

Running head: TRIGGER WARNINGS ON SOCIAL MEDIA

A qualitative investigation of people's attitudes towards the use of
trigger warnings on social media.

Orla Hyland
X19756151

National College of Ireland
BA (Hons) Psychology

Supervisor: Dr. Fergal O'Brien
Year of Submission: 2023

Submission of Thesis and Dissertation

National College of Ireland
Research Students Declaration Form
(Thesis/Author Declaration Form)

Name: Orla Hyland

Student Number: X19756151

Degree for which thesis is submitted: BA (Hons) Psychology

Title of Thesis: A qualitative investigation of people's attitudes towards the use of trigger warnings on social media.

Date: 12/03/2023

Material submitted for award

- A. I declare that this work submitted has been composed by myself. ✓
- B. I declare that all verbatim extracts contained in the thesis have been distinguished by quotation marks and the sources of information specifically acknowledged. ✓
- C. I agree to my thesis being deposited in the NCI Library online open access repository NORMA. ✓
- D. *Either* *I declare that no material contained in the thesis has been used in any other submission for an academic award.
Or *I declare that the following material contained in the thesis formed part of a submission for the award of

(State the award and the awarding body and list the material below)

✓

Acknowledgements

Firstly I would like to thank every person who agreed to take part in my study. Without your commitment and participation this research study would not have been possible.

I would also like to thank my Final Year Project supervisor Dr. Fearghal O'Brien and Final Year Project module lecturer Dr. Michelle Kelly for your advice and guidance throughout the year. In addition I would like to thank the NCI student support services for providing care and assistance throughout my degree.

I wish to acknowledge my amazing family and friends who without their love, reassurance and support this research project would not have been possible. I would particularly like to thank my Mam, Dad, brother Conor and cousin Niamh who were particular pillars of support throughout this time.

Finally, I wish to give a big thanks to the Linn Dara School and Transition Programme. Thank you so much for reintroducing me back into education at a time of great vulnerability. The care and support I received within the school helped me grow and heal and fall back in love with education. Without the amazing support of the Linn Dara school none of this would have been possible.

Table of Contents

Abstract.....	1
Introduction.....	2
Methods.....	7
Design.....	7
Research’s Positions.....	7
Participants	8
Materials.....	8
Electronics and Software	8
Interview Agenda	9
Procedures	10
Data Analysis	11
Ethical Considerations.....	12
Approval and Exclusion criteria	12
Obtaining and Documenting Consent	12
Data Handling and Protection	13
Results.....	14
Thematic map of results.....	14
Care and Awareness of Others	15
Misuse as a Prominent Issue	16
Importance of Choice	19
Discussion.....	21
Additional Limitations, Strengths and Implications of Research	26
References.....	28
Appendices.....	32

Abstract

Psychological research has found that exposure to distressing or sensitive content on social media can have severe psychological effects on individuals. Trigger warnings are a widely used method of warning individuals of distressing content they may encounter both on social media and across other mediums. However, despite their use across mediums, research regarding the use of trigger warnings in various settings is lacking, with the majority of research being carried out in the context of their use within an educational setting. Due to the intended use of trigger warnings to warn and protect of distressing content and the problems associated with distressing content on social media, it appeared pertinent to conduct research in this area. A qualitative study was carried out to investigate individual's attitudes towards the use of trigger warnings in social media. The research study aimed to understand possible motivations for the use of trigger warnings on social media, perceived benefits and drawbacks regarding their use and attitudes regarding regulations of implementation of trigger warnings on social media. Following research three themes were identified; (i) Care and Awareness of Others, (ii) Misuse is a Prominent Issue and (iii) Importance of Choice. These findings demonstrated that similar attitudes towards trigger warnings occurred across platforms and highlighted possible areas of concern and interest for future research into the use of trigger warnings on social media.

Introduction

The term “trigger warning” refers to a statement or notice that is provided prior to content that is viewed as sensitive in nature (George & Hovey, 2019). They are used to inform individuals that continued engagement with said content may cause distress. The term trigger is defined by the American psychology association as a “stimulus that elicits a reaction” (American Psychological Association, 2023). In this definition the trigger does not have to be one of negative connotation, however, the term trigger warning derives more from terminology associated with PTSD. In these cases a trigger is used to describe something that causes the individual to have an adverse reaction due to exposure to a stimulus that reminds the individual of their previous trauma (Bellet et al., 2018; McDonald et al., 2019). While trigger warnings are not only aimed at those suffering from PTSD the implication of trigger warnings is that the content following is one that can cause upset or distress particularly for those sensitive to the topic specified. The term “content warning” is also used alongside or in place of the term trigger warning. The difference in terms being one term references the content specifically, content warning, and one term references the effect the content may have on an individual, trigger warning. Content warning and trigger warnings are often used interchangeably and are regularly referred to as such within the research literature and in lay man’s terms (Charles et al., 2022). However, some content warnings can be used to precursor content that may elicit a physical reaction such as flashing images for those with epilepsy whereas the following paper is focused on warnings for content that would cause an adverse emotional response. It is for this reason that the term trigger warning will be the primary term used throughout this research study. However, literature using the term “content warning” will also be referenced throughout.

Trigger warnings are used throughout many aspects of life, primarily in relation to engagement with information or entertainment. Warnings are regularly stated on the news

prior to segments deemed particularly distressing such as violent images of an ongoing war, discussions of abuse or racial discrimination and natural disasters. These warnings are also regularly found prior to other entertainment such as television programmes, on the packaging of video games and in movies (Bridgland et al., 2019). Initial age advisories are usually provided as well as a brief reference to any explicit content such as violence, profanity or nudity that is depicted. Trigger warnings are also used within school and college settings when topics of sensitive nature are being taught and discussed (Dickman-Burnett, 2019). The use of trigger warnings in educational settings has been the predominant focus of trigger warnings related research to date (Dickman-Burnett, 2019; Wyatt, 2016).

Although the idea of trigger warnings may seem initially straightforward and helpful there are many who oppose and dispute its usefulness (Halberstam, 2017; Wyatt, 2016). A US college sparked major debate amongst both faculty and students when it introduced a policy of trigger warnings within lectures. Some staff within the college viewed this as an act of courtesy and acknowledgement to those with triggers or traumatic pasts, while others viewed it as censorship and a prevention of free speech (George & Hovey, 2019). Some argued that it is an insult to students to assume they need to be protected from such content and that it will prevent students from being able to have an overall realistic view of the world. In this context people argue that trigger warnings could be a barrier to education.

Another argument regarding the use of trigger warnings was developed in relation to sufferers of Post-Traumatic Stress Disorder (PTSD). Triggers of past traumatic events can cause a negative and adverse reaction in people with PTSD. This can lead to PTSD patients avoiding situations that could expose them to this trigger. The avoidance can become so severe that it has a debilitating effect on a sufferer's capability of performing everyday tasks and significantly impact their quality of life. Therefore, part of PTSD treatment is to refrain from and combat these avoidance behaviours. With this being the case arguably trigger

warnings, by forewarning people of triggering content, could continue to encourage people with PTSD to avoid said content. This in turn would lead to reinforcing the avoidance behaviour and preventing recovery (Jones et al., 2020; Wyatt, 2016).

Almost all arguments and research within the area of trigger warnings has been conducted with regards to their usage in an educational setting (Dickman-Burnett, 2019; Sanson et al., 2019; Wyatt, 2016) despite their regular use across other mediums. One prolific medium in which they are used in modern society is social media. Social media refers to platforms where people can post, share and engage with content and other users online (Social Media Overview, 2021). Social media is used regularly in relation to entertainment, communication and distribution/consumption of information. Content displayed on these platforms can be in the form of text posts, images or videos and are usually casual in nature. Posts regarding day-to-day life events and humour is predominantly the type of content seen on social media. However, it is not unusual for content of a sensitive or distressing nature to also be shared on these platforms. This content can revolve around subjects of violence, abuse, rape, eating disorders and animal cruelty. A study conducted by Chancellor et al., 2021 looked to analyse some of this content, in particular content on Instagram related to mental illness and eating disorders. This study analysed over 26 million posts across five years and categorised them into levels of "Mental Illness Severity". Content that would be considered of high mental illness severity would be graphic pictures of self-harm, suicidal ideation posting and dangerous depictions of low body weights and dangerous methods of losing weight. The study found that not only were there an extremely large amount of posts which could be considered distressing in nature they also found the mental illness severity of posts on social media has increased significantly overtime approximately 13% a year from 2012. Further studies investigating the implications of viewing these posts found that individuals reported negative emotional reactions when they viewed content of a sensitive

nature (Radovic et al., 2016). A study relating to self-harm behaviours found individuals reported an increase in suicidal thoughts and self-harm behaviours when exposed to distressing content surrounding the topic of self-injury on social media (Duggan et al., 2011). This study found that the majority of individuals effected did not seek out this content purposefully and engagement with said content was accidental, possibly adding to their distress.

Social media sites have been subject to scrutiny for facilitating a platform in which people can be exposed to such content. To combat this platforms have introduced trigger warnings alongside certain sensitive content. On some sites this warning would place a filter which would blur a post alongside a statement that cautions users of the nature of the content. It is at this point the individual has to click a statement explicitly saying they still wish to see the content. However, it is difficult for platforms to review and keep up with all the sensitive content being produced (Stratta et al., 2020). Additionally common users posting sensitive content may also post a trigger warning alongside their posts. However, often this does not involve blurring of the post instead there may be a trigger warning stated as a caption of an image, at the beginning of a written post or momentarily seen prior to a video. As it is the individual users posting trigger warnings in these instances the type of posts and severity of sensitive posts containing and not containing trigger warnings can vary greatly on these platforms (Charles et al., 2022). The lack of continuity and regulation regarding the use of trigger warnings may possibly cause an effect on their perceived seriousness or necessity, for this reason it could be beneficial to see if users perceive trigger warnings to be used appropriately within these platforms and their attitudes towards their effectiveness as a result.

Previous research has indicated that the perception of trigger warnings and attitudes relating to their use vary greatly (George & Hovey, 2019). In addition debate has emerged as a result of implementation of their use within educational settings (Dickman-Burnett, 2019).

However, it is conceivable to believe that attitudes may differ in the context of their use on social media, particularly as the exposure to distressing content on these platforms has shown to cause significant emotional distress to users (Duggan et al., 2011). Due to an extreme lack of research regarding the use of trigger warnings in social media the following research study wishes to obtain an in depth understanding of perceptions and attitudes towards the use of trigger warnings on social media. Possible areas of interest or concern may emerge which could influence future areas of study. A qualitative research design will allow the researcher to obtain an in depth understanding of the reasoning behind attitudes and perceptions held by individuals. Regher (2010) discusses the importance of the use of qualitative methods to analyse social issues such as this. He argues that the “generalisable simplicity” needed for quantitative research inhibits our abilities to understand the “complexity” of some social issues in context. The research aim for the following study is to gain a clearer understanding of people’s perceptions and attitudes towards the use of trigger warnings. The research study particularly wishes to better understand possible motivations and reasons behind their use, perceived benefits and drawbacks of their use and attitudes towards regulations regarding trigger warning use on social media.

Methods Section

Design

The research study follows a qualitative design which focuses on developing an understanding of the participant's attitudes towards the use of trigger warnings on social media. Data collection takes place in person via individual semi structured interviews. Questions are designed to allow individuals expand on the topic and give meanings behind their answers.

A reflexive thematic analysis was deemed appropriate for the current study. As there has been limited research into the use of trigger warnings on social media a reflexive approach allowed for flexibility in code generation and removal throughout the analytic process. The researcher also used previous debate in literature of the use of trigger warnings in different environments to social media to guide the code and theme generation through a theoretical, realist approach.

Researcher's Position

The researcher is a 23-year-old Caucasian female of Irish nationality. The researcher is a regular user of multiple social media platforms and often comes across sensitive content with and without trigger warnings. It is of the researcher's belief that trigger warnings should be used when content of a sensitive nature is being shared online so that people are afforded the opportunity to disengage with the content if they so wish. The researcher aligns herself with some of the arguments stated in previous literature that trigger warnings are mainly a sign of respect which allows individuals with sensitivities and past traumas to feel safe and included within the community. It is of the researcher's belief that the use of trigger warnings as a sign of respect and inclusion should be upheld on social media platforms. However, the researcher also acknowledges arguments against the use of trigger warnings and feels that they are not a finite solution to combating the effects of sensitive or distressing posts on

social media. Based on previous readings of literature the researcher approaches this study with the assumption that attitudes towards the use of trigger warnings may vary greatly between participants and that new arguments may emerge when discussed within the context of social media.

Participants

A total of 9 participants took part in the study all of which were of Irish nationality. As per Braun and Clarke, (2006; 2022) guidelines the decision to stop recruitment was made once data saturation had been reached, meaning participant answers were becoming repeated and no new themes were emerging from the data set. Participants were divided relatively evenly between males (N=4) and females (N=5). The age range for participants was between 20 and 43 and the mean age of participants was 26.6 (SD=8.4). Participants were required to be over the age of 18 and an active member of at least one social media platform to take part in the study. It should be noted that all participants were known to the researcher prior to taking part in the study which is likely due to the use of convenience sampling and advertisement of the study via personal social media accounts. However, it is not believed this would have had a significant impact on the data other than the possibility that participants better known to the researcher might feel more at ease during the interview and therefore better able to articulate their feelings.

Materials

The following is a detailed account of all materials used throughout the research study.

Electronics and Software

The researcher used google forms to create the information sheet (Appendix A), consent form (Appendix B) and debriefing sheet (Appendix D) provided to the participant throughout the study. Attachments of the google forms were sent via email to participants.

Upon reading the information sheet and filling out the consent form participants consent was documented within the researchers google forms account.

All interview audio was recorded via the built in voice memos app on the researchers personal iPhone 11. The audio recordings were later played via the app into the transcription software Otter.ai so that interviews could be transcribed. Transcription was carried out on the researchers personal laptop; a HP Intel i5.

Interview Agenda

All interviews were conducted in a semi structural conversational style with open ended questions that allowed participants to elaborate on the topic. Demographic questions regarding age, sex and nationality were included to allow a deeper understanding of the sample during data analysis (Appendix C). Additionally, participants were asked to describe what a trigger warning was to establish they understood what it was they were being interviewed about. Questions were based off areas of debate in previous literature regarding the use of trigger warnings in a different setting to social media (George & Hovey, 2019). As these areas of debate were most prominent when discussing trigger warnings in a different setting it is believed that these topics of debate would carry over into a social media context; therefore, targeting areas of interest and obtaining the richest data from participants. It also allows the researcher to get an idea of if the debate is changed in a social media context. Prompts such as “why do you feel that way” or “can you explain that a bit more” were provided by the researcher when deemed appropriate to help obtain rich data. Additionally reassuring feedback such as “perfect thank you” or “I understand” was provided at the end of question answers so that individuals felt more at ease and confident while stating their opinion on the topic.

Procedures

Participants were recruited via convenience sampling. A post to recruit participants was made on the researcher's personal social media accounts of Instagram and Snapchat. The post contained a brief overview of the nature of the study along with the college email of the researcher so that individuals could get in contact and register their interest to take part in the study. Recruitment via social media sites was deemed appropriate as engagement with at least one social media platform is an inclusion criterion for the study. Upon registering their interest to take part in the study participants were sent a detailed information form providing an in depth overview of the study, the use of recording and transcription software and their rights regarding data protection and right to withdraw (Appendix A). Participants were also sent a consent form to confirm they read the information sheet and agreed to take part in the study. Participants were required to confirm they were over the age of 18 and an active user of at least one social media when filling out the consent form (Appendix B). Upon completion of the consent form participants were contacted to arrange a time and place to be interviewed individually. Participants were provided with the option to carry out their interview online or in person and all elected to carry out an in-person interview. Interview locations varied between participants as the researcher allowed the individual to elect a quiet private location of their choosing to allow them to feel more at ease throughout the interview. Five of the interviews took place in an empty room on the campus of the National College of Ireland, one interview took place in an empty room on the campus of University College Dublin, one interview took place in the private home of the participant and one interview took place in the private home of the researcher. An initial pilot study was carried out with the first participant recruited. The participant was informed they were part of the pilot study prior to interview. Upon completion of the interview the participant was asked if there were any questions they felt were not clear or if they had any recommendations for change within

the interview. No recommendations were provided and upon review of the data collected the researcher believed the data collected was rich and thick, as required for thematic analyses. As a result no changes were made to the interview guide and the pilot study interview was included as part of the complete data set for analyses. Each participant was interviewed individually by the researcher in their chosen location over a period of 6 weeks between December and January. Prior to interview interviewees were reminded they could pause or stop the interview at any point and to ask questions if anything regarding the study was unclear. Alongside consent forms additional verbal consent was obtained and recorded prior to interview. Interviews were recorded on the researchers password protected iPhone 11 via the built in voice memos app. Interview times ranged from 6 to 12 minutes long with a mean length of 8.3 minutes. Due to the short interview agenda interviews were relatively short compared to some qualitative studies however, the interview agenda still provided rich data which adequately answered the research aims. Upon completion of the interview all participants were thanked and asked if they had any further questions regarding the study. A debrief form was sent to each participant and the researchers college e-mail was provided in case participants had any concerns after the study. Once interviews had been completed transcription of the audio recordings was carried out using Otter.ai software. Audio recordings were played into the software and transcription was automatically carried out on the researcher's laptop. The researcher then reviewed the transcription and fixed any errors made by the software. In order to abide by ethical guidelines, the audio recordings were then deleted, and all transcription were saved on the researchers password protected laptop.

Data Analysis

A reflexive thematic analyses was carried out using Braun and Clarkes 6 step method to thematic analysis (2006; 2022). A reflexive approach allowed for flexibility throughout the data coding process which was beneficial due to the novel subject matter. Repeated review of

data alongside self-reflection was carried out by the researcher throughout analysis to ensure personal biases were not influencing data analysis. The researcher carried out Braun and Clarkes six steps as follows:

Familiarization of the data set was first carried out through the transcription of audio data. It was at this time that the researcher noticed patterns emerge within the data. The researcher noted that some patterns were consistent with arguments from previous research. The researcher further examined transcriptions and highlighted repeated areas of discussion, from this initial codes were generated. Codes were documented and reviewed by the researcher to attempt to form themes throughout the data set. Initially a total of five themes were generated however upon further review themes were merged to form 3 main themes with subsequent sub themes. Themes were named, explained and discussed within the result and discussion sections.

Ethical Considerations

Approval and Exclusion criteria

The study was reviewed and approved by the National College of Ireland ethics committee and adhered to PSI guidelines. Individuals under the age of 18 were ineligible to take part in the study as they are deemed a vulnerable population and researchers for undergraduate projects would not have the skills or expertise necessary to handle this population in a research setting.

Obtaining and Documenting Consent

Prior to consenting to the study participants were provided with an information sheet via Google Forms which gave a detailed account of the nature of the study, information regarding recording and transcription software, their rights regarding data protection and right to withdraw from the study. Participants were then required to fill out a consent form also

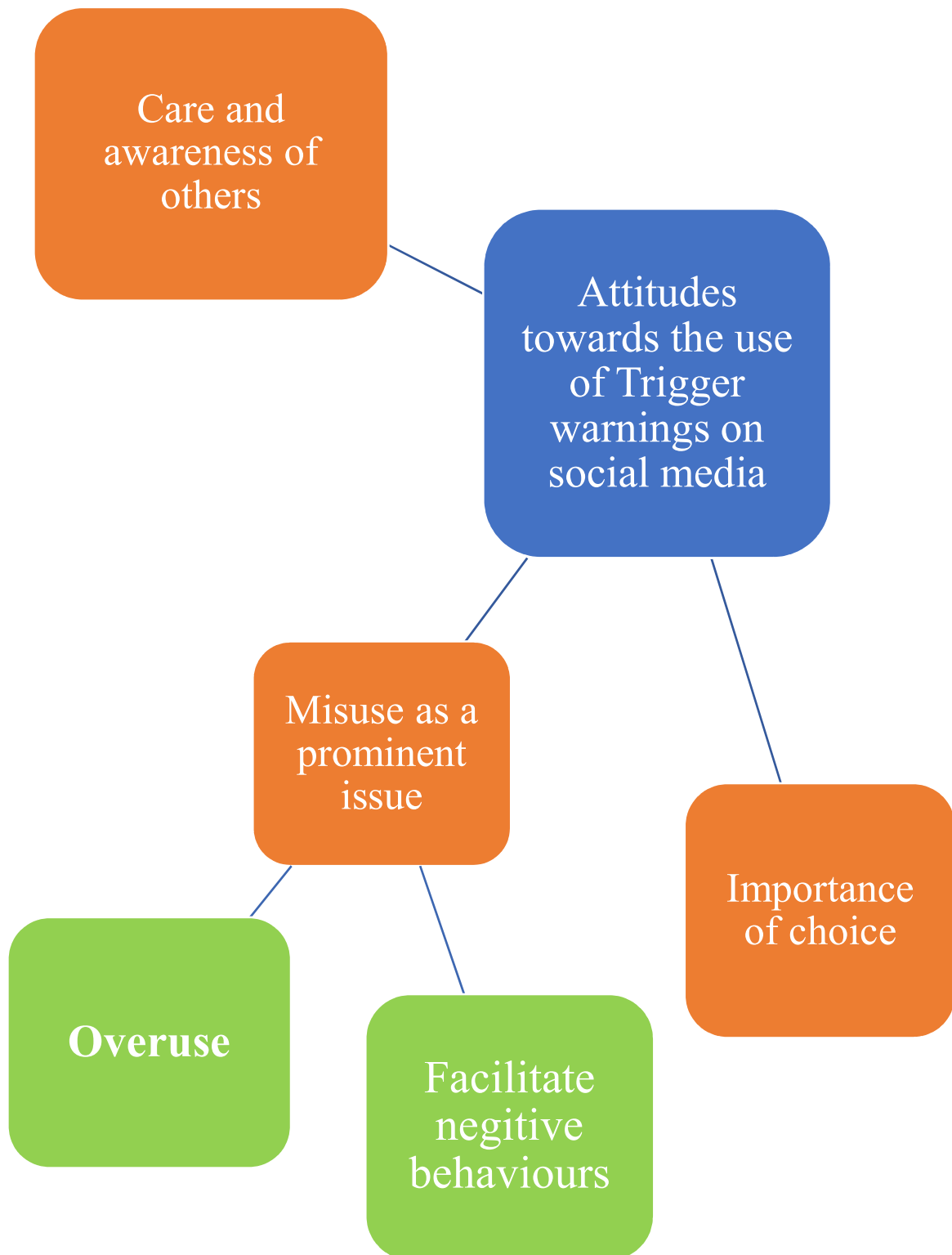
sent via Google Forms and confirm that they read and understood the information sheet, were over the age of 18 and consented to partake in the study. Prior to commencement of the interview the researcher reiterated the participant's right to withdraw if they so wish and obtained further verbal consent for the participants to be recorded and take part in the study. Following the interview participants were e-mailed a debrief document via Google Forms which provided helpful resources regarding data protection and online resources regarding online safety. In addition the researcher's college e-mail was provided for participants to get in contact with any questions or concerns which could arise later.

Data Handling and Protection

Participant interviews were recorded using the voice memos app on the researcher's personal iPhone 11. The iPhone was password protected and remained on the researchers person at all times. Upon completion of the interview the researcher used Otter.ai to transcribe the voice recordings on the researchers laptop which was also password protected. Transcriptions were labelled using a participant number and all potentially identifiable information provided during the interviews were redacted during transcription making data unidentifiable. Upon completion of transcription all audio recordings were deleted from the researcher's phone.

E-mail correspondence also took place between the researcher and participants. This occurred to allow participants to register interest, organise interview time and location and send and receive consent forms. Upon completion of the study all contact information and email correspondence of participants will be deleted by the researcher.

Results



Thematic Map of Results

A thorough examination of the data was carried out through the use of reflexive thematic analyses. Upon following Braun and Clarke's 6 step method of thematic analyses, (2006; 2022), three prominent themes emerged from the data. The themes are as follows (i) Care and Awareness of Others, (ii) Misuse as a Prominent Issue, (iii) Importance of Choice. Themes and subsequent subthemes will be discussed below with reference to interview data.

Care and Awareness of Others

Throughout the interviews almost all participants referenced care and concern for others when discussing the use of trigger warnings. P3 stated *"if I thought it was going to cause any distress to someone I certainly would put a trigger warning"*. Care and consideration for the possibility of causing emotional distress to others appeared to be a motivation for the use of trigger warnings. P9: *"I believe in using them because like there's no point in unnecessarily upsetting somebody you know."*

At times a particular emphasis was placed on protecting those who were emotionally vulnerable.

P2: *"I think like they are beneficial.... They give a warning to people who suffer from the kind of content posted...like if they are in a bad headspace particularly that day they can know to keep scrolling."*

P7: *"For people with mental health issues, eating disorders all that kind of stuff I think like it can really help them so that they don't have to worry about like accidentally coming across things."*

P9: *"I think it is important to acknowledge when something could be potentially upsetting or problematic to a certain type of people"*

Favourable and positive attitudes towards the use of trigger warnings on social media appeared to be associated with an awareness and care for other people's emotions.

Participants seemed conscientious of the fact that certain content could have a negative effect on individuals and indicated that trigger warnings could be used to alert others and prevent the distress.

Participants also acknowledged the uniqueness of others experiences and highlighted the importance of understanding the difference in people's sensitivities when deciding whether to use trigger warnings.

P3: "You don't know the person viewing the information like what effect it could potentially have on the individual because everyone's experiences are so unique...I think it is important to prepare people that they may see something that could cause them harm"

P8: "you never know what could trigger them you know.... You have to kind of treat triggers carefully and be respectful to post the warnings for like others who might need it".

A sense of awareness towards the fact that people's inner emotional experiences differ greatly and a need to protect others online from being impacted negatively seemed to be associated with positive attitudes towards the use of trigger warnings on social media.

Misuse as a Prominent Issue

Several participants spoke of misuse of trigger warnings on social media. These participants felt that this misuse could impact the effectiveness of trigger warnings and even have negative effects on individuals if used incorrectly. *P1: "I definitely think they are overused"*. Overuse was the main way participants felt trigger warnings were misused.

P1: "I once saw someone put up a trigger warning on Tumblr and it.... It was like 'Trigger warning: Cows', for people like if they had a bad experience with cows and I just feel like that is just not necessary"

P6: "Like I've seen people put up, "Trigger warning: Sharp Objects", when there is something sharp in the background and it isn't really the focus of the video and I just feel that its irrelevant".

Participants expressed frustration at trigger warnings being used for seemingly mundane topics on videos that wouldn't usually cause people harm. P8 *"There's a lack of critical thinking with their use sometimes"*. Participant 6 goes on to discuss the impact this could have on the effectiveness of trigger warnings. *".....it can be used so often and it's not actually relevant like people can become desensitised to it"*. The participant believes that if trigger warnings are used too often people may not recognise it as a warning of serious content as people will become accustomed to seeing trigger warnings on mundane content and this could lead to trigger warnings becoming ineffective.

Participant 4 believes that overuse could have a different effect. *"if you put it on nothing over and over again it might make people think that something is more harmful than it actually is"*. Participant 4 feels that trigger warnings when misused could prime people into perceiving content as sensitive or harmful when they may not have viewed it as such without a trigger warning.

Participants also spoke of how they felt some people misuse trigger warnings to engage in negative behaviour.

Participant 1 "I think at times it can actually be very harmful like...some people are posting terrible things and they think like that because they put a trigger warning on it it's okay...they think they can just like trauma dump or put really bad things up if they are using

a trigger warning...I just feel it like gives them an excuse for posting it and kind of perpetuates toxic behaviour on social media.”

This participant speaks of how trigger warnings could potentially be facilitating people putting up distressing content. They express that through using trigger warnings some people may feel they can justify posting “*really bad things*” so long as they have a warning on the content which they feel in turn makes social media less safe. Participant 7 continues this point when asked about their feelings on mandatory trigger warnings they go on to say “*I mean yes and no because like yes I think people need to be warned about something upsetting but also no because I feel that maybe the post should not be made in the first place like maybe if you make mandatory trigger warnings people might think it is okay to post triggering things online when they shouldn’t be in the first place*”.

The participant reiterates the point stated previously that trigger warnings may be used as an excuse to be able to post content on social media that should not be shared. They feel people need to be warned of distressing content but also that this content should not be posted in the first place and that maybe trigger warnings allow people to feel it is acceptable to post this content so long as they preface it with a warning.

Participants also spoke about how they feel trigger warnings could be misused by consumers of the content. They specifically spoke of individuals who may be vulnerable due to mental illness.

P2: “they can act as a kind of lure to pull people in who has the specific trigger... like lets say the post, it’s about self-harm, and a person is struggling with that like a lot of people who struggle with that get drawn into that content and trigger themselves and then go on and selfharm and I feel like by putting a trigger warning on the post it might draw their attention to it they click on it to trigger themselves and it does them more harm in the end

because they might have glossed over it if there wasn't a trigger warning to draw attention to it"

The participant feels that trigger warnings can almost attract people to negative content. They say, *"humans we are curious"*, and by seeing a trigger warning the person is more likely to engage with content they would have ignored. They feel people who are at more risk of hurting themselves as a result of seeing distressing content are particularly vulnerable as they may not have noticed if there was not a trigger warning. Participant 7 continues to speak of misuse of trigger warnings by people with eating disorders.

"I feel like some people especially with eating disorders, they go out looking for things with trigger warnings....like the hashtag trigger warnings and stuff because they get ideas and like motivation from it and it kind of feeds their eating disorder kind of similar to pro ana sites and stuff but recently it is really bad like with Instagram and Tik Tok I feel"

The participant feels that some people with eating disorders use trigger warnings as a way of facilitating disordered behaviour. They discuss how by placing a trigger warning hashtag or notice on these posts it allows people to simply look up trigger warnings as a hashtag and find a wealth of triggering content that feeds their eating disorder. They even speak of people getting *"ideas"* of disordered behaviours from these videos which could worsen the effects of their eating disorder.

Importance of Choice

The concept of choice was a repetitive and predominant theme throughout each interview. However the context in which people spoke of choice differed between participants. Some participants felt that trigger warnings were important for people to be able to choose whether they engage with content or not.

P9: "I think we need trigger warnings for people who don't want to view certain things so they know and they are able to have the choice about whether or not to view that content like they can decide that they want to see it or not"

P3: "I think it's important that somebody can have the choice to say no I'm not going to watch it or maybe I will it...(trigger warnings)...gives us the choice really"

Some showed positive attitudes towards the use of trigger warnings on social media as they felt it provided people with an opportunity to avoid unwanted content. Participant 7 says *"its good to have them so if someone is trying to avoid those pictures or videos or anything you have it there to warn them"* They speak of how without the trigger warning people may become *"blindsided"* by content they do not expect and trigger warnings allow people to stop and choose whether they want to see a post or not.

P6: "Having the warning there so that people know if there is something that could possibly trigger them they have the option to not watch it"

Trigger warnings were perceived as a way in which choice of engagement with content can be facilitated and without it people may see content they do not wish to against their will.

However, other participants spoke of the content posters right of choice when it came to deciding whether they wanted to use trigger warnings or not. When asked about their opinions on mandatory trigger warnings they were opposed to the idea.

P2: "People have a right to express themselves and if you say to them they have to use trigger warnings then you are kind of curtailing their freedom of speech"

P4: "I'd be very much freedom of speech I think people should have their own autonomy to make these decisions"

P5: "I don't think we should be telling people what to do like leave it up to themselves, I think they should use it but I also don't think we should force them"

These participants believed that the use of trigger warnings had to be down to the content user's discretion. Some speak of how forcing trigger warnings would have an effect on their "freedom" and limits people's ability to "express" how they feel. One participant spoke about the fact that even if they personally felt a trigger warning should be used on sensitive posts they don't think that forcing people to use them is right.

Discussion

This research study aimed to gain a better understanding of the perception and attitudes towards the use of trigger warnings on social media. Possible motivations for the use of trigger warnings on social media, perceived benefits and drawbacks of their use and attitudes towards regulations regarding trigger warnings on social media were areas of particular interest within the study. An examination of the interviews through reflexive thematic analysis uncovered the three following themes, (i) Care and awareness of others, (ii) Misuse as a Prominent Issue and (3) Importance of Choice.

The theme "Care and Awareness of Others" demonstrated positive attitudes towards the use of trigger warnings on social media. Participants expressed an understanding of the impact negative content on social media could have on a person. They discussed a need to protect others from emotional distress and how trigger warnings provided a way in which they could facilitate this. Care and consideration for others' emotions appeared to be a motivation for the use of trigger warnings amongst participants. The study findings echo previous reasoning for their use in educational settings this (Dickman-Burnett, 2019; George & Hovey, 2019). Participants who spoke favourably regarding the use of trigger warnings appeared particularly cognisant of the uniqueness of people's experiences and how different

topics could have an effect on different individuals; particular consideration for individuals suffering from mental illness was expressed. Knowledge and recognition of vulnerable populations also frequently arose within previous literature and contributed to debates supporting the use of trigger warnings with a specific emphasis on protecting those with past traumas. As a result of limited research surrounding the use of trigger warnings these papers were also focused on the use of trigger warning within an educational setting. It appears having an awareness and care for others is associated with positive attitudes towards trigger warnings and are a motivation for their use regardless of the environment in which they are used. It should be noted that due to the use of convenience sampling and the area of study of the researcher over half of participants were psychology students. This may explain the depth of knowledge, interest and awareness of others emotional vulnerabilities expressed amongst participants. However, this theme was also present in interviews of participants who were not students of psychology indicating that this is not the sole reason this attitude emerged from the data.

The theme “Misuse as a Prominent Issue” highlights the concerns participants conveyed relating to perceived inappropriate use of trigger warnings. Multiple participants referenced times in which they had seen posts on social media that contained trigger warnings which they felt were unnecessary or irrelevant. Participants spoke of frustration and annoyance regarding the overuse of trigger warnings within social media on seemingly trivial content. They discussed how this overuse could undermine the seriousness of trigger warnings and effect their potency. A previous systematic review aiming to better understand the different typology of trigger warnings also made reference to this issue. The review discussed a lack of continuity in relation to the use of trigger warnings across different types of content and a lack of clarity regarding their appropriate use (Charles et al., 2022). Future research should be carried out to investigate whether the use of trigger warnings on mundane

posts on social media impacts their perceived seriousness and potency on users. If this is the case, clear guidelines may need to be established to protect their effectiveness.

One participant discussed a further implication of the overuse of trigger warnings. The participant discussed how they felt that if trigger warnings were used repeatedly on topics that were not necessarily sensitive in nature social media users may begin to perceive that content as distressing. A previous research paper regarding the use of trigger warnings in the classroom argued that by precursing content with a trigger warning you automatically cause a “negative relationship” between the individual and the content (Carniel, 2017). This idea that trigger warnings may prime an individual to negatively perceive content has been further studied through experimental studies. These studies found that participants who were shown trigger warnings before engagement with content showed a greater anxiety response prior to and when exposed to said content than control groups (Bellet et al., 2018). However, replications of these findings have been ambiguous (Jones et al., 2020; Sanson et al., 2019) and further research is needed in the area to establish these results.

Multiple participants also expressed the belief that some individuals misuse trigger warnings to facilitate and engage in problematic behaviours. A number of participants spoke about how they felt some people would use trigger warnings in order to justify posting distressing content on social media. They expressed that trigger warnings provided a way in which individuals could condone their actions and place the responsibility of engaging with content on the user. Participants felt this negation of responsibility was wrong and people should be held accountable for their own posts. Participants highlighted an avenue of study that has not yet been explored surrounding trigger warnings on social media. We know from previous research that exposure to distressing content has harmful effects on individuals. In extreme cases with vulnerable populations this exposure can even lead to self-harm and injury (Duggan et al., 2011). If participants theories are correct and trigger warnings are used

as a justification for posting distressing content it is possible that trigger warnings are causing more harm than good. Further studies should investigate whether posters of distressing content on social media feel justification of their actions through the use of trigger warnings and whether this correlates to more frequent posting of distressing content.

Additionally, some participants expressed their belief that members of vulnerable populations, specifically those suffering from a mental illness, may use trigger warnings to facilitate negative behaviour. They portrayed unfavourable attitudes towards the use of trigger warnings in this manner and felt that at times trigger warnings negatively impacted and even worsened disordered behaviours; thus worsening the symptoms and impact of their mental illness. Previous research mainly discussed this argument in relation to sufferers with PTSD and how trigger warnings encouraged the problematic avoidance behaviours associated with PTSD (Jones et al., 2020). However, participants of this study spoke of other possible mental health disorders and behaviours that could be facilitated through trigger warnings. Specifically participants spoke of self-harm engagement and eating disorder behaviours. They felt that some individuals would use trigger warnings as a form of search word that allowed them to quickly and easily access content which would expose them to triggers. Previous research has explored this concept and found that self-triggering is a problematic area of concern amongst individuals who self-harm and struggle with eating disorders (Bellet et al., 2020; Chancellor et al., 2016) . One participant even spoke of how they believed trigger warnings were a form of lore that almost attracted people to the content as a result of a natural curiosity within human nature. This participant outlined an important area of future research into whether people are more likely to look at distressing due to trigger warnings. If this is found to be the case, it is possible that trigger warnings are resulting in more harm than good for individuals on social media.

It became evident throughout analysing interviews that the “Importance of Choice” was a shared attitude amongst all participants. The majority of participants spoke discussed how they believed trigger warnings provided time for social media users to stop and choose whether or not they wanted to see certain content. They discussed how without trigger individuals may come across sensitive content suddenly and accidentally. Most participants felt this was morally wrong and people should be provided with an opportunity to decide for themselves whether they wanted to see distressing content. This concept of providing individuals with the agency to make their own decisions regarding engagement with content has been a recurring argument within literature for the use of trigger warnings within educational settings (Dickman-Burnett, 2019; George & Hovey, 2019).

Notably however, common arguments against the use trigger warnings within educational setting were also repeated in the context of their use in social media. These arguments focused mainly around the concept of mandatory trigger warnings which had been enforced in different setting previously. Participants reiterated the belief in freedom of speech, a common argument against the enforcement of trigger warnings (DeGroot & Carmack, 2020; Hume, 2016) . They described a need for people to have the ability to express themselves however they see fit and that an enforcement of trigger warnings prevents them from doing so. The current study suggests that argument for and against the use of trigger warnings which have only been discussed in relation to educational settings transfer over to their use other platforms such as social media. This has allowed us to obtain a clearer perception of their use within the broader community and may influence regulations regarding enforcement of their use in the future.

Additional Limitations, Strengths and Implications of Research

One of the limitations of this study is its lack of ethnic diversity within the sample due to the fact that all participants were people of Irish nationality. This could have an impact on the samples attitudes towards the use of trigger warnings on social media. Some experts argue that trigger warnings can block or inhibit content that's enables narrative surrounding important issues faced by people in society such as racism and prejudice. In 2020 a video was shared which depicted the murder of George Floyd, a black man from the United States, by local police officers (Hill et al., 2020). Although the video was inarguably a harrowing and disturbing watch it highlighted the ongoing racism and police brutality within the United States. This subsequently spurred conversations and prompted action against racism across the globe. There is some who would debate that feelings of distress and discomfort are needed to move forward in society. It is argued that trigger warnings could be used as a way for people to avoid topics that make them uncomfortable (Kyrölä, 2018). It is possible due to a lack of minority groups and racial diversity within the sample this perspective was missed as racial prejudice is not usually an issue faced by white Irish nationals. Further studies with a more diverse sample should be conducted to investigate whether different arguments for and against trigger warnings emerge due to a person's personal experience of racial issues.

It is also possible that lack of experience on the part of the researcher may have impacted the quality of data collected. As an undergraduate student this study was the first time the participant conducted their own research involving qualitative study. The researcher noticed throughout the process of interviewing that they were more confident in asking participants to elaborate and explain their answers in more depth. This could have impacted the quality of data obtained from initial interviews. However, it should also be noted that earlier interviewees appeared more confident and fluid in their speaking and therefore little

prompting was necessary. It is of the researcher's opinion that each interview managed to obtain rich and thick data despite early inexperience and confidence on the researcher's part.

Due to the qualitative design of the study the researcher was able to obtain a detailed account of the participant's perceptions and attitudes towards trigger warnings in social media. The conversational and flexible style of qualitative research facilitated investigation into the "how" and "why" these perceptions were held amongst participants by allowing the researcher to prompt participants into further elaborating on the reasoning behind their perceptions. As a result, a broader spec of analyses and understanding of the data was enabled.

The study has contributed to the extremely limited research regarding the use of trigger warnings on social media. Previous research has predominantly focused on the opinions and impacts of trigger warnings in an educational setting, despite their use across a variety of environments and mediums. The research study uncovered topics of concern and interest regarding the use of trigger warnings in social media amongst participants via three identified themes; (i) Care and Awareness of Others , (ii) Misuse as a Prominent issue and (iii) Importance of Choice. These topics of concern and interest form a basis for areas of future research regarding the use of trigger warnings in social media.

References

- American Psychological Association. (n.d.). *Apa Dictionary of Psychology*. American Psychological Association. Retrieved February 11, 2023, from <https://dictionary.apa.org/trigger>
- 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. <https://doi.org/10.1016/j.jbtep.2018.07.002>
- Bellet, B. W., Jones, P. J., & McNally, R. J. (2020). Self-triggering? an exploration of individuals who seek reminders of trauma. *Clinical Psychological Science*, *8*(4), 739–755. <https://doi.org/10.1177/2167702620917459>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE Publications.
- Bridgland, V. M., Green, D. M., Oulton, J. M., & Takarangi, M. K. (2019). Expecting the worst: Investigating the effects of trigger warnings on reactions to ambiguously themed photos. *Journal of Experimental Psychology: Applied*, *25*(4), 602–617. <https://doi.org/10.1037/xap0000215>
- Carniel, J. (2017). [insert image here]: A reflection on the ethics of imagery in a critical pedagogy for the humanities. *Pedagogy, Culture & Society*, *26*(1), 141–155. <https://doi.org/10.1080/14681366.2017.1364784>

- Chancellor, S., Lin, Z., Goodman, E. L., Zerwas, S., & De Choudhury, M. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. <https://doi.org/10.1145/2818048.2819973>
- Charles, A., Hare-Duke, L., Nudds, H., Franklin, D., Llewellyn-Beardsley, J., Rennick-Egglestone, S., Gust, O., Ng, F., Evans, E., Knox, E., Townsend, E., Yeo, C., & Slade, M. (2022). Typology of content warnings and trigger warnings: Systematic review. *PLOS ONE*, *17*(5). <https://doi.org/10.1371/journal.pone.0266722>
- DeGroot, J. M., & Carmack, H. J. (2020). Unexpected negative participant responses and researcher safety: “Fuck your survey and your safe space, trigger warning bullshit.” *Journal of Communication Inquiry*, *44*(4), 354–375. <https://doi.org/10.1177/0196859920921752>
- Dickman-Burnett, V. L., & Geaman, M. (2019). Untangling the trigger-warning debate. *Journal of thought*, *53*(3/4), 35-52.
- Duggan, J. M., Heath, N. L., Lewis, S. P., & Baxter, A. L. (2011). An examination of the scope and nature of non-suicidal self-injury online activities: Implications for School Mental Health Professionals. *School Mental Health*, *4*(1), 56–67. <https://doi.org/10.1007/s12310-011-9065-6>
- George, E., & Hovey, A. (2019). Deciphering the trigger warning debate: A qualitative analysis of online comments. *Teaching in Higher Education*, *25*(7), 825–841. <https://doi.org/10.1080/13562517.2019.1603142>

- Halberstam, J. (2017). Trigger happy: From content warning to censorship. *Signs: Journal of Women in Culture and Society*, 42(2), 535–542. <https://doi.org/10.1086/688189>
- Hill, E., Tiefenthäler, A., Triebert, C., Jordan, D., Willis, H., & Stein, R. (2020, June 1). *How George Floyd was killed in police custody*. The New York Times. Retrieved March 13, 2023, from <https://www.nytimes.com/2020/05/31/us/george-floyd-investigation.html>
- Hume, M. (2016). *Trigger warning: Is the fear of being offensive killing free speech?* Harpercollins Publishing Aust.
- Jones, P. J., Bellet, B. W., & McNally, R. J. (2020). Helping or harming? the effect of trigger warnings on individuals with trauma histories. *Clinical Psychological Science*, 8(5), 905–917. <https://doi.org/10.1177/2167702620921341>
- Kyrölä, K. (2018). Negotiating vulnerability in the trigger warning debates. *The Power of Vulnerability*. <https://doi.org/10.7765/9781526133113.00007>
- McDonald, A. D., Sasangohar, F., Jatav, A., & Rao, A. H. (2019). Continuous monitoring and detection of post-traumatic stress disorder (PTSD) triggers among veterans: A supervised machine learning approach. *IISE Transactions on Healthcare Systems Engineering*, 9(3), 201–211. <https://doi.org/10.1080/24725579.2019.1583703>
- Radovic, A., Gmelin, T., Stein, B. D., & Miller, E. (2016). Depressed adolescents' positive and negative use of social media. *Journal of Adolescence*, 55(1), 5–15. <https://doi.org/10.1016/j.adolescence.2016.12.002>
- Regehr, G. (2010). It's not rocket science: Rethinking our metaphors for research in health professions education. *Medical Education*, 44(1), 31–39. <https://doi.org/10.1111/j.1365-2923.2009.03418.x>

Sanson, M., Strange, D., & Garry, M. (2019). Trigger warnings are trivially helpful at reducing negative affect, intrusive thoughts, and avoidance. *Clinical Psychological Science*, 7(4), 778–793. <https://doi.org/10.1177/2167702619827018>

Social Media Overview. Communications. (2021, March 5). Retrieved March 13, 2023, from <https://communications.tufts.edu/marketing-and-branding/social-media-overview/#:~:text=What%20is%20Social%20Media%3F,Instagram%2C%20LinkedIn%20and%20YouTube%20accounts>.

Stratta, M., Park, J., & deNicola, C. (2020). Automated content warnings for sensitive posts. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3334480.3383029>

Wyatt, W. (2016). The ethics of trigger warnings. *Teaching Ethics*, 16(1), 17–35. <https://doi.org/10.5840/tej201632427>

Appendix A

Participant Information Sheet

Thank you for registering your interest to take part in my final year research project. Please read the following information sheet carefully before you consent to taking part in my study.

The study will be investigating individuals attitudes towards the use of trigger warnings on social media in Ireland. The research will be carried out through one on one interviews which can be conducted either in person or via video call on Microsoft Teams depending on the individuals personal preference.

Participants should be aware that the audio of the interviews will be recorded and transcribed so that I, the researcher, can conduct a thorough analyses once all interviews have been completed. These recordings will be recorded using either Microsoft Teams or a mobile voice recording app. Transcriptions will be carried out using the online transcription software Otter.ai. Recordings and transcriptions will be protected on a password protected laptop and/or phone that will only be accessible by the researcher. All data will be stored and processed in accordance with GDPR guidelines. Participants may withdraw from the study prior, during or after the interview if they so wish. Unfortunately due to the nature of this study participants will be unable to withdraw their interview from the data once their interview has been reviewed and transcribed. This is because upon completion of transcription audio recordings will be deleted and all transcriptions will be stored unidentifiably. Prior to this, interviews can be removed from the data set. Upon completion of the research study all audio recordings, email correspondents and transcriptions will be deleted by the researcher.

Overall findings of the research project will be written up and discussed as a whole however, some individual quotes may be used to demonstrate specific themes within the research. The research study will also be presented within NCI by the researcher as part of their final year project. There will be no questions that could cause you to be identifiable within the interview. Due to the nature of the research subject there is a chance questions may cause individuals to be reminded of upsetting content they have seen on social media.

Although graphic or sensitive content will not be presented to participants discussion of such content may cause distress to some individuals. It is advised that any individual who believes they may be particularly sensitive to the research subject should not take part in this study.

In addition, no person under the age of 18 will be eligible for this study.

Please click and read the following links to receive all information in regards to your rights of data protection confidentiality and freedom of information.

[Data Protection Act 2018 \(irishstatutebook.ie\)](https://www.irishstatutebook.ie/eli/2018/act/23/enacted/en/html)

[Data Protection Act 2018 \(irishstatutebook.ie\)](https://www.irishstatutebook.ie/eli/2018/act/23/enacted/en/html)

I have read and understand the information sheet provided



Appendix B

Consent form

The following document is a consent form to partake the final year project of National College of Ireland Student Orla Hyland student no.x19756151. Please ensure to read the information sheet provided before filling out the consent form. Please note that through filling out the following consent form you are not obligated to take part in the research study. You may withdraw from the study at any point prior to the of the transcription of your interview

I can confirm that I have read and understood the information sheet relating to the details of this study.

I understand my rights to withdraw from the study prior to the transcription of my interview.

I can confirm I am over the age of 18 and use one form of social media.

I have read and understood the information relating to the recording of my interview and I consent to its recording.

Appendix C

Interview Agenda.

- What is your age?
- What is your identifying gender?
- Are you working/studying? If so what area are you working/studying in?
- Can you describe in your own words what you believe a trigger warning to be?
- Do you see trigger warnings often on social media?
- Have you ever used a trigger warning in your own posts on social media? If so why did you choose to use the trigger warning?
- Do you feel trigger warnings are beneficial in anyway? Why?
- Do you feel trigger warnings are harmful in anyway? Why?
- Do you feel trigger warnings are used enough, too much or just the right amount on social media?
- Do you feel trigger warnings should be made mandatory on sensitive posts on social media? Why?
- Do you have any other views or opinions you would like to share about the use of trigger warnings on social media that you have not already spoken about?

Appendix D

Debrief Sheet

Thank you for taking the time to be a participant of my study. The following sheet has been provided to answer any questions you may have regarding the study. It will also provide resources to contact in case you have been upset by any topics brought up during the study. If you have any further questions or concerns that have not be answered, please contact me at x19756151@student.ncirl.ie

Please remember that you may have your data withdrawn from the study prior to the transcription of your interview at which point it will not be identifiable. If you wish to withdraw, please contact me at the email provided above. A reminder that personal data such as emails or correspondence will be deleted upon completion of the study. Any audio or transcriptions of interviews will be deleted also. Any data reported or published will refer to the data set provided by the interviews as a whole; no references to individual interviews will be written up. If you have become upset due to any topics brought up by the research project please see the resources below.

[Mental health - HSE.ie](https://www.hse.ie/eng/health/mental_health/)

[gov.ie - Be Safe Online \(www.gov.ie\)](https://www.gov.ie/en/about-us/2017-03-14-be-safe-online/)

[Homepage | Data Protection Commission](https://www.data-protection-commission.ie/)