Usage guidelines

Evaluating a Bystander Intervention to Disrupt Interpersonal Violence/Abuse

SCHOLARONE™
Manuscripts

MANUSCRIPT DETAILS

TITLE: Evaluating a Bystander Intervention to Disrupt Interpersonal Violence/Abuse

ABSTRACT:

It is recommended that universities implement bystander interventions to disrupt the interpersonal violence and abuse that students experience in this context. Yet, there are few evaluations of bystander interventions in the UK. Building on an existing evaluation carried out on a bystander intervention at a university in 2017/18, the purpose of this research was to evaluate the intervention.

Using a one-way repeated measures design, ANOVA was used to analyse pre- and post-intervention data gathered from 121 students, during 2018/19.

As the aims of the session were met, it can be inferred individuals who participate in the bystander intervention have the potential to disrupt interpersonal violence and abuse.

The small sample size and design of the survey limited the research. Further evaluations of bystander interventions are needed in the UK that utilises large samples and a validated survey.

This paper notes the importance of engaging many students in a cohort to participate on a bystander intervention.

CUST_SOCIAL_IMPLICATIONS_(LIMIT_100_WORDS) :No data available.

This study adds to the paucity of evaluations of bystander interventions in the UK. Knowing that the intervention has the potential to disrupt interpersonal violence and abuse builds the momentum for other similarly designed interventions to be implemented in universities in the UK.

Evaluating a Bystander Intervention to ~~Tackle~~ Disrupt Interpersonal Violence/Abuse

Abstract

Purpose

It is recommended that universities implement bystander interventions to ~~tackle~~ disrupt the interpersonal violence and abuse that students experience in this context. Yet, there are few evaluations of bystander interventions in the UK ~~to assess whether they work~~. Building on an existing evaluation carried out on a bystander intervention at a university in 2017/18, the purpose of this research was to evaluate the intervention.

Design/methodology/approach

Using a one-way repeated measures design, ANOVA was used to analyse pre- and post-intervention data gathered from 121 students ~~participating in the intervention,~~ during 2018/19.

Findings

As the aims of the session were met, it can be inferred ~~that participants who~~ ~~engage~~ individuals who participate in the bystander intervention have the potential to ~~tackle~~ disrupt interpersonal violence and abuse.

Research limitations/implications

The small sample size and design of the survey limited the research. Further evaluations of bystander interventions are needed in the UK that utilises large samples and a validated survey. ~~for bystander evaluations.~~

Practical implications

This paper notes the importance of engaging ~~all~~ many students in a cohort to participate on a bystander intervention. ~~Further research is needed to understand the difficulties of achieving this.~~

Originality/value

This study adds to the paucity of evaluations of bystander interventions in the UK. Knowing that the intervention has the potential to ~~tackle~~ disrupt interpersonal violence and abuse builds the momentum for other similarly designed interventions to be implemented in universities in the UK.

Key words

Interpersonal violence/abuse, universities, bystander interventions, evidence-base, statistical evaluations, recommendations

Paper type

Research paper

Introduction

~~Students~~ Female students at ~~universities in the~~ UK universities experience interpersonal violence and abuse, particularly sexual violence (NUS, 2011~~, 2018~~; Phipps and Young, 2015; Roberts *et al*., 2019; Stenning *et al*., 2013). Research with 2,058 female students identified that 68% had experienced verbal and physical sexual harassment, e.g., flashing and groping, and almost one in four had experienced unwanted sexual contact within and outside university. Most respondents reported that they knew the perpetrators (NUS, 2011). Although, this is not the hidden sexual violence that exists within domestic violence relationships, but is as Roberts *et al.* (2019) argue, the perceived hetero-normative practices of pre-dating and courtship. For example, in Phipps and Young's (2015) research with 40 women, they found that over two-

thirds discussed sexual harassment and sexual physical assault as normative of university life. Consequently, Universities UK (2016) recommend that universities implement evidence-based bystander interventions.

Key components of bystander interventions educate participants about interpersonal violence and abusive behaviours, including the confidence and knowledge to intervene safely to disrupt behaviours such as sexual violence (Fenton *et al*., 2015; Labhardt *et al.,* 2017). The aim is for students to develop a culture of intolerance to interpersonal violence and abuse, this may benefit wider society due to the large intake of students progressing to higher education (Universities UK, 2016). Most evaluations of bystander interventions have been carried out in the United States (US) (Fenton and Mott, 2017; Labhardt *et al*., 2017). Very few evaluations have been carried out in the UK (Fenton *et al*., 2015). More research is needed in this context; it is not clear how applicable the US research is to British context (Fenton and Mott, 2017; Labhardt *et al*., 2017).

This paper begins by outlining the bystander effect and how to overcome it. We then review the recent evaluations of bystander interventions carried out in the US and the UK. We also review research on students' experiences of interpersonal violence and abuse carried out at a university in the north of England, as well as an evaluation of a bystander intervention at that university. Building on an this existing evaluation of a bystander intervention at the same university, we discuss our methodology to evaluate the intervention. provide a statistical evaluation to assess whether the intervention meets its aims. After this, we present our findings and a discussion of these, concluding with the limitations of the study and recommendations for future research.

The Bystander Effect

The story about The bystander effect originates from usually begins with Kitty Genovese: who was a victim of murder and sexual assault in New York in 1964. The story that was presented in the Media and academia reports suggested that there were 38 bystanders who witnessed the

attack and did not act. ~~did nothing about it~~. This led to the notion of the bystander effect: that individual responsibility is diffused when in the presence of others (Manning *et al*., 2007). Individuals feel~~s~~ less responsible when in ~~groups or~~ groups/crowds because they assume others will or have acted; or, if no one is acting, they assume that the situation is not an emergency. Individuals thus look to others in a group or crowd setting to ascertain ~~a~~ course of action (Garcia *et al*., 2002). ~~In this sense then,~~ Thus, urban crowds/groups ~~and groups of people~~ are powerful and 'a dangerous threat to social stability' (Manning *et al*., 2007, p. 560). ~~As a collective, they~~ Crowds/groups thereby promote passiveness and inactivity on individuals, meaning~~. In doing so,~~ social values are either undermined and/or upheld (ibid).

The case of Genovese addresses this, the sexual and physical attack of a woman in public undermined the value of social responsibility for others (ibid). ~~In the case of Kitty Genovese, the story of the 38 inactive bystanders upheld the view that the sexual and physical attack of a woman in public is no-one's business and undermined the view that it was everyone's business (ibid).~~ This perpetuates attitudes towards victim blaming: 'larger community and societal issues such as gender inequality, along with male social control and entitlement, permeate the foundation of attitudes that condone violence against women' (Banyard *et al.,* 2004, p. 63). If ~~individuals in a group or crowd setting~~ victims are viewed as the cause of violence, ~~view the victim as causing the violence~~ surrounding bystanders are less likely to feel responsible for the violence themselves, and thus, they are less likely to intervene (ibid).

Despite Manning *et al*. (2007) arguing that the ~~the story of Kitty~~ Genovese case has been frequently misrepresented, there has been much research since then to support the bystander effect (see Fischer *et al*., 2011; Garcia *et al*., 2002). A review conducted by Banyard *et al*. (2004) ~~review of the research, they argue that~~ identified that bystanders are more likely to intervene if they are: aware of the problem, ~~including the negative impact on the victim;~~ view victims as not causing their victimisation; view themselves as the solution to the problem ~~responsible in some way for solving the problem; view others who intervene;~~ and feel they ~~have the skills to~~ believe they are equipped with the skills to intervene~~; and are asked to help~~. ~~Fenton and Mott (2017) argue that~~ Arguably, it is important that bystanders actively intervene because it sends a message to the perpetrator and other bystanders that the behaviour observed

is unacceptable (Fenton & Mott, 2017), creating a culture of intolerance. They also argue that Bystander intervention also sends a message about the acceptability of challenging inappropriate behaviours (Fenton and Mott, 2017). Furthermore, from a victim's perspective, bystander intervention is often 'associated with feeling safe and with decreased harm' (Fileborn, 2017, p. 200). The next section reviews whether bystander interventions work.

Bystander Evaluations in the US

Bystander interventions have developed as a strategy that situates the responsibility of preventing interpersonal violence and abuse with the community (Jewkes *et al*., 2014 cited in Fenton *et al*., 2015, p.12). Most of the research research underpinning the effectiveness of bystander interventions has been carried out in the is conducted in the US (Fenton and Mott, 2017; Labhardt *et al*., 2017). Bystander interventions in this context have predominantly focused on addressing sexual and dating violences at on college/university. campuses. The findings from the evaluations suggest that such interventions have the potential to enhance the likelihood of bystanders intervening and thereby reducing violence (see Jouriles *et al*., 2018; Katz and Moore, 2013). The effects of bystander interventions have been shown to last over time: at least 3 months after the intervention (Jouriles *et al*., 2018; Senn and Forrest, 2016).

Coker *et al*. (2016) for example, evaluated the *Green Dot* bystander intervention over a 4-year period on one college campus in Kentucky, against two other comparing the effects against two other campuses that which did not have the had no intervention. First year students Students (intervention n=2,979 and comparison campuses n=4,132) were randomly selected forto the study. As such, the evaluation was experimental in its design. The intervention sought aimed 'to empower potential bystanders to actively engage their peers' (ibid, p. 296) in order to change social values and reduce the likelihood of violence. The researchers therefore hypothesised that violence would be reduced on the campus as a result of the intervention. The researchers used a survey to measure Violent victimisation rates of dating and sexual violence were measured across the campuses and over 4 years and data was analysed using log-binomial regression. Victimisation rates for sexual and dating violence were significantly lower on the campus that had the bystander intervention compared to the campuses that did not have the intervention, excluding 2013. To account for this finding, Coker *et al*. (2016) argue that the

reduced sample size may have contributed to a TYPE II error by reducing the power of the tests to find differences in the data.

Bystander Interventions in the UK

Until recently, there were no published studies providing an evidence-base for ~~using~~ bystander interventions in ~~universities in the~~ UK universities (Fenton *et al*., 2015). ~~This was largely due to the recent addition of such interventions in this context (Fenton *et al*., 2015).~~

Research has identified that students hold myths about ~~Fenton and Jones (2017) found that myths about~~ sexual and domestic violence (Fenton and Jones, 2017) ~~were held by new students.~~ Such myths served to 'scaffold gender-based violence in university settings' (ibid, p. 147). As such, *The Intervention Initiative* (TII) was designed to disrupt such myths and thus the potential for violence (Fenton and Mott, 2018b)~~. Fenton and Mott (2018a) were the first to evaluate a bystander intervention, the TII, on university students in the UK.~~ To evaluate the TII, Fenton and Mott (2018a) adopted a within subjects repeated measure design, administering a questionnaire before and after 4 two-hour sessions. ~~, and used paired t-tests for 131 matched cases on the measures at pre and post-test~~intervention to analyse the data (after 4 two-hour sessions). ~~Thus, a within subjects repeated measures design was used. They found positive results after~~Following the intervention, ~~with law students~~ the authors found~~:~~ decreases of rape and domestic violence myth acceptance and denial; and increases in bystander efficacy, readiness and intent to help, and responsibility scores ~~(Fenton & Mott, 2018a)~~ ~~all key factors which are well founded in the American literature to test when evaluating bystander interventions (Labhardt *et al*., 2017).~~ ~~However,~~ Although, there was no effect of the intervention on students engaging in bystander behaviours.~~, albeit the researchers note that~~ To account for this finding, the authors argued that participants may not have had sufficient time at the point they completed the intervention (1 – 8 weeks following) to apply their learning into practice.

There is ~~certainly~~ a need for bystander interventions in the UK given the extent of interpersonal violence and abuse experienced by students at universities in the UK (NUS, 2011~~, 2018~~; Phipps and Young, 2015; Roberts *et al*., 2019; Stenning *et al*., 2013). In Roberts *et al's.* (2019) research, using an online university-wide survey in 2016, students were asked about their experiences of interpersonal violence and abuse whilst at university. Women were more likely to report experiencing verbal sexual harassment and sexual physical assault than men. ~~Roberts et al. (2019), using a university wide online survey, found that women reported experiencing sexual violence more than men.~~ Such sexual violence, the perceived hetero-normative practices of pre-dating and courtship, happened both on and off campus in public spaces, with the sexual physical assault particularly occurring ~~in the night-time economy, i.e.,~~ in the pubs and clubs. An analysis of the qualitative responses to the survey found that women reported adopting a range of ~~measures~~ strategies to resist the sexual violence as well as others intervening. For example, a bystander witnessing sexual harassment of 'a drunk girl' reported to the survey how she actively intervened to remove her from the perpetrators (Roberts *et al*., 2019, p. 332).

The authors identified that women rarely reported their experiences of sexual violence due to their perceptions that such sexually violent behaviours against women are normal and are accepted at university ~~behaviours, in at a university~~ context (Roberts *et al*., 2019; see also NUS, 2011~~, 2018~~; Phipps and Young, 2015), and in ~~the night-time economy~~pubs and clubs (Roberts *et al*., 2019; see also Kavanaugh, 2013). The authors further identified that most of the interpersonal violence and abuse occurred in public spaces. This is important to note because there is some evidence of the bystander effect, where individuals do not intervene because they feel less responsible for the interpersonal violence and abuse when in groups and crowd settings. There were fewer reports of bystanders, other than friends, intervening. Students also reported that the sexual violence had affected their mental health and academic study (Roberts *et al*., 2019).

Bystander interventions are arguably of great importance when applied to the issues referred to above by enhancing the likelihood of bystanders actively intervening, encourage potential victims to feel safe and reducing harm caused (Fileborn, 2017). The next section discusses our methodology to evaluate a bystander intervention. ~~This is particularly so given the reported~~

impact of sexual violence where students said it had affected their mental health and academic study (Roberts *et al*., 2019).

Given these findings, a 90 minute taster session was designed to enhance bystander intervention. It was piloted and evaluated during 2017/18 (Donovan and Corr, 2018), at the university where the online survey (discussed above) was carried out (see Roberts *et al*., 2019). The intervention used students' reports (anonymously) from the survey to illustrate students' experiences and impact of verbal abuse, including sexual verbal abuse, racism, body shaming, homo-bi-transphobia, disablism, islamophobia. Consequently, the bystander session at the university was broader than the other bystander interventions discussed in the existing evaluations that covered sexual and domestic/dating violence, albeit the underlying model of change is the same. Findings from the pilot evaluation suggested the taster session had a positive impact (Donovan and Corr, 2018). As such, the bystander session was delivered at the university in 2018/19, which we evaluated. The next section discusses our methodology.

Methodology

The Bystander Session

A a 90 minute taster session[1] had been designed. The session teaches participants about a broad range of interpersonal violence and abusive behaviours, the impact of these, and how to safely intervene to disrupt such behaviours. For example, facilitators delivering the session use students' anonymous reports to illustrate student experience and impact of verbal abuse, including sexual verbal abuse, racism, homophobia. The current bystander session was broader in its education of examples of verbal abuse than other bystander interventions reviewed in the existing evaluations that covered primarily sexual and dating/domestic violence, albeit the underlying model of change is the same (see Donovan and Corr, 2018).

Findings from the evaluation of the pilot in 2017/18 suggested the taster session had a positive impact (Donovan and Corr, 2018). Consequently, the bystander session was delivered at the university in 2018/19, which we evaluated to ascertain whether it met its four aims. The 90

~~minute bystander session had four aims.~~ The first two aims of: i) 'increase knowledge about and confidence to be active bystanders'; and ii) 'increase knowledge and awareness about harms/impacts of interpersonal violence/abuse' were informed by TII. The session aimed to provide practical information about being an active bystander in a range of contexts and to enhance empathy about victims (Donovan and Corr, 2018, unpaginated; see Fenton and Mott, 2018a, 2018b). If these aims are met, it is thought that the intervention will enhance individual responsibility for the interpersonal violence and abuse observed, and thus increase the likelihood of bystander intervention (see Fenton *et al*., 2015; Labhardt *et al*., 2017). The last two aims of: iii) 'raise knowledge and awareness about services provided by the University'; and iv) 'encourage reporting/help-seeking regardless of where incident(s) take place' (Donovan and Corr, 2018, unpaginated) were informed by the findings from the ~~university-wide online survey, discussed above,~~ university-wide online survey that identified ~~carried out at the university with students~~ (see Roberts *et al*., 2019) non-reporting of incidents of interpersonal violence and abuse to the university, which happened off campus (Donovan and Corr, 2018~~, unpaginated~~).

The Survey and the Sample

We used an adapted version of the survey from the previous evaluation (see Donovan and Corr, 2018). Some questions were added/refined to ensure testing of the programme aims. We also changed the response format of the questions to scales, for example, 'from 1 not at all to 10 yes definitely', to ensure parametric tests, which are more powerful in finding differences, could be used (see Carifo and Perla, 2007, 2008). The survey was piloted on social science undergraduate students to ensure accuracy in the design and delivery. The study was approved by the university Ethics Research Group. Students were not randomly allocated to the session nor did they self-select to take part in the session, which may avoid introducing bias in such non-probability samples (de Vaus, 2014). The session is often scheduled into students' timetables, as part of their curriculum, because of the synergy between their programmes of study and the bystander session. Before the start of the bystander session, participants were asked to read a study information sheet, confirm that they understood, and that they consented to take part. There were no incentives offered to students to take part in the study.

Sessions ran in late 2018 and in early 2019. The participants completed the survey, mostly electronically in Survey Monkey, but some hard copies were completed, ~~then administered~~ before and immediately after the 90 minute session. The aim was to measure differences over time, i.e., before and after the session, to ascertain if the aims of the session were met. ~~(hard copy or electronic in Survey Monkey), to measure differences over time, and thus, the impact of the session in meeting its 4 aims. Students were not randomly allocated to the session. The session is often scheduled into students' timetables because of the synergy between their programmes of study and the bystander session, and/or, it is important to students' future careers that they undertake the bystander session. The design of the evaluation was thus a within subjects, one-way repeated measures design.~~

Data Analysis

Data ~~was~~ were cleaned and entered into a master SPSS dataset. During 2018/19, 212 new students to the university, participated on the 90 minute bystander session.[2] Of these students, 152, most of whom were at level 3 study,[3] completed the pre-test survey. However, not all of these students completed the post-test survey immediately after the session[4]. Given this and the deletion of cases that were not useable[5], 121 matched cases at pre- and post-intervention were used for testing whether aims i and iv of the session had been met. There were 116 cases paired at pre- and post-intervention to test whether aim ii had been met. This is because there were 5 missing responses in total, not at random (see Kang, 2013), on both scale items[6] of victim-blaming, which were used to test aim ii (see findings section). There were 120 matched cases at pre- and post-intervention testing aim iii because there was 1 missing random response on the scale item testing aim iii. Table I shows the four aims of the intervention, the participants tested at each, and the missing responses.

'Table I' here

Paired t-tests were unsuitable to test the data due to violations of the assumption of normality found in the data. As one-way within subjects ANOVA is closely related to the paired t-test, and it is more robust to violations of the normality assumption, particularly with large samples >30 (Reid, 2014), we present the findings of one-way repeated measures ANOVA to test whether the programme met its aims. It is hypothesised that if there are statistically significant positive shifts in the mean scores on the scale/s for each of the aims of the programme, from

before the intervention to after the intervention, then the session will have met its aims. Time is the independent variable and the scale scores are the dependent variables. The design of the evaluation was thus a within subjects, one-way repeated measures design.

Findings

The Sample and the Session

The sample consisted of 45% (n=55) male and 55% (n=66) female students. The modal age of the sample was 20, the mean age was 25, and 67% (n=81) of the sample were aged between 18 and 24. This is a usual distribution of age of students at universities (see Universities UK, 2016) and an age group most at risk of interpersonal violence and abuse (see the *Crime Surveys for England and Wales*). As most of the pre- and post-test surveys were completed electronically, we were able to work out the length of the 90 minute bystander session as delivered in practice. We found that the modal time was 63 minutes, the mean and median time were 74 minutes (based on 88 cases). This excludes time to complete the surveys. For some sessions then, the delivery of the bystander session fell short of its prescribed time. Qualitative evaluation of the bystander session, including observations of the session, have been carried out during 2019/20, to assess this finding further. Moving on, frequency data shows that after the session, 98% (n=118) of students rated the quality of the bystander session on a scale from 1 very weak to 10 excellent, 5 and above. and Seventy nine per cent 79% (n=96) of students rated the quality of the bystander session 8 and above (where 1 is very weak and 10 is excellent). Students after the session selected one of the following responses: they already considered themselves to be an active bystander (46%, n=54); the session had given them the confidence to be an active bystander (48%, n=57); I feel apprehensive/not confident to be an active bystander (6%, n=7) (based on 118 cases[7]). It might be argued that almost half of students do not need the session to be an active bystander, however, shifts in the mean scores of the scales before and after the bystander session indicate that the intervention met its aims, as the next section tabulatesshows.

Evaluating the Session

As the survey was kept short to maximise the response rate (see for example Fenton and Mott, 2018a), there were 10 scale items testing the 4 aims of the session. This meant that aims iii and iv were measured using 1 scale item each. Aims i and ii were each tested using 4 scale items. We felt that *knowledge* and *confidence* were two different constructs to test aim i: increase knowledge about and confidence to be active bystanders. Knowledge was therefore tested with 1 scale item: on a scale from 1 not at all to 10 yes definitely, do you know what an active bystander is? We found a significant effect of time on knowledge score, $F(1, 120) = 536.63$, p $<.001$, $n^2=.82$. There was a significant increase in knowledge score from before (mean = 3.48) and after (9.66) the intervention. There was a large effect size. Table II shows this.

'Table II' here

Confidence was tested with 3 scale items: on a scale from 1 not confident to 10 very confident, how confident would you feel intervening if you saw someone being physically assaulted; how confident would you feel intervening if you saw someone being verbally abused; how confident would you feel intervening if you saw one person being bullied by a group? We found a significant effect of time on confidence scores, $F(1, 120) = 97.86$, p$<.001$, $n^2=.45$. There was a significant increase in confidence scores from before (mean = 19.02) and after (mean = 23.75) the intervention. There was a large effect size. Table III shows this.

'Table III' here

To test aim ii of the session: increase knowledge and awareness about harms/impacts of interpersonal violence/abuse we used four scale items. First, we asked on a scale from 1 not at all to 10 yes definitely, do you understand how people feel who are victimised by violence/abuse/bullying/hate comprising one scale item about victim-empathy. We found a significant effect of time on victim-empathy score, $F(1, 120) = 13.41$, p$<.001$, $n^2=.10$. There was a significant increase in victim-empathy score from before (mean = 8.23) and after (mean = 8.94) the intervention. There was a medium effect size. Table IV shows this.

'Table IV' here

Second, we asked on a scale from 1 not at all to 10 yes definitely, do you think that girls that get drunk are asking for trouble and do you think people who are offended by banter are too sensitive, comprising two scale items about victim-blaming.[8] We found a significant effect of time on victim-blaming scores, $F(1, 115) = 7.26$, p$<.008$, $n^2=.06$. There was a significant

decrease in victim-blaming scores from before (mean = 5.87) and after (mean = 5.07) the intervention. There was a medium effect size. ~~Table V show this.~~

~~'Table V' here~~

Third, we asked on a scale from 1 not at all to 10 yes definitely, do you think it is important to promote attitudes of respect, tolerance and care to challenge bullying (including body shaming), hate acts (e.g., racism, homo-bi-transphobia, disablism, Islamaphobia) and sexual harassment, which comprised one scale item about inclusive attitudes. Results are not significant for the inclusive attitudes scale, $n^2 \underline{< .01} = \underline{.00}$. Students embraced inclusive attitudes before the bystander session, as 70% (n=85) scored 10 on the scale at that time, rising to 79% (n=96) after the session.

To test aim iii: raise knowledge and awareness about services provided by the University, we asked on a scale from 1 not at all to 10 yes definitely, if something happened to you/a student you know that impacted upon your/their ability to study, would you know where to go/send them in the university for help? We found a significant effect of time on raising awareness about services score, $F(1, 119) = 188.72$, $p<.001$, $n^2=.61$. There was a significant increase in raising awareness about services score from before (mean = 5.34) and after (mean = 9.07) the intervention. There was a large effect. ~~Table VI shows this.~~

~~'Table VI' here~~

To test aim iv: encourage reporting/help-seeking regardless of where incident(s) take place, we asked on a scale from 1 not at all to 10 yes definitely, if something happened to you or a student you know, regardless of where the incident took place, would you report it to the university for help? We found a significant effect of time on reporting score, $F(1, 120) = 54.63$, $p<.001$, $n^2=.31$. There was a significant increase in reporting score from before (mean = 5.99) and after (mean = 7.97) the intervention. There was a large effect. ~~Table VII shows this.~~

~~'Table VII' here~~

Discussion

There were statistically significant positive shifts in the mean scores on the scale/s from before the intervention to after the intervention for each of the aims of the programme. Although, the difference of the scores on the one scale item about inclusive attitudes, which part measured aim ii, from before the session to after the session, was not statistically significant, the frequency data indicates that participants held fairly inclusive attitudes at the start of the session. Moreover, the other three scale items, which measured aim ii were statistically significant. Like Coker et al. (2016), we may have suffered a TYPE II error with our small sample size and inability to find a difference in the data. That said, our evaluation was a repeated measures design where the same participants completed both pre- and post-test surveys. This reduces the error in the data from the 'pre-existing subject differences', as each participant is matched to themselves, they thus serve as their own controls (Reid, 2014, p. 325). Consequently, we are able to make powerful inferences from the data. This is particularly salient given students did not 'self-select' to participate in the session (see also Fenton and Mott, 2018a, p. 655) potentially avoiding bias in the sample (de Vaus, 2014). Moreover, frequency data indicated that most participants rated the quality of the bystander session as veering towards excellent.

As the scale items measuring such changes in knowledge and skills were mostly all statistically significant, there is potential, albeit theoretically, for those participants who participated in the session and the evaluation to become active bystanders. Given the extent of interpersonal violence and abuse that university students experience, it is likely that these participants will observe interpersonal violence and abusive behaviours in the future. As such, they have the potential to challenge the passiveness and inactivity of individuals in a group or crowd setting in such situations, whether that is in the moment or afterwards e.g., by reporting it. In doing so, they have the potential to reduce the bystander effect, uphold important social values about collective responsibility (Manning et al., 2007) and acceptability of challenging inappropriate beliefs and interpersonal violence and abusive behaviours (Fenton and Mott, 2017).

Given the bystander session evaluated in this paper is much broader in educating participants about a wide range of verbal abuse than existing bystander interventions reviewed in the literature, which primarily focus on sexual and dating/domestic violence, students who

participated in the session are likely to have a broader understanding of what constitutes interpersonal violence and abusive behaviours. There is a theoretical argument that supports participants becoming active bystanders after participating in the 90 minute taster session, but there is no empirical support from this research that participants have become active bystanders after participating in the taster session (see also Fenton and Mott, 2018a).

Limitations

There were a number of several limitations. to the research. As we were treating a Likert scale as scale data, analysing single items (i.e., the dependent variables) on an 'item-by-time' basis is limiting (Carifio and Perla, 2007, p. 106). If there are sufficient scale points, for example 7, analysing items on a single basis is not so limiting (Carifo and Perla, 2007). Our scales had 10 points. Yet, as the survey was not validated, questions might be raised over whether we were sufficiently measuring the aims of the session. Furthermore, our small sample size as a result of attrition, meant that we could not measure the longer-term impact of the session. Despite the advantages of using a repeated measures design they are notable for loosing sample numbers over time (Fenton and Mott, 2018a; Knapp, 2018; Reid, 2014). Small sample sizes also reduce the power of ANOVA to find significant differences in the data. Similar to other UK bystander evaluation (see Fenton and Mott, 2018a), we could not generalise the findings to a wider sample of students due to non-random sampling. (see also Fenton and Mott, 2018a).

Recommended Changes

A number of recommendations arise from the evaluation. for research methodologies, bystander interventions, and universities. Firstly, in terms of evaluations, if they are to measure the longer-term impact of an intervention upon participants, they must begin with a large sample high numbers of participants them to allow for their attrition over time (Knapp, 2018). Carrying out a prospective power analysis to indicate the appropriate sample size is fundamental This will to ensure the power of the tests, such as the tests, such as ANOVA, are powerful to find differences in the data. In doing so, are commensurate with appropriate sample sizes and thereby avoiding a Type II error is avoided (see Clark-Carter, 1997).

Moreover, surveys need to be validated for their use in evaluations of bystander interventions (see also Fenton and Mott, 2018a). Using Cronbach's alpha to test the internal consistency of the scales prior to analysis is paramount (Allen and Seaman, 2007 cited in Leung, 2011, p. 413). A pilot of the survey is therefore advised. Scales on the survey should start with 0 to indicate complete absence of the attribute, which is being measured (Leung, 2011). If we had done this, we could have reverse coded our data, where needed.

While our statistical evaluation can tell us that the session works in meetingmet its aims, we do not know if the positive results lead to a reduction in the bystander effect when participants are in a group or crowd setting observing interpersonal violence and abuse, and thus, we do not know if they become active bystanders. Nor does our data it does not tell us inform how and why it the session met its aimsworks. Therefore, a recommendation for future research is to assess the long term behaviour change of participants (see Banyard *et al*., 2004, 2005; Fenton and Mott, 2018a), including why they do or do not intervene when observing interpersonal violence and abuse (Banyard *et al*., 2005; Fileborn, 2017; Labhardt *et al*., 2017). The surveys used in our statistical evaluation gathered qualitative data about how and why the session works (which is the focus of another paper), and students were invited for interview to discuss this, but take up the response was low. The data gathered from these sources will add to the qualitative evaluation of the bystander intervention during 2019/20, which we are carrying out to address these gaps in the research.

In terms of the content of the bystander session, it is advised that one of the aims of the session explicitly states 'to increase individual responsibility in a group/crowd setting' because this is the theoretical premise for overcoming the bystander effect. Whilst this is implicit in aims ii and iv of the current session, making it explicit ensures it is central, and the design of the session can be adapted accordingly. This is important because much of the interpersonal violence and abusive behaviours against students found at the university occurred in public spaces amongst crowds and groups and it was also normalised (Roberts *et al*., 2019). If bystanders are less likely to view the behaviours as a problem (Fileborn, 2017), and/or, they blame the victim for causing their own victimisation, they will not consider themselves as

individually responsible for the violence and abuse, and thus, they are less likely to intervene (Banyard *et al*., 2004). Lastly, the bystander session is short. Single sessions are not thought to be effective at changing behaviours long term (DeGue, 2014 cited in Fenton and Mott, 2018b, p. 182).[9] Other bystander interventions have longer and multiple sessions (see Fenton and Mott, 2018a).

In summary, despite the limitations, a specific recommendation for universities would be to implement the bystander session to all students in a year group (University of Exeter, 2020) or at least most students in a year group. As Coker *et al*., (2016, p. 301) claimed that the *Green Dot* bystander intervention had the potential to impact upon students who had not participated on the intervention by those who had participated on the intervention 'modelling bystander behaviours'. This creates a culture of intolerance to violent and abusive behaviours, which needs to be driven by cohorts so they all act together to challenge the beliefs that uphold such behaviours because all members of a community are responsible for acting: not just the perpetrators and victims (Banyard *et al*., 2004). Given the small numbers of new students to the university who participated in the bystander session in 2018/19, further research is needed to understand how to extend the delivery of the session to more students in a year group. particularly noting the difficulties of harnessing staff and students in to engage in the sessionachieving this.

References

Banyard, V.L., Plante, E.G. and Moynihan, M.M. (2004) "Bystander Education: bringing a broader community perspective to sexual violence prevention"*, Journal of Community Psychology,* Vol. 32 No. 1, pp. 61-79.

Banyard, V.L., Plante, E.G. and Moynihan, M.M. (2005) *Rape Prevention Through Bystander Education: bringing a broader community perspective to sexual violence prevention,* available at: https://www.ncjrs.gov/pdffiles1/nij/grants/208701.pdf (accessed 13 February 2020).

Carifio, J. and Perla, R. (2007) "Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes", *Journal of Social Sciences,* Vol. 3 No. 3, pp.106-116.

Carifio, J. and Perla, R. (2008) "Resolving the 50-Year Debate Around Using and Misusing Likert Scales", *Medical Education,* Vol. 42, pp. 1152-1154.

Clark-Carter, D. (1997) "The Account taken of Statistical Power in Research Published in the *British Journal of Psychology*", *British Journal of Psychology,* Vol. 88, pp. 71-83.

Coker, A.L., Bush, H.M., Fisher, B.S., Swan, S.C., Williams, C.M., Clear, E.R. and DeGue, S. (2016) "Multi-College Bystander Intervention Evaluation for Violence Prevention", *American Journal of Preventive Medicine,* Vol. 50 No. 3, pp. 295-302.

de Vaus, D. (2014) *Surveys in Social Research,* 6th ed.*,* Routledge, London.

Donovan, C. and Corr, E. (2018) *Be the Difference: an evaluation of a 90-minute Active Bystander Pilot,* unpublished.

Fenton, R.A., Mott, H.L., McCartan, K. and Rumney, P.N.S. (2015) *A Review of Evidence for Bystander Intervention to Prevent Sexual and Domestic Violence in Universities,* available at: https://www2.uwe.ac.uk/faculties/BBS/BUS/law/Law%20docs/dvlitreviewproof0.6.forCLR.pdf (accessed 5 February 2020).

Fenton, R. and Jones, C. (2017) "An Exploratory Study on the Beliefs about Gender-Based Violence held by Incoming Undergraduates in England", *Journal of Gender-Based Violence,* Vol. 1 No. 2, pp. 147-67.

Fenton, R.A. and Mott, H.L. (2017) "The Bystander Approach to Violence Prevention: considerations for implementation in Europe", *Psychology of Violence,* Vol. 7 No. 3, pp. 450-458.

Fenton, R.A. and Mott, H.L. (2018a) "Evaluation of the Intervention Initiative: a bystander intervention program to prevent violence against women in universities", *Violence and Victims,* Vol. 33 No. 4, pp. 645-662.

Fenton, R.A. and Mott, H.L. (2018b) "The Intervention Initiative: theoretical underpinnings, development and implementation", in Anitha, S. and Lewis. R. (Ed.s.) *Gender Based Violence*

in University Communities: policy, prevention and educational initiatives, Policy Press, Bristol, pp, 169-188.

Fileborn, B. (2017) "Bystander Intervention from the Victims' Perspective: experiences, impacts and justice needs of street harassment victims"**,** *Journal of Gender-Based Violence,* Vol. 1 No. 2, pp. 187-204.

Fischer, P., Krueger, J.I., Greitemeyer, T., Vogrincic, C., Kastenmuller, A., Frey, D., Heene, M., Wicher, M. and Kainbacher, M. (2011) "The Bystander-Effect: a meta-analytic review of bystander intervention in dangerous and non-dangerous emergencies", *Psychological Bulletin,* Vol. 137 No. 4, pp. 517-537.

Garcia, S.M., Weaver, K., Moskowitz, G.B. and Darley, J.M. (2002) "Crowded Minds: the implicit bystander effect", *Journal of Personality and Social Psychology,* Vol. 83 No. 4, pp. 843-853.

Jouriles, E.N., Krauss, A., Vu, N.L., Banyard, V.L. and McDonald, R. (2018) "Bystander programs addressing sexual violence on college campuses: a systematic review and meta-analysis of program outcomes and delivery methods", *Journal of American College Health,* Vol. 66 No. 6, pp. 457-466.

Kang, H. (2013) "The Prevention and Handling of the Missing Data", *Korean Journal of Anesthesiology,* Vol. 64 No. 5, pp. 402-406.

Katz, J. and Moore, J. (2013) "Bystander education training for campus sexual assault prevention: an initial meta-analysis", *Violence and Victims,* Vol. 28 No. 6, pp. 1054-1067.

Kavanaugh, P.R. (2013) "The Continuum of Sexual Violence: women's accounts of victimisation in urban nightlife", *Feminist Criminology,* Vol. 8 No. 1, pp. 20-39.

Knapp, H. (2018) *Intermediate Statistics Using SPSS, Sage: London.*

Labhardt, D., Holdsworth, E., Brown, S. and Howat, D. (2017) "You See But You Do Not Observe: a review of bystander intervention and sexual assault on university campuses", *Aggression and Violent Behavior,* Vol. 35, pp. 13-25.

Leung, S. O. (2011) "A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-Point Likert Scales", *Journal of Social Service Research,* Vol. 37 No. 4, pp. 412-421.

Manning, R., Levine, M. and Collins, A. (2007) "The Kitty Genovese Murder and the Social Psychology of Helping: the parable of the 38 witnesses", *American Psychologist,* Vol. 62 No. 6, pp. 555-562.

NUS (2011) *Hidden Marks: a study of women students' experiences of harassment, stalking, violence and sexual assault*, available at: https://www.nus.org.uk/global/nus_hidden_marks_report_2nd_edition_web.pdf (accessed 29 April 2020).

NUS (2018) *Power in the Academy: staff sexual misconduct in UK higher education*, available at: https://www.nusconnect.org.uk/resources/nus-staff-student-sexual-misconduct-report (accessed 10 February 2020).Phipps, A. and Young, I. (2015) "Neoliberalisation and 'lad culture' in higher education", *Sociology,* Vol. 49 No. 2, pp. 305-322.

Reid, H. M. (2014) *Introduction to Statistics: Fundamental concepts and procedures of data analysis,* Sage, London.

Roberts, N., Donovan, C. and Durey, M. (2019) "Agency, Resistance and the Non-'Ideal' Victim: how women deal with sexual violence", *Journal of Gender-Based Violence,* Vol. 3 No. 3, pp. 323-338.

Senn, C. Y. and Forrest, A. (2016) "'And Then One Night When I Went to Class…': the impact of sexual assault bystander intervention workshops incorporated in academic courses", *Psychology of Violence,* Vol. 6 No. 4, pp. 607-618.

Stenning, P., Mitra-Kahn, T. and Gunby, C. (2013) "Sexual Violence against Female University Students in the U.K.: a case study", *Rivista di Criminologia, Vittimologia e Sicurezza,* Vol. 7 No. 2, pp. 100-119.

University of Exeter (2020) *The Intervention Initiative,* available at: https://socialsciences.exeter.ac.uk/research/interventioninitiative/universities/foruniversities/ (accessed 13 February 2020).

Universities UK (2016) *Changing the Culture: Report of the Universities UK Taskforce examining violence against women, harassment and hate crime affecting university students*, available at: http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Pages/changing-the-culture-final-report.aspx (accessed 6 December 2019).

[1] There is an 8 week programme at the university that follows-on from the taster/'introductory' session.

[2] Rough estimates suggest that this is around 6% of new students at the university who ~~engaged~~ participated in the bystander session in 2018/19.

[3] This is the foundation stage: the year prior to first-year undergraduate study.

[4] The reasons why some students did not complete the post-test survey were because they did not take part in the whole session, e.g., they had walked-out, or, for some students who remained in the session, they chose not to complete it.

[5] Examples of why cases were not useable were because students did not provide usernames and there were non-responses across the two surveys so cases/variables could not be paired for testing.

[6] Failure to remove cases of non-responses on scale items could deflate/inflate the score of the scale for testing (see for example de Vaus, 2014).

[7] Difference in figures are non-responses.

[8] This was a reverse scale and the equal number of scale points did not allow us to reverse code the data.

[9] ~~There is an 8 week programme at the university that follows-on from the short bystander session, but n~~Not many students ~~engage~~ participate in ~~it~~ the 8 week programme that follows the taster session, at the university~~or its entirety~~.

Tables for Article (SCJ)

Table I: Aims of the intervention, number of participants, and missing responses

Aim	n = participants (pre-intervention)	n = participants (post-intervention)	n = missing responses
i	121	121	0
ii	116	116	5
iii	120	120	1
iv	121	121	0

~~**Table II**: Means (and standard deviations) of aim i – knowledge scale:~~

	~~Before Bystander~~	~~After Bystander~~
~~Mean (SD)~~	~~3.48 (2.87)~~	~~9.66 (0.89)~~

~~**Table III**: Means (and standard deviations) of aim i – confidence scale (3 scale items):~~

	~~Before Bystander~~	~~After Bystander~~
~~Mean (SD)~~	~~19.02 (7.18)~~	~~23.75 (6.05)~~

~~**Table IV**: Means (and standard deviations) of aim ii - victim-empathy scale.~~

	~~Before Bystander~~	~~After Bystander~~
~~Mean (SD)~~	~~8.23 (2.21)~~	~~8.94 (1.77)~~

~~**Table V:** Means (and standard deviations) of aim ii – victim-blaming scale (2 scale items):~~

	~~Before Bystander~~	~~After Bystander~~
~~Mean (SD)~~	~~5.87 (3.75)~~	~~5.07 (3.54)~~

Table VI: Means (and standard deviations) of aim iii.

	Before Bystander	After Bystander
Mean (SD)	5.34 (3.01)	9.07 (1.65)

Table VII: Means (and standard deviations) of aim iv:

	Before Bystander	After Bystander
Mean (SD)	5.99 (2.84)	7.97 (2.47)