	.17			
2. Others got mad at me for what I did. ^c	.77	–.07	.10			
3. Others said I did the wrong thing. ^c	.72	–.11	–.07	.59		
4. Others told me it was none of my business.	.63	.13	.02	.63		
5. Others tried to stop me. ^c	.67	.02	.13			
6. Others bad-mouthed me for what I did. ^c	.74	–.01	.07	.75		
Bystander apathy						
11. Others acted like nothing happened (ignored what I did).	.14	.66	–.09			
12. Others walked away from the situation before I did something.	–.06	.83	.12		.59	
13. Others walked away from the situation after I did something.	.08	.82	.12		.60	
14. Others walked away from the situation while I was doing something.	.17	.81	.16		.92	
Positive feedback						
7. Others praised me for what I did.	–.12	–.02	.82			.81
8. Others thanked me for what I did.	.10	.12	.83			.69
9. Other joined in to help, because of my example.	.03	.08	.67			

10. Others told me I did the right thing.	.04	.10	.86			.77
Cronbach's α .				.64	.80	.79

Note. Items in bold for EFA and CFA indicate final item for each subscale.

^a N = 169, Kaiser–Meyer–Olkin = .78. ^b N = 255. ^c Indicates endorsement by less than 20% of participants.

Participants

Participants were recruited through a university psychology subject pool at a medium-sized New England state campus ($N = 674$) and Amazon's Mechanical Turk (MTurk; $N = 717$). The university subject pool produced a fairly homogenous sample in terms of ethnicity, sexual orientation, and gender, and therefore we collected additional data using MTurk to expand the diversity of the sample.

The psychology subject pool participants were recruited from a number of lower level psychology classes. Specific screening was prohibited by the SONA department process. The average age was 19.07 ($SD = 1.28$). The sample was 71.8% female ($n = 479$) and 27.9% male ($n = 356$); two participants identified as "other." Most of the sample identified as heterosexual (95.3%) and Caucasian (89.6%). The majority of participants were in their first or second year (81.9%). Only 354 of these participants reported helping in an instance of interpersonal violence and received the newly developed CBC measures.

We obtained the MTurk sample through an initial screening of 4,322 individuals who chose to participate in this study. A total of 717 participants met the screening criteria (described below) and took the full survey (16.6% of the sample who selected the study on MTurk). For those sampled from MTurk ($n = 717$), the average age was 22.25 ($SD = 1.57$). There was a fairly even gender distribution, where males made up 49.7% ($n = 356$) and females made up 50.1% ($n = 359$) of the sample; there were two participants who identified as "other." The majority identified as heterosexual (83.1%). A little more than half the sample identified as Caucasian (57.4%), followed by Asian (18.3%), Hispanic (7.8%), African American (7.7%), American Indian (4.6%), Pacific Islander (0.7%), and other (0.6%); a small percentage (2.9%) identified as two or more races. The majority had at least some college experience, an associate's degree, or a bachelor's degree (81.9%).

The psychology subject pool and MTurk samples were significantly different in terms of their gender (psychology subject pool included more females), age (MTurk participants were older), ethnicity (MTurk had more racial diversity), and sexual orientation (psychology subject pool included more heterosexual participants). The two samples also differed on the bystander consequences measures, particularly the negative consequences (MTurk sample reported significantly higher scores on negative consequences; please see available online supplemental table for more details).

Procedures

All participants were first given a consent form that described the study and the potential risks. Participants answered questions about a list of bystander behaviors specific to sexual and dating violence (Banyard, Moynihan, Cares, & Warner, 2014), including the frequency of consequences of these behaviors, and then about a specific recent bystander incident (some additional scales, analyzed and reported on elsewhere, were left off the MTurk survey to make it shorter [Moschella & Banyard, 2018]). To ensure that participants understood the types of situations intended, they were given definitions of "harassing comments, dating violence, unwanted sexual advances, and controlling behavior by a partner." Participants were also asked to briefly describe the incident they were using to answer questions about consequences, and inspection of one-third of these responses indicated that participants generally understood the question and were answering based on incidents that are included in widely used measures of bystander behavior. It is beyond the scope of the current article to describe these analyses, but two raters achieved adequate interrater reliability following established content coding procedures (McHugh, 2012). Institutional review board approval was obtained for data collection with both samples, and participants were given a debriefing information sheet with resources for national crisis centers at the end of the survey.

Several aspects of the survey procedures were slightly different based on recruitment method (i.e., participant pool vs. MTurk). For the MTurk sample, the study was described using the following text: "We are conducting an

academic survey about outcomes of bystander behavior that you may have experienced as a result of helping in a situation at risk for sexual assault, harassment, or dating violence. If you agree to participate in this study, you will be asked to take part in a brief 3–5 minute screening that will ask about demographic information and helping behavior in which you will be paid \$.10. If you qualify for this survey, you will be directed to a 40 minute survey in which you will be paid a bonus of \$2.50. Select the link below to complete the survey. At the end of the survey, you will receive a code to paste into the box below to receive compensation for taking our survey.” For MTurk participants to proceed to the full study, they had to be 18–24 years old, reside in the United States, and have indicated that they helped someone in a situation involving harassment, dating violence, unwanted sexual advances, or controlling behavior from a partner in the past 6 months (defined for participants during the first set of questions) using the following instructions: “In the past 6 months have you intervened or helped someone when you thought they might be at risk for one of the following types of problems (remember, we are not asking if you have experienced these—just if you were in a situation where you saw someone else at risk for one of these and tried to help)” (see footnote for defined types of incidents). Participants received \$0.10 for answering the screening questions. Participants who qualified for the survey were redirected to another survey that took approximately 40 min to complete and earned an additional \$2.50.

Individuals in the Psychology Department subject pool could take the survey online or in-person. The study was described using the following summary: “This study is about understanding bystander action that you may have engaged in and the situational outcomes that you may have experienced in response to risk for sexual assault, dating violence, and harassment.” Participants could choose an alternate assignment to receive course credit if they did not wish to do the survey: “To receive research credit, you can either choose to: (a) complete one 40–60 min online questionnaire worth 1 research credit. You must be at least 18 years old to participate, or (b) complete an alternative assignment for equivalent course credit. You will be assigned to read several research articles and briefly meet with a research team member to discuss them.” Most took the survey online (80.9% of this sample, $n = 545$), whereas 19.1% ($n = 129$) completed the survey in-person in groups of 30 to 50 students. Participants answered demographic items and questions about bystander behavior and general consequences while being given the option to indicate that they had not intervened in a bystander situation. Participants then received the same list of definitions as the MTurk sample and were asked to indicate which of those types of situations they had most recently been in (or to indicate they had not been a bystander in the past 6 months). Participants who indicated they did not help in those situations were redirected (in the online survey) or instructed to skip ahead (in the in-person survey) to questions about bystander attitudes that are not included in the current analyses. Psychology participant pool participants received partial course credit for completing the survey.

Measures

Bystander Action Consequences—General

General Outcomes of Bystander Behavior is a newly developed questionnaire consisting of 13 items inquiring about the frequency of outcomes experienced after “trying to help someone who was at risk for sexual assault, relationship abuse, or stalking.” The measure is designed to provide a general overview of consequences not tied to any one specific incident but across all of the times a person may have helped. We created this measure as a briefer alternative to the incident-specific measures described below. Participants indicated how often they had experienced both positive and negative outcomes in the past 6 months on an 8-point scale, ranging from 0 to “8 or more” times. “Not applicable” was an answer choice for individuals who had not been in a “situation where you tried to help someone who was at risk.” Sample items included, “people praised me for what I did” and “I was threatened verbally because of what I did.” Table 1 presents the list of items.

In addition to the Bystander Action Consequences—General, participants completed four newly developed

measures that were designed to be used in relation to a specific incident of bystander action. For each, participants were asked to “please think about the most recent time you were in a situation like the ones we just asked you about and you did something to try to help.” Participants were asked to describe whether the situation involved harassment, dating violence, unwanted sexual advances, or controlling behavior. For each of the following measures, participants were asked to bring the situation to mind and to answer questions related to what happened.

Bystander Feelings

This measure included 17 items that inquired about positive, negative, and unsure feelings after engaging in bystander behavior. Participants were asked to, “think about how you felt about the situation and how much you had each of the following thoughts or feelings immediately following that specific situation.” An example of one of the five positive items was, “felt like I helped.” Negative feelings included eight items like, “wished I hadn’t stepped in” or “felt embarrassed about what I did.” Three items tapped feelings of uncertainty, such as “felt unsure what I did was helpful” or “felt confused about what to do.” Participants indicated how much they felt a certain way on a 4-point scale ranging from “not at all” to “a lot.” Table 2 presents a full list of items.

Bystander Action Response—Victim

This measure included nine items asking about positive and negative reactions of the victim after bystanders have taken action. Participants indicated if they had experienced a range of positive or negative outcomes in response to a situation where they took action by answering “Yes” or “No.” Instructions stated, “think about how the person you thought needed help reacted during that specific situation.” Sample items included, “person I tried to help told me to mind my own business” or “person I tried to help thanked me for stepping in” (see Table 3).

Bystander Action Response—Perpetrator

This scale included nine items asking about how the perpetrator responded following the participant’s bystander behavior. Participants indicated if they had experienced a variety of positive or negative outcomes by answering “Yes” or “No.” Instructions were, “think about how the person who was potentially harming reacted during that specific situation.” Example items include, “the action strengthened my friendship with the person who was harming” or “person who was harming got mad or upset with me.” Table 4 presents these items.

Bystander Action Responses—Other

Finally, participants completed a series of questions related to the reactions of other people present in the situation. This measure included 14 items that ask about positive, negative, and apathetic reactions of “other bystanders like your friends who were around during that specific situation, or maybe saw or heard about it but were not directly involved.” No one specific bystander was asked about to keep the measure short. Rather, participants were asked to describe “others” more generally, whether that was one or several other people. “Please circle how much they reacted like the following statements.” Sample items included, “others acted like nothing happened” or “others praised me for what I did” (see Table 5 for full list of items).

Social Desirability Scale-17

The psychology subject pool sample completed the Social Desirability Scale-17, which measures the degree to which participants portray themselves in a positive light (Stöber, 2001). The scale consists of 17 true (coded as 1) or false (coded as 0) items, such as “I take out my bad moods on others now and then.” Average scores were calculated for each participant, with higher scores indicating higher rates of social desirability ($M = .59$, $SD = .13$).

Cronbach's α for the current sample was .70. This scale has demonstrated validity in a U.S. context (Blake, Valdiserri, Neuendorf, & Nemeth, 2006).

Data Analysis

We sought to establish the preliminary reliability and validity of the new measures based on both internal structure of the scales and their relationship to other variables. To explore the factor structure of these new measures, the psychology subject pool and MTurk samples who indicated having taking bystander action were combined ($n = 966$), and then the full sample was randomly divided into three approximately equal subsamples (Cudeck, & Browne, 1983; Floyd & Widaman, 1995).

The first subsample ($n = 327$) was used to conduct Principal Components Exploratory Factor Analyses (EFA) using SPSS 24.0 to estimate the underlying factor structure of each scale (Warner, 2012) using the default orthogonal rotation (Varimax). MacCallum, Widaman, Zhang, and Hong (1999) noted that the minimum N for factor analyses varies, but is reduced for analyses with high item communalities (above .6). With high communality and sufficient overdetermination (we hypothesized two or three factors per scale and had close to 10 or more items for each scale) they found sample sizes as low as 60 to be sufficient. For the analyses reported below, moderate to high communality was found for items in all analyses except for one item in the Bystander Action Response—Perpetrator scale, suggesting that a sample size above 100–200 would be sufficient. Warner (2012) recommended $N > 10$ times the number of items in the analysis which also supports that range. Items were retained if they achieved a factor loading of at least .4 and were not factorially complex (i.e., they did not demonstrate a high loading on multiple factors; Pedhazur & Schmelkin, 1991; Warner, 2012).

Once the initial factor structure was identified using EFA, a confirmatory factor analyses (CFA) using MPlus Version 8 was conducted on the second subsample ($n = 329$; Geiser, 2013). The purpose of this CFA was to evaluate the factor structure from the EFA and then to trim items that did not fit the model well using modification indices and model fit indices. Then, a second CFA was conducted with the final subsample ($n = 310$) to confirm the results of the model resulting from the previous CFA. Consistent with previous research, we used several indicators of model fit: root mean square error of approximation (RMSEA) with good fit defined as $\leq .06$ and adequate fit by $\leq .08$, comparative fit index (CFI) $\geq .95$, Tucker–Lewis index $\geq .95$, and standardized root mean square residual (SRMR) $\leq .08$ (Hu & Bentler, 1998). We used the full study sample to compute Cronbach's α s for the final subscales as one indicator of reliability. We explored gender similarities and differences using multivariate analysis of variance.

Validity was assessed in only a very preliminary way. Specifically, using only the sample from the university psychology participant pool, correlations were computed between the new measures and the measure of social desirability. In addition, Mayer (2016) described how “group differences” can serve as another key form of validity evidence. For the current study, we examined gender differences between male and female participants given that previous bystander measures often find these differences, with women reporting more positive bystander attitudes (e.g., confidence or likelihood to take action).

Approximately one-third of the sample answered only the general consequences or only the incident-specific, but not both (two-thirds answered both sets of questions). For the general consequences measure, participants were asked to think generally about the outcomes they have experienced when helping in situations involving harassment, sexual assault, or dating violence. After completing this measure, participants were then asked about a specific incident of helping and selected the type of situation they most recently helped in (detailed definitions for harassment, sexual assault, dating violence, and controlling behavior were provided). This served as a second screening measure, as participants who said they had not helped in any of these situations were redirected or instructed to skip ahead to questions about bystander attitudes. Listwise deletion of missing data was used for each analysis.

Results

All items were initially inspected for skewness and frequency of endorsement. Overall, only seven items had skewness values outside of -2 to 2 (George & Mallery, 2010). Three items related to responses from victims and four items to responses of others. Overall, at least 20% of participants reported each type of consequence.

Hypothesis 1: Factor Structure for the CBC Bystander Action Consequences

Bystander Action Consequences—General

All items had skewness values within -2 to 2 and were experienced as a consequence by at least 20% of participants. For the 13-item General Outcomes scale, a two-factor solution was initially obtained using EFA, with the first factor explaining 63.07% of the variance and the second factor 13.41% of the variance. All items had high communality values (at least .7). The items on each factor corresponded to positive and negative consequences, respectively. One item about positive posts on social media was factorially complex (it loaded above .4 on both factors) and was dropped (Yong & Pearce, 2013). Although a one-factor solution explained more than 50% of the variance, the fit of a two-factor solution that distinguished positive and negative consequences made more sense theoretically and is more useful practically, and thus was retained.

The first CFA model used all of the 12 items with adequate loadings in the EFA and two factors, but did not produce adequate model fit (RMSEA = .13). Modification indices suggested that Item 11 (i.e., being called as a witness) also fit with the “positive” factor and was dropped as it suggested factorial complexity. Further, three items had covariances that suggested, for parsimony that they could be dropped for redundancy. The CFA model that achieved excellent model fit retained eight items and two factors, RMSEA = .03 (90% confidence interval [CI] [.00, .07]); CFI = .99; TLI = .99, SRMR = .03. Three items made up the Positive subscale and five items made up the Negative Consequences subscale (see Table 1 for full list of items and EFA and CFA loadings). A second CFA was conducted with the third randomly drawn subsample using the eight items and the two-factor model, and it achieved adequate fit (RMSEA = .08, CFI = .98, TLI = .97, SRMR = .02).

Incident-related measures

Bystander Feelings

All items had skewness values less than 2 and were endorsed by at least 20% of participants. In the EFA, the 17-item index of bystander feelings produced a three-factor solution using Varimax rotation, explaining 69.04% of the variance. Table 2 shows the factor loadings. All but one item had communalities above .5. The first factor reflected negative feelings, contained seven items, and explained 40.69% of the variance. The second factor captured positive feelings, included six items, and explained 22.43% of the variance. The third factor explained 6.19% of the variance, included three items, and seemed to capture a sense of uncertainty about what happened. Two items, “worried I made the situation worse” and “felt scared for my own safety,” were factorially complex (loaded above .4 on two factors) and an additional item, “felt like others would not have intervened if I did not”, had a very low communality score (.24) and were dropped from further analyses.

This three-factor solution was examined in a CFA using the second subsample. In the CFA, we excluded the two items that were factorially complex in the EFA and the third item with low communality. Overall the model fit for the 14 items was less than adequate and was improved by trimming two items that had strong covariance with others, producing a final 12-item model with RMSEA = .07 (CI [.06, .09]), CFI = .96, TLI = .95, and SRMR = .04. This also made the number of items for positive and negative feelings equal and the overall measure more parsimonious. A second CFA was performed using the third subsample of data and 12 items produced good model fit that was consistent with the first CFA. The RMSEA = .05 (CI [.03, .07]), CFI = .98, TLI = .98, SRMR = .04.

Bystander Action Response—Victim

Three items (2, 3, 4) had skewness values over 2 but less than 3 (i.e., person I tried to help got in trouble, my helping harmed my friendship with the victim, the victim told me to mind my own business). These items also had low rates of endorsement (only 10–14% of the sample said they experienced the consequence). Nonetheless, given the exploratory nature of this study and our desire to understand even low base rate negative consequences, the items were retained. The nine items related to victim responses to the bystanders' actions produced a two-factor solution in the EFA, explaining 56.36% of the variance. Factor one was made up of four items describing negative reactions by the victim (39.55% of variance), and factor two included five items indicating positive responses by the victim (16.82% of variance). Communalities were moderate to high (.3–.7), but all items had factor loadings above .4 and were retained after the EFA.

The CFA model of these two factors produced adequate fit with an RMSEA = .07 (90% CI [.04, .09]), CFI = .96, TLI = .95, SRMR = .05. The model fit improved by dropping Item 2 ("the person I tried to help got in trouble"), which was not experienced by many participants and had the lowest communality (.32) of all of the items. The new RMSEA = .06 (CI [.02, .08]), CFI = .98, TLI = .97, SRMR = .04. Table 3 presents factor loadings for these analyses. The second CFA also produced adequate fit for the full model from the EFA with RMSEA = .06 (CI [.04, .09]), CFI = .96, TLI = .94, SRMR = .05. Adequate fit was obtained for the revised model without Item 2: RMSEA = .07 (CI [.04, .10]), CFI = .96, TLI = .94, SRMR = .05.

Bystander Action Response—Perpetrator

All items had skewness values between –2 and 2. Two of the positive items (Items 6 and 7) were experienced by less than 20% of participants, but were initially retained given the exploratory nature of this study and the desire to capture even low base rate consequences. A set of nine items (parallel to those asked about victim reactions) aimed to assess reactions by the person who was potentially harming in the situation. The EFA with Varimax solution produced a two-factor solution explaining 52.99% of the variance. The first factor included five items describing positive reactions and explained 31.52% of the variance. The second factor included four items reflecting negative reactions and explained 21.46% of the variance with one item (Item 5) showing low communality (.23), but acceptable factor loading, and other items having moderate communalities (.3–.7).

CFA using these two factors produced adequate model fit but had low loadings for two items (one from each factor) that also had the lowest communalities on the EFA, RMSEA = .07 (CI [.05, .09]), CFI = .95, TLI = .92, SRMR = .07. Dropping those two items marginally improved the model; RMSEA = .07 (CI [.04, .10]), CFI = .97, TLI = .95, SRMR = .05. See Table 4 for items and factor loadings. A second CFA was performed on the third subsample and produced good model fit; RMSEA = .05 (CI [.00, .09]), CFI = .98, TLI = .97, SRMR = .03.

Bystander Action Responses—Other

Finally, the 14 items designed to describe how others may have reacted to a bystander's actions were factor analyzed using EFA and Varimax rotation. All but four items (1, 2, 5, 6) had skewness values below 2 and all but 5 (1, 2, 3, 5, 6) had frequencies above 20%. The result was a three-factor solution explaining 59.85% of the variance. The first factor explained 26.09% of the variance and included six items best described as "negative feedback" by others. The second factor included four items best described as "bystander apathy" and explained 19.85% of the variance. The third, describing positive feedback, included four items and explained 13.91% of the variance. All items had communalities between .4 and .7 and factor loadings above .4 and were all retained (see Table 5).

These three factors were used in the first CFA model, which did not have good fit. As a result, models were tested that used a higher cut-off, removing items with loadings less than .6 on the EFA and following model fit improvement indices from the MPlus analyses. The result was a set of three subscales each with three items that had adequate fit (RMSEA = .07, 90% CI [.05, .10], CFI = .95, TLI = .93, SRMR = .05). Table 5 lists items with

both EFA and CFA factor loadings. This revised model was tested with the third subsample and achieved good fit; RMSEA = .03 (CI [.00, .06]), CFI = .99, TLI = .99, SRMR = .04.

Hypothesis 2: Reliability and Preliminary Validity

We used the full study sample ($n = 966$) to compute Cronbach's α s for each scale, and these are presented in Tables 1 through 5. All but two subscales were greater than .65, indicating satisfactory reliability for most of the scales (DeVellis, 1991; Nunnally & Bernstein, 1994). We also used the full study sample to investigate the utility of the different consequence measures as separate scales by examining intercorrelations among the new measures; Table 6 presents these intercorrelations. They ranged from $-.39$ to $.77$; 85.45% of the correlations were less than .5, suggesting that the scales are measuring constructs that differed from one another (Mayer, 2016).

Table 6. Intercorrelations Between Bystander Consequences Subscales (N = 458–789)

Scale	1	2	3	4	5	6	7	8	9	10	11	12
1. Feelings–unsure	—	-.27***	.62***	-.35***	.39***	.11***	.17***	.34***	-.21***	-.02	.17***	.50***
2. Feelings–positive		—	-.14***	.38***	-.14***	.12***	-.01	-.05	.29***	.15***	.15***	-.13**
3. Feelings–negative			—	-.39***	.49***	.10**	.32***	.41***	-.11**	.09*	.34***	.77***
4. Vic react–positive				—	-.31***	.16***	-.01	-.09*	.32***	.15***	.01	-.32***
5. Vic react–negative					—	.26***	.24***	.45***	.01	.15***	.13**	.44***
6. Perp react–negative						—	-.07	.30***	.17***	.24***	.14**	.15***
7. Perp react–positive							—	.22***	.13**	.16***	.28***	.35***
8. Other react–negative								—	.02	.29***	.14**	.40***
9. Other react–positive									—	.17***	.24***	-.06
10. Other apathy										—	.07	.13***
11. General positive											—	.48***
12. General negative												—

Note. Vic = victim; Perp = perpetrator.

* $p < .05$. ** $p < .01$. *** $p < .001$.

In addition, using only the sample from the university psychology participant pool who completed the Social Desirability Scale, correlations were conducted between the new measures and the measure of social desirability. There were no significant correlations, which is preliminary evidence that the scales are not reflective of a response bias toward looking good (Warner, 2012).

Hypothesis 3: Gender Difference Analyses

Gender similarities and differences were explored using a series of analyses of variance, one for each of the 12 subscales of bystander consequences using the full sample. A dichotomous variable indicating male/female was used as the grouping variable. Table 7 presents overall means and means by gender, as well as univariate tests. Given the number of tests, a Bonferroni correction was used (Warner, 2012) and results were considered significant only at $p < .004$. Gender differences were found for seven scales; men reported significantly more positive and negative general outcomes, more unsure and negative feelings about their actions, more positive responses from the perpetrator, and more negative feedback from potential victims and others. There were no gender differences on positive feelings about action, apathy of other bystanders, negative feedback from the perpetrator, or positive feedback from others and victims.

Table 7. Means From One-Way Analyses of Variance of Sex Differences in Bystander Consequence Measures

Bystander outcome	Overall	Female	Male	<i>F</i>
General outcome positive ^a	3.60 (2.40)	3.30 (2.36)	4.00 (2.41)	15.48***
General outcome negative ^b	1.49 (2.19)	1.04 (1.87)	2.09 (2.43)	40.48***
Feelings—unsure ^c	2.00 (.96)	1.93 (.92)	2.11 (1.02)	7.30**
Feelings—positive ^d	3.25 (.75)	3.29 (.73)	3.19 (.77)	3.32
Feelings—negative ^e	1.53 (.81)	1.35 (.62)	1.84 (.99)	77.67***
Victim—positive ^f	0.67 (.37)	0.70 (.38)	0.63 (.36)	6.62
Victim—negative ^g	0.14 (.28)	0.11 (.27)	0.18 (.30)	11.72***
Perp—negative ^h	0.39 (.38)	0.37 (.38)	0.43 (.39)	4.15
Perp—positive ⁱ	0.18 (.31)	0.14 (.30)	0.24 (.33)	18.20***
Others—apathy ^j	0.30 (.39)	0.30 (.40)	0.31 (.37)	0.10
Others—positive ^k	0.51 (.42)	0.52 (.43)	0.50 (.41)	0.49
Others—negative ^l	0.14 (.27)	0.12 (.24)	0.18 (.29)	11.10***

Discussion

The current study introduces preliminary new measures of bystander consequences, the Compendium of Bystander Consequences. Much work has gone into understanding the process by which bystanders choose to intervene in a potentially threatening situation, but few studies have examined what happens afterward. The measures evaluated in this study were developed using a comprehensive process that began with pilot work identifying the kinds of consequences that people who engaged in bystander intervention have described. The item pool included a range of potential consequences, including the responses of the potential victim, potential perpetrator, and others present in the situation, as well as bystanders' feelings after intervening. We evaluated the factor structure, reliability, and conducted analyses to begin to explore the construct validity of the scales using a sample of 18–24-year-olds large enough to divide in thirds so that the factor analytic results could be examined in two replication samples.

Factor analyses of each scale produced factors representing positive and negative consequences, and in the cases of bystander feelings and responses from others, a third type of consequence. Specifically, some bystanders reported feeling unsure after intervening, and indicated that other bystanders seemed indifferent to

their intervention, expressing neither positive or negative reactions. Thus, the scales provide a tentative, though potentially meaningful, assessment of the different kinds of consequences that could occur following bystander intervention. They also offer the capacity to assess both the consequences that individuals may have experienced in multiple situations and to focus in on a specific situation.

The correlations of the general measure with the specific scales were modest with a few exceptions, suggesting that there is variability in the kinds of consequences that occur across situations. The correlations for negative consequences were higher, suggesting that the briefer measures of general negative consequences might be an adequate choice when researchers are seeking a general overview of negative perceptions of helping. For researchers interested in obtaining a fairly brief, global assessment of the consequences individuals have experienced following bystander intervention, the nine-item General scale may suffice (following further evidence of its validity). The general positive index did not seem to be a close approximation of incident-specific positive consequences and indeed was positively related to the general negative feelings scale. This may indicate that responding on this general measure may be reflecting a response set—where some participants are more sensitive to noticing higher levels of positive and negative reactions overall and others are less attuned to these. It may also be influenced by memory, where participants with better memories report more of both positive and negative responses. This finding also suggests that positive and negative consequences are not inversely related or mutually exclusive and bystanders may receive both in relation to any one instance of helping. Further, it should be noted that the general and specific measures are scored on different scales with the general measure indicating numeric frequency across a broader scale than the scale used for the incident-specific measures. Such differences may make these scales less comparable with one another.

Interestingly, and consistent with previous studies of bystander behavior, we found gender differences in consequences. However, these differences were in the opposite direction of what was hypothesized. Men tended to report greater perceptions of negative consequences from taking action than women. Given previous research that finds women more likely to take action (Banyard & Moynihan, 2011; McMahon, 2010), perhaps this is related to their perception and/or experience of fewer negative consequences when they do offer help. Further research is needed to better understand how gender influences the ways that people intervene and how those strategies may lead to different consequences, and to explore how the reactions of victims, perpetrators, and other bystanders may differ depending on the gender of the bystander. For example, it would be interesting to study the importance of particular consequences to men and women; do they perceive different salience or meaning for each if they experience them? Leone, Parrott, Swartout, and Tharp (2016) found that men's adherence to traditional norms of masculinity was related to bystander action: Men who believed men should be tough, for example, perceived greater barriers to action and expressed lower bystander confidence. Much more work is needed to understand how the complexity of gender identity impacts how individuals think about bystander action. What is more, gender nonconforming individuals may experience a unique set of bystander consequences that is important to understand.

Limitations

It is important to keep a number of limitations in mind when interpreting these findings. The current study was exploratory and cross-sectional and did not attempt to evaluate theory-driven hypotheses about how positive and negative consequences of bystander actions predict future bystander behavior. Further research is needed to examine such associations. The sample was not racially diverse, which prohibited analyses of the impact of ethnicity on bystander consequences. Similarly, the current research focused on young adults, and the results may not generalize to other age groups. Specifically, we know from recent work (Wee, Todd, Oshiro, Greene, & Frye, 2016) that bystander behavior looks different among adults in communities than among college students. The consequences of taking action likely differ as well. Although some consistency was found between EFA and CFA, one sample was divided in thirds for these analyses. A clear next step would be to further confirm the

factor structure and item loadings for each scale in a new sample, because issues specific to these samples might be influencing results. All the data were collected using self-reports from the same respondent, which may have increased the magnitude of the associations among them. Obtaining reports of the consequences of bystander intervention from other individuals would provide another perspective on what happens after bystanders act, although ultimately it may be the bystanders' perception of those reactions that is critical for shaping their future behavior. What is more, measures of negative reactions have many more items than those indicating positive reactions. It may be helpful to conduct further focus groups that concentrate more specifically on positive outcomes to generate more items to enhance these measures.

Research Implications

Although some consistency was found between EFA and CFA, a clear next step would be to further confirm the factor structure and item loadings for each scale in a new sample. Refinement of instructions may also be needed. For example, it is possible that more than one bystander was present and thus it is not clear on the current measure Bystander Consequences-Other whether a participant is describing the reaction of one bystander or a summary across several bystanders. Finally, more work needs to be done on constructs that showed low reliability (e.g., feeling unsure and reporting negative feedback from other bystanders, negative perpetrator reactions) and on establishing convergent and discriminant validity with other constructs, such as rape myth acceptance and intent to help. These next steps should be taken before these bystander consequence measures are widely used.

Once the psychometric properties of these new measures are better established, they have a number of possible uses. For example, they can be used to test theories on predicting when individuals do and do not intervene to prevent violence. The Action Coils model of bystander intervention predicts that past experiences with taking action, especially the consequences bystanders experience, will affect whether individuals are more or less likely to try to intervene in the future (Banyard, 2015). To test this model, it is important to have a clear measurement of the types of responses that bystanders experience and to continue research to understand exactly how these consequences impact future bystander behavior.

Prevention and Clinical Implications

Further, given the proliferation of bystander intervention training programs, we need a much better understanding of the assumed positive consequences of these actions and also of the potential negative effects stepping in to help might have on bystanders. This will enable programs to better design safety training strategies to help potential bystanders make better and safer choices and to measure possible iatrogenic effects of training. Such measures would also be a first step in measuring whether different types of bystander action are more likely to have more positive or more negative impacts and consequences. If such patterns could be established through research, this information would help preventionists focus skill building workshops on types of action that are most likely to be effective and to have the least negative impact on helpful bystanders. Many programs note the importance of safety, but there is very little empirical evidence about the scope of positive or negative effects of taking action for bystanders.

Footnotes

- 1 Terms used were defined as: Harassing comment (i.e., catcalling someone, sexist, racist, or homophobic joke, rape joke, blaming a victim of sexual assault or relationship abuse); Dating violence (i.e., any situation where you observed emotional, psychological, or physical abuse between two relationship partners, including use of physical force or threats of force against a partner such as slapping, punching, throwing objects, threatening with weapons or threatening any kind of physical harm; any situation where one or both relationship partners were arguing, physically fighting, name calling, or insulting, intimidating,

making fun of, or blaming one another); Unwanted sexual advances (i.e., any situation where you observed behaviors that are unwanted by the recipient and include remarks about physical appearance, persistent sexual advances that are undesired by the recipient, and/or unwanted touching and unwanted oral, anal, or vaginal penetration; any situation where one person make obscene gestures at or dances (ex., grinding) with another person who does not want to engage in those actions); Controlling behavior (i.e., any situation where you observed someone receiving incessant texting or phone calls, being asked where he or she is or what he or she is doing, and the victim's phone being monitored by another person and/or where you observed someone sending incessant texts or phone calls, demanding to know where another person is or what he or she is doing, and monitoring another person's phone).

References

- Banyard, V. L. (2015). Springer briefs in criminology. Toward the next generation of bystander prevention of sexual and relationship violence: Action coils to engage communities. New York, NY: Springer Science + Business Media. 10.1007/978-3-319-23171-6
- Banyard, V. L., & Moynihan, M. 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, 287–301. 10.1037/a0023544
- Banyard, V. L., Moynihan, M. M., Cares, A. C., & Warner, R. (2014). How do we know if it works? Measuring outcomes in bystander-focused abuse prevention on campuses. **Psychology of Violence**, 4, 101–115.
- Bennett, S., & Banyard, V. L. (2016). Do friends really help friends? The effect of relational factors and perceived severity on bystander perception of sexual violence. **Psychology of Violence**, 6, 64–72. 10.1037/a0037708
- Bennett, S., Banyard, V. L., & Garnhart, L. (2014). To act or not to act, that is the question? Barriers and facilitators of bystander intervention. *Journal of Interpersonal Violence*, 29, 476–496. 10.1177/0886260513505210
- Black, M. C., Basile, K. C., Breiding, M. J., Smith, S. G., Walters, M. L., Merrick, M. T., . . . Stevens, M. R. (2011). The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 Summary Report. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.
- Blake, B. F., Valdiserri, J., Neuendorf, K. A., & Nemeth, J. (2006). Validity of the SDS-17 measure of social desirability in the American context. *Personality and Individual Differences*, 40, 1625–1636. 10.1016/j.paid.2005.12.007
- Burn, S. M. (2009). A situational model of sexual assault prevention through bystander intervention. *Sex Roles*, 60, 779–792. 10.1007/s11199-008-9581-5
- Coker, A. L., Williams, C. M., Follingstad, D. R., & Jordan, C. E. (2011). Psychological, reproductive and maternal health, behavioral, and economic impact of intimate partner violence. In J. W. White, M. P. Koss, & A. E. Kazdin (Eds.) *Violence against women and children: Vol. 1. Mapping the terrain* (pp. 265–284). Washington, DC: American Psychological Association.
- Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, 18, 147–167. 10.1207/s15327906mbr1802_2
- DeVellis, R. F. (1991). *Scale development: Theory and applications*. London, United Kingdom: SAGE.
- Dovidio, J. F., Piliavin, J. A., Gaertner, S. L., Schroeder, D. A., & Clark, R. D., III. (1991). The Arousal: Cost-reward model and the process of intervention: A review of the evidence. In M. S. Clark (Ed.), *Prosocial behavior* (pp. 86–118). Newbury Park, CA: Sage.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286–299. 10.1037/1040-3590.7.3.286
- Geiser, C. (2013). *Data analysis with MPlus*. New York, NY: Guilford Press.
- George, D., & Mallery, M. (2010). *SPSS for windows step by step: A simple guide and reference, 17.0 update*. Boston, MA: Pearson.

- Gidycz, C. A., Orchowski, L. M., King, C. R., & Rich, C. L. (2008). Sexual victimization and health-risk behaviors: A prospective analysis of college women. *Journal of Interpersonal Violence*, 23, 744–763. 10.1177/0886260507313944
- Hamby, S. L., Weber, M. C., Grych, J., & Banyard, V. L. (2016). What difference do bystanders make? The association of bystander involvement with victim outcomes in a community sample. **Psychology of Violence**, 6, 91–102. 10.1037/a0039073
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424–453. 10.1037/1082-989X.3.4.424
- Krauss, A., Jouriles, E. N., Yule, K., Grych, J. H., Sargent, K. S., & Banyard, V. L. (2017). Adverse consequences to assisting victims of campus violence: Initial investigations among college students. *Journal of Interpersonal Violence*. Advance online publication. 10.1177/0886260517746944
- Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help*. New York, NY: Appleton-Century-Crofts.
- Leone, R. M., Parrott, D. J., Swartout, K. M., & Tharp, A. T. (2016). Masculinity and bystander attitudes: Moderating effects of masculine gender role stress. **Psychology of Violence**, 6, 82–90. 10.1037/a0038926
- Levine, M., & Crowther, S. (2008). The responsive bystander: How social group membership and group size can encourage as well as inhibit bystander intervention. *Journal of Personality and Social Psychology*, 95, 1429–1439. 10.1037/a0012634
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84–99. 10.1037/1082-989X.4.1.84
- Mayer, J. D. (2016). *The elements of mental tests* (2nd ed.). New York, NY: Momentum Press.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemical Medicine*, 22, 276–282. 10.11613/BM.2012.031
- McMahon, S. (2010). Rape myth beliefs and bystander attitudes among incoming college students. *Journal of American College Health*, 59, 3–11. 10.1080/07448481.2010.483715
- Moschella, E. A., & Banyard, V. L. (2018). Action and reaction: The impact of consequences of intervening in situations of interpersonal violence. *Journal of Interpersonal Violence*. Advance online publication. 10.1177/0886260518782983
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Mahwah, NJ: Erlbaum.
- Planty, M. (2002). *Third-party involvement in violent crime, 1993–1999*. Rockville, MD: U.S. Department of Justice, Office of Justice Program, Bureau of Justice Statistics.
- Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Assessment*, 17, 222–232. 10.1027//1015-5759.17.3.222
- Warner, R. M. (2012). *Applied statistics: From bivariate through multivariate techniques: From bivariate through multivariate techniques*. Atlanta, GA: Sage.
- Wee, S., Todd, M.-J., Oshiro, M., Greene, E., & Frye, V. (2016). Modifiers of neighbors' bystander intervention in intimate partner violence: A concept mapping study. *Violence and Gender*, 3, 55–63. 10.1089/vio.2015.0012
- Witte, T. H., Casper, D. M., Hackman, C. L., & Mulla, M. M. (2017). Bystander interventions for sexual assault and dating violence on college campuses: Are we putting bystanders in harm's way? *Journal of American College Health*, 65, 149–157. 10.1080/07448481.2016.1264407
- Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9, 79–94. 10.20982/tqmp.09.2.p079
- Yule, K., & Grych, J. (2017). College students' perceptions of barriers to bystander intervention. *Journal of Interpersonal Violence*. Advance online publications. 10.1177/0886260517706764