Trigger Warning Efficacy: The Impact of Warnings on Affect, Attitudes, and Learning

Preprint · May 2019

CITATIONS
0

READS
170

4 authors, including:

Guy Boysen
McKendree University
58 PUBLICATIONS **805** CITATIONS

SEE PROFILE

Trigger Warning Efficacy: The Impact of Warnings on Affect, Attitudes, and Learning

Guy A. Boysen

Raina Issacs

Lori Tretter

Sydnie Markowski

McKendree University

In press at *Scholarship of Teaching and Learning in Psychology* (5/15/2019).

Guy A. Boysen, Ph.D., Department of Psychology, McKendree University. Raina Issacs, Department of Psychology, McKendree University. Lori Tretter, Department of Psychology, McKendree University. Sydnie Markowski, Department of Psychology, McKendree University.

Correspondence concerning this article should be addressed to Guy A. Boysen, Department of Psychology, McKendree University, 701 College Rd., Lebanon, IL 62254, USA. Telephone: 618-537-6882. E-mail:gaboysen@mckendree.edu

Abstract

The purpose of trigger warnings is to prevent distress by giving prior notice about sensitive topics, but there is little empirical evidence to support their effectiveness in psychology education. The current research examined the effects of trigger warnings on affect, learning, and attitudes. Study 1 (N = 353) presented an online sample of adults with a video lecture about sexual assault, and participants reported their positive and negative affect before and after the video. They also took a test on the content and reported their attitudes about the necessity of warnings. Learning about sexual assault led to significant changes in affect for participants with and without personal experience related to the topic. Trigger warnings had no significant impact on changes in affect or test scores. However, participants who received a trigger warning had significantly increased belief that warnings are necessary for the topic of sexual assault. Study 2 (N = 412) replicated Study 1 using the topic of suicide. Trigger warnings had no significant effect on changes to affect or test scores but did significantly increase perceptions of warnings as necessary. Study 3 examined a sample of college students (N = 105) learning about sexual assault, and it also showed no significant effect of trigger warnings on changes to affect or test scores but a significant effect on belief that warnings are necessary. Overall, trigger warnings appear to have little impact on affect or learning, but they do increase people's belief that warnings are necessary for sensitive topics.

Keywords: education and training, undergraduate education, trigger warning, teaching of psychology

Trigger Warning Efficacy: The Impact of Warnings on Affect, Attitudes, and Learning

Imagine that a novice teacher has approached you for advice about an upcoming lesson. The teacher has the perfect research example for class, but the research is about crime victims and contains definitions of terms like sexual assault and rape, as well as detailed statistics about the prevalence of these crimes among college students. The teacher has heard about the practice of warning students before mentioning sensitive topics like sexual assault and wants to know "Should I give a trigger warning?"

Novice teachers are not alone in needing guidance on trigger warnings. Despite a growing empirical literature on the topic (Bellet, Jones, & McNally, 2018; Boysen, Wells, & Dawson, 2016; Boysen & Prieto, 2018; Gainsburg & Earl, 2018), the effectiveness of trigger warnings as a pedagogical technique remains unclear. Trigger warnings are relevant to psychology teachers because sensitive topics frequently emerge in psychology courses and because some students are attracted to psychology as a means of understanding their own mental health. As such, the purpose of the current research was to explore the impact of trigger warnings on affect, learning, and attitudes in the context of psychology education.

Trigger warnings originated as an informal method for handling symptoms of posttraumatic stress disorder. People with posttraumatic stress disorder often experience distress when presented with reminders of their trauma, so a norm developed in online communities to preface discussions of trauma-related topics with warnings about content (Medina, 2014; Vingiano, 2014). Eventually, the norm expanded to include warnings related to other forms of psychopathology and potentially distressing topics in general. After a few high-profile cases of college students requesting warnings for course content, a national debate emerged about the legitimacy, efficacy, and practicality of trigger warnings (Essig, 2014; Lukianoff & Haidt, 2015;

Stokes, 2014). Logical arguments can be made for and against the use of trigger warnings in the classroom, but an evidence-based approach to pedagogy requires documentation of their effectiveness using empirical methods.

One challenge in evaluating the effectiveness of trigger warnings is the absence of a standard definition of the concept. Some researchers have adopted the colloquial definition of trigger warnings as prior notice about any material that might be distressing (Beverly, Diaz, Kerr, Balbo, Prokopakis, & Fredricks, 2018). However, such a definition would include warnings about topics that are distressing to people in general (e.g., depictions of extreme violence, images of medical injuries), and the debate over trigger warnings is not about whether warnings are needed for universally distressing material but if warnings are needed for topics that are distressing to narrow group of people with mental health problems. Thus, the more precise definition of trigger warnings is prior notification of material that could evoke distress related to a clinical mental health problem (Boysen, 2017; Cares, Franklin, Fisher, & Bostaph, 2018). Using this more precise definition, the key research question is if warnings about mental health triggers are effective when implemented as part of standard instructional practice.

When considering trigger warnings within the context of mental health problems, there is indirect evidence for their short-term effectiveness. As outlined in a previous review of basic and clinical research related to trigger warnings (Boysen, 2017), some mental disorders include symptoms of distress that are automatically activated by external stimuli, and avoidance of those stimuli reduces immediate distress. For example, warnings could allow people with posttraumatic stress disorder to avoid reminders of their trauma that automatically trigger negative emotions, and warnings could allow people with specific phobias to avoid phobic stimuli that automatically trigger anxiety. Furthermore, warnings should benefit the learning of

students with mental health problems if it allows their arousal levels to stay within the normal range rather than being significantly elevated; it is difficult to learn while panicking (Boysen, 2017). If this basic and clinical research generalizes to the classroom, trigger warnings could be an appropriate educational accommodation for people diagnosed with certain mental disorders.

In contrast to people with mental health problems, the potential effects of trigger warnings on people without mental disorders are unclear. Trigger warnings may have a generalizable effect, causing both clinical and nonclinical populations to experience less distress when exposed to sensitive topics. However, trigger warnings could also have no effect or even a negative effect on people without mental health problems. Trigger warnings label topics as potentially harmful. As such, trigger warnings could induce negative emotional reactions through nocebo effects (Barsky, Saintfort, Rogers, & Borus, 2002) or demand characteristics (Orne, 1962; Orne & Scheibe, 1964). The effect on learning is also unclear. Distractions hinder the ability to learn (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996), but trigger warnings might increase or decrease distracting thoughts among nonclinical populations. Furthermore, describing educational material as potentially alarming could facilitate learning by increasing students' interest and attention, or it could inhibit learning by causing overarousal.

Although the number of studies is small, some empirical evidence exists on the effectiveness of trigger warnings. One measure of effectiveness is students' satisfaction with teachers' use of trigger warnings. Surveys reveal that students have mixed attitudes about trigger warnings. When considering trigger warnings in general, students hold slightly positive attitudes about their use and helpfulness (Boysen et al., 2018; Cares et al., 2018). However, when considering specific topics and specific courses, students view them as unnecessary in many cases (Bentley, 2017; Beverly et al., 2018; Boysen et al., 2018; Cares et al., 2018). Furthermore,

they tend to believe that sensitive topics are an unavoidable part of education that students must learn to face (Beverly et al., 2018; Boysen et al., 2018). The largest of these satisfaction studies occurred using samples of psychology students (Boysen et al., 2018), so the trends are relevant to psychology teachers. Although students' satisfaction with trigger warnings is a valuable starting point, additional research is needed because perceptions of educational effectiveness can be inaccurate.

The ultimate test of trigger warning effectiveness is to randomly assign students to receive or not receive warnings when learning about a sensitive topic. If effective, trigger warnings should decrease distress and increase learning. Some researchers have taken this experimental approach, but the results have been inconsistent. Although one study showed that trigger warnings decrease anxiety (Gainsburg & Earl, 2018), another study showed that they increase anxiety among people who believe in the harmfulness of offensive words (Bellet et al., 2018). In yet another study, trigger warnings had no significant impact on negative affect, a finding the researcher replicated across multiple online and student samples (Sanson, Strange, & Gary, 2019). The same study found that trigger warnings had no consistent effect on learning. Overall, it is impossible to draw a conclusion about the effect of trigger warnings on distress or learning because of the limited number of experimental studies and their inconsistent findings.

Experimental studies have demonstrated consistent effects of trigger warnings on various attitudes. Experimentally delivered trigger warnings have increased participants' expectation that educational material would be negative (Sanson et al., 2019), increased their belief that material would cause anxiety and should be avoided (Gainsburg & Earl, 2018), and increased their perceptions of themselves and others as vulnerable to the effects of trauma (Bellet et al., 2018).

Overall, trigger warnings seem to increase the perceived impact of sensitive material and decrease perceptions of resiliency to those effects.

The decision to offer or not offer trigger warnings should be based on empirical data, but experimental evidence for the effectiveness of trigger warnings is inconclusive. In addition, previous studies did not emphasize sensitive topics that are typical of psychology courses (Bellet et al., 2018; Sanson et al., 2019). As such, the current research examined the impact of trigger warnings in educational contexts with relevance to psychology. Across three studies, participants either received or did not receive a trigger warning before learning facts about a sensitive topic related to psychology. Specifically, participants listened to brief, informative lectures about sexual assault or suicide, which are topics that might emerge in variety of psychology courses. To assess the impact of warnings, participants reported their positive and negative affect and they took a test about the sensitive topic. In addition, they reported their attitudes about the need for warnings. These procedures occurred using online samples from the general population (Studies 1 and 2) and a face-to-face sample of college students (Study 3).

Study 1

To examine the efficacy of trigger warnings, participants in Study 1 received or did not receive a trigger warning before watching an educational video lecture about sexual assault. Participants reported their positive and negative affect before and after the video. They also evaluated the need for warnings and completed a test on the video content. Based on these methods, Study 1 addressed the following research questions. Do trigger warnings influence affective responses to sensitive material? Do trigger warnings affect test performance related to sensitive material? Do trigger warnings affect people's perceptions of warnings as necessary? Do trigger warnings have the same effect on people with and without personal experience related to

the warning topic? The answers to these research questions will provide evidence that teachers can use when making decisions about the use of trigger warnings in their courses.

Method

Participants. Recruitment of participants (N = 353) occurred through Amazon's Mechanical Turk website. Participants included adults in the United States and Canada who were mostly female (55%) and White (66%), and who had a mean age of 37 (SD = 12). The sample was highly educated; 62% of participants had completed a college degree program, and 29% had completed some college. Although the purpose of this research was primarily to inform instruction of face-to-face college courses, our initial studies used online samples for two reasons. First, Mechanical Turk is a valid source of participants for many forms of basic research (Buhrmester, Kwang, & Gosling, 2011; Paolacci & Chandler 2014), and our plan was to establish the efficacy of trigger warnings using more basic research methods (Studies 1 and 2) and then replicate those results using a sample of college students (Study 3). Second, researchers have previously used Mechanical Turk for studies of educational topics, including trigger warnings (e.g., Gainsburg & Earl, 2018; Wilson, Martin, Smilek, & Risko, 2018).

Materials and procedure. Participants volunteered for a study on how people learn from educational materials related to psychology. Random assignment placed participants into the no-warning or trigger-warning condition. Instructions in the no-warning condition informed participants that they would be watching a video, there would be a test over the content, and they should act as if they were students. The trigger-warning condition contained the same information but also warned participants that the video was about sexual assault and that it could "trigger extreme distress among some people, especially survivors of trauma." After the warnings, participants evaluated their preexposure positive and negative affect by rating four

items (“enthusiastic,” “excited,” “nervous,” “upset”) on a 5-point scale ranging from *very slightly or not at all* to *extremely.* Although the source of the items was an established measure of positive and negative affect (Thompson, 2007), the scales have not been validated as standalone measures. Nonetheless, internal consistency reliability was acceptable across both administrations of the scales (α = .73-.84). Also, in the current study, scores on the scales varied as expected after exposure to sensitive material, and this is suggestive of construct validity.

Participants then watched a 3-minute video about sexual assault consisting of narration over lecture slides. The video mimicked the type of fact-based, informational lecture students encounter in college courses. The video contained educational information taken from an anti-sexual-violence organization (rainn.org), including definitions of sexual assault and rape, prevalence rates for sexual assault among college-age adults, and sources of assistance for victims of sexual assault. The videos included frank use of sexual terms but contained no graphic depictions nor personal stories of sexual assault. For example, the definition of sexual assault, taken verbatim from rainn.org, read as follows: “Contact or behavior that occurs without explicit consent of the victim. Some forms of sexual assault include: attempted rape; fondling or unwanted sexual touching; forcing a victim to perform sexual acts, such as oral sex or penetrating the perpetrator’s body; penetration of the victim’s body, also known as rape.” After the video, participants completed a 4-item multiple-choice test covering factual content from the lecture. Participants could receive a small monetary bonus if they answered all test questions correctly, and there were no limitations placed on their ability to replay the video to find test answers.

After completing the test, participants reported their postexposure positive and negative affect using the previously outlined items. To measure beliefs about the necessity of trigger

warnings, participants rated three statements using a 5-point scale ranging from *strongly disagree* to *strongly agree*. Statements included "I needed a warning about the topic of sexual assault to prevent it from causing distress," "People should always receive a warning before hearing about sexual assault," and "I needed a warning about the topic of sexual assault in order to learn from it." Finally, participants indicated *yes* or *no* if they had "personal experience" with the topic of sexual assault; the survey did not define personal experience, and participants were free to determine if the descriptor applied to them.

The survey included items designed to ensure the validity of the procedures. Several items throughout the survey tested participants' attention to the questions, and the analyses excluded 17 participants who failed to attend to these items. In addition, items included after the warning assessed its impact. Before exposure to the educational materials, participants utilized a 6-point scale ranging from *none* to *an extreme amount* to indicate how much anxiety "the average person" and "someone who had a personal experience with sexual assault" would feel about the educational materials. The trigger warning condition had significantly higher ratings of anticipated anxiety (people on average: $M = 3.09$, $SD = 1.18$; people with personal experience: $M = 4.46$, $SD = 1.35$) than the no-warning condition (people on average: $M = 2.37$, $SD = 1.24$; people with personal experience: $M = 3.56$, $SD = 1.41$), all ts > 5.75, all ps $< .001$, all ds > 0.59. These results indicate that the trigger warning manipulation was effective at signaling the sensitive nature of the content to participants. One participant in the trigger warning condition did not continue with the survey after receiving the warning, but it is impossible to know the cause of the participants' discontinuation.

Results

The first set of analyses examined positive and negative affect. The analysis consisted of a 2 (Experimental condition: no-warning, trigger warning) × 2 (Experience: personal experience, no experience) × 2 (Time: preexposure, postexposure) mixed analysis of variance (ANOVA) with experimental condition and experience as the between-subjects factors and time as the within-subjects factor. Positive affect served as the dependent variable in the first analysis. Results of the ANOVAs can be seen in Table 1, and the means can be seen in Table 2. The main effect of time, the interaction between condition and time, and the three-way interaction were all significant. Post hoc *t* tests compared means to explain the three-way interaction. Examination of the means showed that positive affect significantly decreased after exposure to the educational materials among all groups, all *t*s > 6.76, all *p*s < .001, all *d*s > 0.32. Among people with personal experience, preexposure positive affect was significantly lower when there was a trigger warning than when there was no warning, $t(112) = 2.74$, $p = .007$, $d = 0.54$. No other significant post hoc effects emerged, *t*s < 1.95, all *p*s > .054, all *d*s < 0.31. Overall, these results indicate that positive affect decreased after exposure to the sensitive material irrespective of trigger warning condition or personal experience.

The analysis of negative affect was analogous to that of positive affect. The ANOVA results can be seen in Table 1, and the means can be seen in Table 2. The main effects of time and personal experience were significant. The main effect of time occurred because negative affect was significantly higher after exposure to the educational materials than before exposure. The main effect of personal experience was significant because, across experimental conditions, negative affect was higher among people with personal experience (Preexposure: $M = 1.75$, $SD = 0.94$; Postexposure: $M = 2.13$, $SD = 1.06$) than among those without personal experience (Preexposure: $M = 1.45$, $SD = 0.69$; Postexposure: $M = 1.91$, $SD = 1.00$). Overall, these results

indicate that negative affect increased after exposure to the sensitive material irrespective of trigger warning condition or personal experience

The second set of analyses explored the effects of trigger warnings on test performance. The analysis consisted of a 2 (Experimental condition: no-warning, trigger warning) × 2 (Experience: personal experience, no experience) ANOVA. Across conditions, the overall test score was 3.05 (*SD* = 0.90). There were no significant variations by condition, $F(1, 332) = 1.08$, $p = .229$, $\eta_p^2 = .003$, nor experience, $F(1, 332) = 3.45$, $p = .064$, $\eta_p^2 = .010$, and there was no significant interaction, $F(1, 328) = 0.68$, $p = .409$, $\eta_p^2 = .002$. These results indicate that trigger warnings did not significantly improve test scores among participants with or without personal experience.

The final set of analyses explored the effects of trigger warnings on attitudes. The analyses consisted of a series of 2 (Experimental condition: no-warning, trigger warning) × 2 (Experience: personal experience, no experience) ANOVAs examining participants' ratings of personal need for a warning to prevent distress, personal need for a warning to learn, and belief that warnings should always be provided. Due to the large number of tests, the results presented here emphasize significant effects. There was a significant effect of experimental condition on belief that warnings should always be provided, $F(1, 332) = 4.88$, $p = .028$, $\eta_p^2 = .014$. After exposure to the educational materials, participants who received a trigger warning (*M* = 3.51, *SD* = 1.22) believed more strongly that people should always receive a warning when learning about sexual assault than participants who received no warning (*M* = 3.16, *SD* = 1.31). There was also a main effect of experience on ratings of personal need for a warning to prevent distress, $F(1, 333) = 3.97$, $p = .047$, $\eta_p^2 = .012$. Participants with personal experience (*M* = 2.58, *SD* = 1.38) agreed more strongly that they needed a warning to prevent distress than participants with no

personal experience (M = 2.26, SD = 1.20). No other significant effects emerged, all Fs < 3.45, all ps > .064, all η_p^2s < .010. Participants' rating of their personal need for a warning to prevent distress did not vary by condition (M = 2.39, SD = 1.27) nor did their rating of their personal need for a warning to learn the material (M = 2.28, SD = 1.23).

Discussion

Study 1's first research question asked if trigger warnings influence affective responses to sensitive material. Trigger warnings did not influence affect: Irrespective of warning, positive affect decreased and negative affect increased. The second research question asked if trigger warnings affect people's perceptions of warnings as necessary. Receiving a warning led participants to believe more strongly that information about sexual assault should always be preceded by a warning, but it did not impact perceptions of their own need for a warning. The third research question asked if trigger warnings affect test performance, and test scores did not differ by warning condition. The final research question asked if trigger warnings have the same effect on people with and without experience related to the warning topic. Although people with personal experience reported more negative affect and believed more strongly that they needed a warning, they experienced no significant benefits from trigger warnings. Overall, these results suggest that trigger warnings convince people that warnings are needed while simultaneously being ineffective at protecting people from negative changes in mood. However, replication of these results is necessary, especially considering the many sensitive topics that might require trigger warnings

Study 2

Sexual assault is the topic that people believe most needs a trigger warning, but students also perceive suicide as a highly sensitive topic that is likely to elicit trigger warnings from

teachers (Boysen et al., 2018). Thus, to test the generalizability of Study 1's results, the methods of Study 2 asked participants to learn about the topic of suicide after receiving or not receiving a trigger warning. As in Study 1, participants in Study 2 reported their positive and negative affect, provided their attitudes about the necessity of warnings, and completed a test. These methods allowed Study 2 to address the same research questions as Study 1.

Method

Participants. Recruitment of participants (N = 412) for Study 2 occurred using the same methods as Study 1. Participants were mostly female (58%) and White (66%), with a mean age of 37 (SD = 12). In addition, 56% had completed a college degree program, and 33% had completed some college.

Materials and procedure. Participants volunteered for a study about educational topics in psychology. To improve on Study 1, the method of Study 2 included a full, validated measure of affect. Specifically, participants completed the Positive and Negative Affect Schedule Short Form (Thompson, 2007). The measure includes 10 items and exhibits good convergent validity with longer measures of affect. The measure includes items such as "inspired," "determined, "upset," and "nervous" that participants rated using a 5-point scale from *very slightly or not at all* to *extremely*. The procedure asked participants to report their positive and negative affect at three points during the study: before the warning (prewarning), after the warning but before exposure to the sensitive material (preexposure), and after exposure to the sensitive material (postexposure). Coefficient alpha was acceptable across all administrations of the scales (.83-.92).

After the prewarning ratings of positive and negative affect, participants either received no warning or a trigger warning analogous to Study 1; this was followed by the preexposure

ratings of positive and negative affect. Participants then watched a 3-and-a-half-minute video containing information from the American Psychological Association (www.apa.org/topics/suicide) about suicide, suicide prevention, and warning signs for suicide. The videos contained frank discussion of suicide and death but no graphic depictions nor personal stories of suicide. For example, the definition of suicide, taken verbatim from apa.org, read as follows: "Suicide is the act of killing yourself, most often as a result of depression or other mental illness." Next, participants completed a 4-item multiple-choice test over the video content with the opportunity to receive a bonus for a perfect score. Participants could watch the video an unlimited number of times. The test was followed by postexposure ratings of positive and negative affect. Participants then completed three items analogous to those from Study 1 assessing their agreements that they needed a warning to prevent distress, that people should always receive a warning, and that they needed a warning to learn. Finally, participants indicated if they had a personal experience with the topic of the video. The survey included items designed to test participants' attention to the materials, and the final analyses excluded 12 participants who failed to attend to these items. Four participants did not continue with the survey after completing the trigger warning portion of the procedure, and four participants did not continue after reaching the same point in the control condition. It is not possible to determine the reasons for these discontinuations.

Results

The first set of analyses examined positive and negative affect. The analyses consisted of a 2 (Experimental condition: no-warning, trigger warning) × 2 (Experience: personal experience, no experience) × 3 (Time: prewarning, postwarning, postexposure) mixed ANOVA with between-subjects factors of experimental condition and experience and a within-subjects

factor of time. Positive affect served as the dependent variable in the first analysis and negative affect served as the dependent variable in the second analysis. The ANOVA results can be seen in Table 3, and the means can be seen in Table 4. For both positive and negative affect, the main effects of time and experience were significant. Post-hoc tests indicated that the main effect of time occurred because exposure to the educational materials significantly reduced positive affect and significantly increased negative affect. The main effect of experience occurred because, across conditions, participants with personal experience had significantly lower positive affect and significantly higher negative affect (Positive: $M = 2.93$, $SD = 0.83$; Negative: $M = 1.63$, $SD = 0.87$) than individuals without personal experience (Positive: $M = 3.23$, $SD = 0.95$; Negative: $M = 1.43$, $SD = 0.71$). There was a two-way interaction between time and experience for positive affect, but this appeared to be a result of an unexpected postwarning elevation of positive affect among participants without personal experience. Overall, these results indicate that exposure to the sensitive material resulted in significant changes to affect irrespective of trigger warning condition or personal experience.

The second set of analyses explored the effects of trigger warnings on test performance. The analysis consisted of a 2 (Experimental condition: no-warning, trigger warning) $\times$ 2 (Experience: personal experience, no experience) ANOVA. There was no significant effect of condition, $F(1, 347) = 1.11$, $p = .739$, $\eta_p^2 < .001$, but there was a significant effect of experience, $F(1, 347) = 8.78$, $p = .006$, $\eta_p^2 = .022$, and a significant interaction, $F(1, 347) = 74.16$, $p = .042$, $\eta_p^2 = .012$. Post hoc t tests indicated that test scores were higher among participants with personal experience ($M = 3.39$, $SD = 0.86$) than those without personal experience ($M = 2.84$, $SD = 1.15$) in the trigger-warning condition, $t(176) = 3.53$, $p < .001$, $d = 0.53$, but there was no difference between participants with experience ($M = 3.12$, $SD = 1.06$) and without experience

(M = 3.03, SD = 1.11) in the no-warning condition, $t(171) = 0.51$, $p = .608$, $d = 0.08$. Test scores were not significantly higher in the warning condition for either participants with or without personal experience, all ts < 1.75, all ps > .082, all ds < .28, which means that, despite the interaction, trigger warnings did not lead to significantly improved test scores. Overall, these results indicate that trigger warnings did not allow participants to significantly improve their test scores.

The final set of analyses explored the effects of trigger warnings on attitudes. The analyses consisted of a series of 2 (Experimental condition: no-warning, trigger warning) × 2 (Experience: personal experience, no experience) ANOVAs. Three separate ANOVAs occurred for participants' ratings of personal need for a warning to prevent distress, personal need for a warning to learn, and belief that warnings should always be provided. Due to the large number of tests, the results presented here emphasize significant effects. There was a significant effect of experimental condition on participants' ratings of whether people in general need warnings before learning about suicide, $F(1, 345) = 7.00$, $p = .009$, $\eta_p^2 = .020$. Participants who received a trigger warning (M = 3.26, SD = 1.23) believed more strongly that warnings are necessary than participants who did not receive a warning (M = 3.00, SD = 1.25). No other significant effects emerged, all Fs < 1.14, all ps > .286, all η_p^2s < .003. Participants' rating of their personal need for a warning to prevent distress (M = 2.17, SD = 1.26) did not vary by condition nor did their rating of their personal need for a warning to learn the material (M = 2.25, SD = 1.28).

Discussion

Using the sensitive topic of suicide rather than sexual assault, Study 2 replicated many of Study 1's answers to the research questions. Once again, exposure to the sensitive topic reduced positive affect and increased negative affect irrespective of warning condition. Nonetheless,

participants who received a warning believed more strongly that warnings are necessary for people in general. These trends were analogous among people with and without personal experience related to suicide. Participants in Study 2 rated their affect three times: before receiving the trigger warning, after receiving the trigger warning, and after exposure to the sensitive material. Although some previous research suggests that warnings can have negative effects (Gainsburg & Earl, 2018), the current results showed that only exposure to the sensitive material, not the warning, led to changes in affect. Analysis of test scores in Study 2 did yield a significant interaction between personal experience and trigger warning condition. When they received a trigger warning, people who reported personal experience related to suicide performed better on the test than people without personal experience. However, there was not a significant difference between participants who received a warning and those who did not receive a warning, so this result does not provide clear evidence for the benefit of trigger warnings.

Study 3

Studies 1 and 2 produced consistent findings across two online samples, 90% of whom had completed at least some college. However, the primary purpose of the current research is to inform teachers' use of trigger warnings in a college setting. The effect of trigger warnings may be different when administered face-to-face inside of a classroom. Thus, the purpose of Study 3 was to replicate the previous studies in college classrooms using a sample of college students.

Method

Participants. Participants (N = 105) consisted of a convenience sample of undergraduate students at a medium-sized private university in the Midwest. The sample was mostly female (61%) and White (68%), with an average age of 20 (SD = 2). Recruitment occurred by asking for

volunteers among students enrolled in upper- (n = 66) and lower-level (n = 39) courses inside and outside of psychology.

Materials and Procedure. Recruitment of participants occurred during class time. The researchers asked students to volunteer for a study on educational topics in psychology. After obtaining informed consent, participants reported their prewarning positive and negative affect using the same scale as in Study 2 (Thompson, 2007). Because no differences emerged between prewarning and postwarning ratings of affect in Study 2, the method of Study 3 only included ratings before the warning and after exposure to the sensitive educational materials. Coefficient alphas for all administrations of the scales were acceptable (.76-.82). The researchers then verbally introduced the sexual assault video from Study 1 by reading the text from the no-warning or trigger-warning condition. Random assignment occurred such that all students in the same classroom were in the same condition. After watching the video, participants completed the same 4-item, multiple-choice test from Study 1, and they rated their postexposure positive and negative affect. Next, participants completed the same three items from Study 1 assessing agreement that they needed a warning to prevent distress, that people should always receive a warning, and that they needed a warning to learn. Finally, the survey included the Anxiety Sensitivity Index for exploration as a potential covariate (Taylor et al., 2007), but the scale was unrelated to the other measures and unaffected by the experimental manipulation. The results do not include these nonsignificant analyses. In addition, the number of participants reporting personal experience with sexual assault (n = 25) provided insufficient statistical power to be included as part of the analysis.

A manipulation-check item asked participants to indicate their agreement that the experimenter provided them with a verbal warning that the video would be about sexual assault,

and there was a significant difference in the expected direction with a large effect size (d = 1.20), which indicates that participants attended to the warning in the trigger-warning condition. One participant chose to leave the room in the control condition for unknown reasons, and no participants did so in the trigger warning condition.

Results

The first set of analyses examined positive and negative affect using a 2 (Experimental condition: no-warning, trigger warning) × 2 (Time: prewarning, postexposure) mixed ANOVA with a between-subjects factor of experimental condition and a within-subjects factor of time. Positive affect served as the dependent variable in the first analysis and negative affect served as the dependent variable in the second analysis. ANOVA results can be seen in Table 5. The main effects of time were significant for both positive and negative affect. Positive affect significantly decreased from prewarning (M = 2.81, SD = 0.85) to postexposure (M = 2.67, SD = 0.89), and negative affect significantly increased from prewarning (M = 1.25, SD = 0.42) to postexposure (M = 1.57, SD = 0.57). There was no significant effect of condition nor an interaction, and these results indicate that exposure to the sensitive material resulted in significant changes to affect irrespective of trigger warning condition.

The second set of analyses examined the effects of trigger warnings on test performance and attitudes. A series of t tests compared test scores and attitudes about warnings in the no-warning and trigger-warning conditions. Results can be seen in Table 6. Test scores were not significantly different between conditions. However, two attitude measures did show significant differences. When there was a trigger warning, participants believed more strongly that they needed a warning to prevent distress and that people should always be warned before hearing about sexual assault.

Discussion

The results of Study 3 suggest that trigger warnings have similar efficacy when tested using a college sample rather than an online sample. Specifically, trigger warnings did not protect participants from changes in affect that occurred after exposure to sensitive material; with or without a warning, positive affect decreased and negative affect increased. In addition, warning about sensitive content significantly increased people's perceptions of warnings as necessary for themselves and others. Overall, these results support the findings of past research indicating that trigger warnings have little effect regardless of whether they are tested on online samples or college students (Sanson et al., 2019).

General Discussion

The current research provided answers to four research questions about the effectiveness of trigger warnings in an educational context. Do trigger warnings influence affective responses to sensitive material? Across three studies, no evidence emerged to support trigger warnings' influence over emotional regulation. With or without a warning, sensitive content significantly reduced positive affect and increased negative affect. Do trigger warnings affect test performance related to sensitive material? No overall effect of trigger warnings on text performance emerged. In Study 2, people reporting a personal experience with suicide had elevated test scores after receiving a trigger warning, but their scores were not significantly different from people with personal experience who did not receive a warning. The results probably reflect the increased knowledge associated with having relevant personal experiences. Do trigger warnings affect people's perceptions of warnings as necessary? Participants who received warnings believed more strongly that people in general need warnings before being exposed to sensitive content, but there was no consistent effect on perceptions of their own need

for warnings. Do trigger warnings have the same effect on people with and without personal experience related to the warning topic? People who reported having relevant personal experiences with sensitive topics tended to have more negative emotions and believe more strongly in their need for warnings, but trigger warnings were similarly irrelevant to emotional regulation regardless of personal experience.

The results of Studies 1 through 3 were consistent with several findings from previous experimental research on trigger warnings. In the current research, trigger warnings significantly influenced attitudes by increasing the belief that people need warnings for sensitive material. Similarly, in previous studies, trigger warnings increased people's belief that sensitive material has a serious impact and decreased their perceptions of people as resilient to that material (Bellet et al., 2018; Gainsburg & Earl, 2018; Sanson et al., 2019). The current research was also consistent with previous research showing that trigger warnings do not significantly influence reactions to sensitive material (Sanson et al., 2019). Although the results of single studies have suggested that trigger warnings may increase or decrease anxiety (Bellet et al., 2018; Gainsburg & Earl, 2018), the effects await independent replication.

Based on the significant and nonsignificant effects of trigger warnings, the current research has several implications for teachers. When making choices about the use of trigger warnings in their classrooms, teachers should consider their goal in offering a warning. If teachers want to reduce distress among students in general or among students who have personal experience with a sensitive topic, then trigger warnings appear to be an ineffective tool. Similarly, if the goal is to increase learning, trigger warnings appear to have limited relevance. The same trends emerged for the topics of sexual assault and suicide, in online and face-to-face

settings, and among traditionally-aged college students and nonstudent adults. As such, teachers in a variety of educational contexts can generalize these results to their work.

Although the current research demonstrated no significant benefit of trigger warnings, teachers should consider some special contexts when making decisions about the use of warnings in their classrooms. When covering potentially sensitive topics in class, teachers should be aware of the severity of sensitive materials, the predictability of sensitive topics within the course content, and the presence of students with relevant personal histories. For example, a warning might be advisable for a teacher showing a video of a person discussing a suicide attempt during the personality module of an introductory psychology course, especially if there was a recent suicide on campus. In contrast, a warning is probably unnecessary for a teacher providing suicide statistics in a graduate course on psychopathology. Furthermore, teachers should consider that research has not eliminated the possibility that trigger warnings reduce automatically-triggered distress among students with mental disorders.

The main argument for trigger warnings is reduction of distress, but a secondary argument is their importance in making students aware of trauma and the sometimes-debilitating effects of trauma-related disorders (Essig, 2014; Wyatt, 2016). For teachers whose goal is to increase students' awareness and acceptance that some people need to be warned about sensitive topics, then trigger warnings appear to be an effective tool because they tend to increase belief in the necessity of trigger warnings. Although the impact of trigger warnings on attitudes was reliable, the size of the effect indicates that teachers are unlikely to completely change students' preexisting attitudes. In addition, teachers should carefully consider the practice of increasing students' belief in the necessity of an intervention with questionable efficacy. Overall, the

implications of trigger warnings for affect, learning, and attitudes must be judged based on their relatively trivial effects (Sanson et al., 2019).

Although the current studies contribute well-controlled, experimental evidence to the trigger warning debate, some limitations should be noted. The most important limitation was the use of samples from the general population rather than people who experience automatically-triggered distress due to a mental disorder. Many participants in the current research reported personally relevant experience with the sensitive topics, but the extent and severity of their experiences was not assessed. Clinical research suggests that trigger warnings may be a strategic way for such individuals to avoid immediate distress (Boysen, 2017). Thus, it is possible that trigger warnings are ineffective among the general population but effective among clinical populations.

Another major limitation was the use of artificial educational scenarios rather than real college courses. As with any laboratory-based research on teaching and learning, the effects need to be tested in real educational settings. For example, tests of immediate comprehension may not generalize to tests of long-term retention. Also, the tightly-controlled procedures of this research omitted many classroom dynamics that might increase or decrease the effectiveness of trigger warnings. For instance, the sensitive material was not part of a broader lesson that might have influenced students' ability to prepare for its emotional effects, and the warnings were not administered by teachers who could use their immediacy skills to assess students' preparedness for sensitive material. A final limitation was the educational materials. Although the materials led to significant changes in affect, they emphasized facts rather than graphic depictions or personal narratives, and the practical effects on emotion were likely small for most participants. More graphic, personal materials could produce different results.

Future research should begin by addressing the limitations of the current research. Researchers should replicate and extend the findings using a large sample of students enrolled in a course with sensitive content. Although the online and student samples produced consistent results, laboratory findings do not always generalize to the classroom. In addition to improving on the current research, future studies should explore the key unanswered question about trigger warnings: Are they effective for people who have a mental disorder that causes automatically-triggered distress? Despite anecdotal reports from students and teachers, no research has directly tested the effectiveness of trigger warnings among students with mental health problems. The current research suggests that people believe in the necessity of trigger warnings even when they have no significant impact, so future research needs to directly test the effectiveness of trigger warnings as a mental health accommodation for students with documented disabilities. Although the trigger warning debate has emphasized trauma-based symptoms, researchers should explore accommodations for students with automatically-elicited distress related to symptoms of phobias, panic disorder, and obsessive-compulsive disorder.

In conclusion, the current research cannot directly answer the question "Should I give a trigger warning?" but it does suggest that teachers can make a decision knowing that trigger warnings are unlikely to impact most students' level of distress or their ability to learn. On average, warning or no warning, people have modest affective reactions while learning about sensitive topics such as sexual assault and suicide. Nonetheless, students are likely to see warnings about sensitive content as helpful, especially if they have received such warnings in the past. Like many educational trends, the importance of trigger warnings appears to have been overstated, and empirical evidence supports, at best, their judicious application rather than wholesale adoption.

References

Barsky, A. J., Saintfort, R., Rogers, M. P., & Borus, J. F. (2002). Nonspecific medication side effects and the nocebo phenomenon. *JAMA: Journal of the American Medical Association, 287*, 622-627. doi:10.1001/jama.287.5.622

Bellet, B. W., Jones, P. J., & McNally, R. J. (2018). Trigger warning: Empirical evidence ahead. *Journal of Behavior Therapy and Experimental Psychiatry, 61*, 134–141. doi:10.1016/j.jbtep.2018.07.002

Bentley, M. (2017). Trigger warnings and the student experience. *Politics, 37,* 470–485. https://doi.org/10.1177/0263395716684526

Beverly, E. A., Díaz, S., Kerr, A. M., Balbo, J. T., Prokopakis, K. E., & Fredricks, T. R. (2018). Students' perceptions of trigger warnings in medical education. *Teaching and Learning in Medicine, 30*, 5–14. doi:10.1080/10401334.2017.1330690

Boysen, G. (2017). Evidence-based answers to questions about trigger warnings for clinically-based distress: A review for teachers. *Scholarship of Teaching and Learning in Psychology, 3*, 163-177. doi:10.1037/stl0000084

Boysen, G. A., & Prieto, L. R. (2018). Trigger warnings in psychology: Psychology teachers' perspectives and practices. *Scholarship of Teaching and Learning in Psychology, 4*, 16-26. doi:10.1037/stl0000105

Boysen, G. A., Prieto, L. R., Holmes, J., Landrum, R. E., Miller, R., Taylor, A., White, N., & Kaiser, D. J. (2018). Trigger warnings and psychology students: What do they think? *Scholarship of Teaching and Learning in Psychology, 4,* 69-80. doi:10.1037/stl0000106

Boysen, G. A., Wells, A. M., & Dawson, K. J. (2016). Instructors' use of trigger warnings and behavior warnings in Abnormal Psychology. *Teaching of Psychology, 43,* 334-339. doi:10.1177/0098628316662766

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3-5. doi:10.1177/1745691610393980

Cares, A. C., Franklin, C. A., Fisher, B. S., & Bostaph, L. G. (2018). "They were there for people who needed them": Student attitudes toward the use of trigger warnings in victimology classrooms. *Journal of Criminal Justice Education*. Advance online publication. doi:10.1080/10511253.2018.1433221

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*, 159–180. https://doi.org/10.1037/0096-3445.125.2.159

Essig, L. (2014, March 10). Trigger warnings trigger me. *The Chronicle of Higher Education.* Retrieved from www.chronicle.com/blogs/conversation/2014/03/10/trigger-warnings-trigger-me/

Gainsburg, I., & Earl, A. (2018). Trigger warnings as an interpersonal emotion-regulation tool: Avoidance, attention, and affect depend on beliefs. *Journal of Experimental Social Psychology*, *79*, 252–263. doi:10.1016/j.jesp.2018.08.006

Knox, E. J. M. (Ed.). (2017). *Trigger warnings: History, theory, context.* Lanham, MA: Rowman & Littlefield.

Lukianoff, G., & Haidt, J. (2015, September). The coddling of the American mind. *The Atlantic.* Retrieved from www.theatlantic.com/magazine/archive/2015/09/the-coddling-of-the-american-mind/399356/?utm_source=hpfb

Medina, J. (2014, May 17). Warning: The literary cannon makes students squirm. *The New York Times.* Retrieved from www.nytimes.com/2014/05/18/us/warning-the-literary-canon-could-make-students-squirm.html

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17,* 776-783. doi:10.1037/h0043424

Orne, M. T., & Scheibe, K. E. (1964). The contribution of nondpreivational factors in the production of sensory deprivation effects: The psychology of the "panic button." *Journal of Abnormal and Social Psychology, 86,* 3-12.

Paolacci, G., & Chandler, J. (2014). Inside the Turk understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23*, 184-188. doi:10.1177/0963721414531598

Sanson, M., Strange, D., & Garry, M. (2019). Trigger warnings are trivially effective at reducing negative affect, intrusive thoughts, and avoidance. *Clinical Psychological Science.* doi:10.1177/2167702619827018

Stokes, M. (2014, May 29). In defense of trigger warnings. *The Chronicle of Higher Education.* Retrieved from www.chronicle.com/blogs/conversation/2014/05/29/in-defense-of-trigger-warnings/

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology, 38,* 227-242. doi:10.1177/0022022106297301

Veraldi, L., & Veraldi, D. M. (2015). Stressors, triggers, and trauma: Considering the DSM-5 in the debate over campus trigger warnings. *American Journal of Forensic Psychology, 33,* 5-17.

Vingiano, A. (2014). How the "trigger warning" took over the Internet. *BuzzFeedNews.* Retrieved from www.buzzfeed.com/alisonvingiano/how-the-trigger-warning-took-over-the-internet?utm_term=.dyOQKXp34#.xv9arPm2Z

Wilson, K. E., Martin, L., Smilek, D., & Risko, E. F. (2018). The benefits and costs of speed watching video lectures. *Scholarship of Teaching and Learning in Psychology, 4,* 243-257. doi:10.1037/stl0000127

Table 1

Study 1 Analysis of Variance Results for Positive and Negative Affect

	Positive affect			Negative affect		
Variable	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Experimental condition (C)	3.72	.055	.011	0.34	.559	.001
Experience (E)	0.31	.576	.001	7.63	.006	.022
Time (T)	142.00	< .001	.300	76.88	< .001	.188
C × E	0.02	.891	< .001	1.14	.286	.003
C × T	12.56	< .001	.037	3.51	.062	.010
E × T	2.18	.141	.007	0.39	.535	.001
C × E × T	10.05	.002	.029	0.06	.808	.000

Table 2

Study 1 Positive and Negative Affect Means and Standard Deviations

	Overall	No warning		Trigger warning	
		Personal experience	No experience	Personal experience	No experience
Time	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
		Positive affect			
Preexposure	2.25 (1.11)	2.52 (1.11)	2.36 (1.13)	1.96 (1.07)	2.15 (1.08)
Postexposure	1.63 (0.93)	1.65 (0.95)	1.66 (0.97)	1.77 (1.04)	1.49 (0.80)
		Negative affect			
Preexposure	1.55 (0.79)	1.62 (0.85)	1.42 (0.67)	1.85 (0.99)	1.48 (0.72)
Postexposure	1.99 (1.02)	2.09 (1.01)	1.98 (1.05)	2.17 (1.10)	1.83 (0.93)

Note. The number of participants per condition was as follows: trigger warning with personal experience = 62, trigger warning without personal experience = 104; control with personal experience = 53, control without personal experience = 118. All preexposure and postexposure comparisons of positive and negative affect are significantly different, $p < .05$.

Table 3

Study 2 Analysis of Variance Results for Positive and Negative Affect

	Positive affect			Negative affect		
Variable	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Experimental condition (C)	0.15	.704	< .001	0.80	.372	.002
Experience (E)	11.50	.001	.034	4.18	.042	.012
Time (T)	26.75	< .001	.075	14.49	< .001	.042
C × E	1.13	.289	.003	0.58	.447	.002
C × T	1.86	.157	.006	0.50	.606	.001
E × T	3.25	.039	.010	0.08	.921	< .001
C × E × T	1.23	.294	.004	2.00	.136	.006

Table 4

Study 2 Positive and Negative Affect Means and Standard Deviations

	Overall	No warning		Trigger warning	
		Personal experience	No experience	Personal experience	No experience
Time	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
		Positive affect			
Prewarning	3.12 (0.93)	2.86 (0.82)	3.24 (0.95)	3.10 (0.90)	3.21 (1.00)
Preexposure	3.16 (0.96)$_a$	2.88 (0.87)	3.39 (1.00)	3.03 (0.84)	3.27 (1.01)
Postexposure	2.96 (0.99)$_a$	2.72 (0.86)	3.16 (1.06)	2.77 (0.93)	3.10 (1.01)
		Negative affect			
Prewarning	1.44 (0.78)	1.47 (0.78)	1.39 (0.71)	1.60 (0.97)	1.34 (0.65)
Preexposure	1.46 (0.77)$_b$	1.47 (0.74)	1.38 (0.73)	1.65 (0.90)	1.38(0.70)
Postexposure	1.57 (0.82)$_b$	1.62 (0.74)	1.46 (0.76)	1.70 (0.92)	1.54 (0.82)

Note. The number of participants per condition was as follows: trigger warning with personal experience = 80, trigger warning without personal experience = 94; control with personal experience = 72, control without personal experience = 89. Means in the overall column that share a subscript are significantly different, $p < .05$.

Table 5

Study 3 Analysis of Variance Results for Positive and Negative Affect

Variable	Positive affect			Negative affect		
	F	*p*	η_p^2	*F*	*p*	η_p^2
Experimental condition (C)	0.81	.369	.008	2.19	.142	.021
Time (T)	8.27	.005	.074	36.85	< .001	.264
C × T	0.12	.902	< .001	1.33	.252	.013

Table 6

Study 3 Means, Standard Deviations, and t-test Results

Item	No warning *M* (*SD*)	Trigger warning *M* (*SD*)	*t*	*p*	*d*
Test score	3.23 (0.87)	3.06 (0.83)	1.01	.311	0.20
I needed a warning about the topic of sexual assault to prevent it from causing distress.	1.62 (1.00)	2.19 (1.27)	2.55	.012	0.50
I needed a warning about the topic of sexual assault in order to learn about it.	1.87 (1.18)	2.06 (1.09)	0.86	.394	0.17
People should always receive a warning before hearing about sexual assault.	3.45 (1.25)	4.19 (1.10)	3.21	.002	0.63

Note. The number of participants per condition was as follows: trigger warning = 52, control with personal experience = 53.