

# **Examining the Self-harm and Suicide Contagion Effects Related to the Portrayal of the Blue Whale Challenge on YouTube and Twitter**

Amro Khasawneh, Kapil Chalil Madathil, Emma Dixon, Pamela Wiśniewski, Heidi Zinzow, Rebecca Roth

Submitted to: JMIR Mental Health  
on: August 22, 2019

**Disclaimer:** © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

## ***Table of Contents***

---

**Original Manuscript..... 4**



# Examining the Self-harm and Suicide Contagion Effects Related to the Portrayal of the Blue Whale Challenge on YouTube and Twitter

Amro KhasawnehMSc, ; Kapil Chalil MadathilPhD, ; Emma DixonBSc, ; Pamela Wi?niewskiPhD, ; Heidi ZinzowPhD, ; Rebecca Roth

## Corresponding Author:

Kapil Chalil MadathilPhD

Phone: +1864-656-0856

Fax: 1 864-656-0795

Email: kmadath@clemson.edu

## Abstract

**Background:** Research suggests that direct exposure to suicidal behavior and acts of self-harm through social media may increase suicidality through imitation and modeling, with adolescents representing a particularly vulnerable population. One example of viral self-harming behavior that could potentially be propagated through social media is the Blue Whale Challenge (BWC).

**Objective:** We investigate how people portray BWC on social media and the potential harm this may pose to vulnerable populations.

**Methods:** We first used a grounded approach coding 60 publicly posted YouTube videos, 1112 comments on those videos, and 150 Twitter posts that explicitly referenced BWC. We deductively coded the YouTube videos based on the Suicide Prevention Resource Center (SPRC) Messaging guidelines.

**Results:** Overall, 83.33%, 28.33%, and 68.67% of the YouTube videos, comments, and Twitter posts were trying to raise awareness and discourage participation in BWC. Yet, about 37% of the videos violated six or more of the SPRC messaging guidelines.

**Conclusions:** These posts might have the problematic effect of normalizing BWC through repeated exposure, modeling, and reinforcement of self-harming and suicidal behavior, especially among vulnerable adolescents. Greater efforts are needed to educate social media users and content generators on safe messaging guidelines and factors that encourage versus discourage contagion effects.

(JMIR Preprints 22/08/2019:15973)

DOI: <https://doi.org/10.2196/preprints.15973>

## Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

**Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

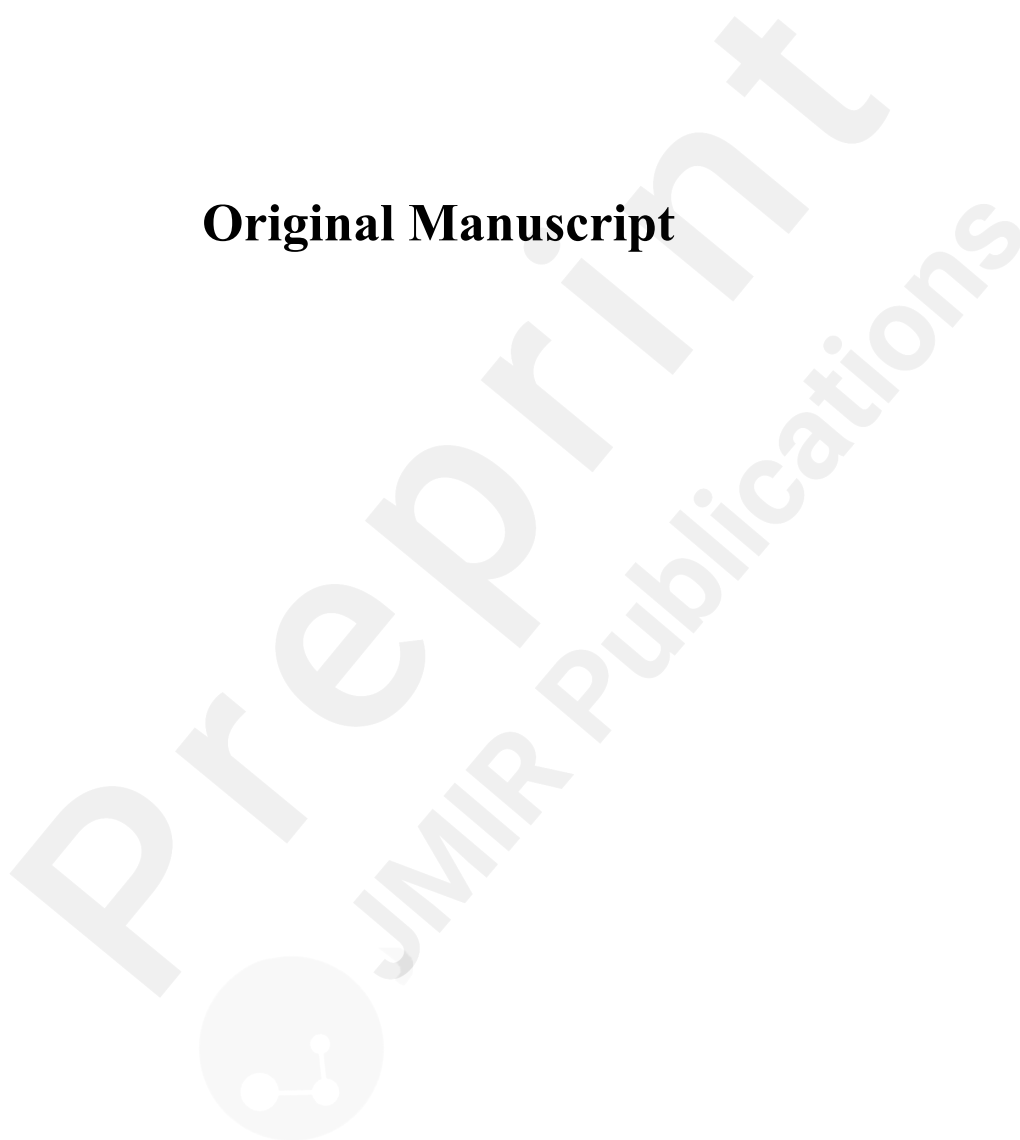
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

**Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in [http](#)

**Original Manuscript**



## Original paper

Amro Khasawneh<sup>1</sup>, MS; Kapil Chalil Madathil<sup>1,2</sup>; Emma Dixon<sup>3</sup>, BS; Pamela Wiśniewski<sup>4</sup>, PhD; Heidi M. Zinzow<sup>5</sup>, PhD; Rebecca Roth<sup>5</sup>

<sup>1</sup>Clemson University, Department of Industrial Engineering, Clemson, SC, United States

<sup>2</sup>Clemson University, Glenn Department of Civil Engineering, Clemson, SC, United States

<sup>3</sup>University of Maryland, College Park, College of Information Studies, College Park, MD, United States

<sup>4</sup>University of Central Florida, Department of Computer Science, Orlando, FL United States <sup>5</sup>Clemson University, Department of Psychology, Clemson, SC United States

Corresponding Author:

Kapil Chalil Madathil, PhD

Assistant Professor

Departments of Civil and Industrial Engineering

S Palmetto Blvd, Clemson, SC 29634

Clemson University

kmadath@clemson.edu

Phone: (713) 294-6499

FAX: (864) 656-0795

## Examining the Self-harm and Suicide Contagion Effects Related to the Portrayal of Blue Whale Challenge on YouTube and Twitter

### Abstract

**Background:** Research suggests that direct exposure to suicidal behavior and acts of self-harm through social media may increase suicidality through imitation and modeling, with adolescents representing a particularly vulnerable population. One example of viral self-harming behavior that could potentially be propagated through social media is the Blue Whale Challenge (BWC).

**Objective:** We investigate how people portray BWC on social media and the potential harm this may pose to vulnerable populations.

**Methods:** We first used a grounded approach coding 60 publicly posted YouTube videos, 1112 comments on those videos, and 150 Twitter posts that explicitly referenced BWC. We deductively coded the YouTube videos based on the Suicide Prevention Resource Center (SPRC) Messaging guidelines.

**Results:** Overall, 83.33%, 28.33%, and 68.67% of the YouTube videos, comments, and Twitter posts were trying to raise awareness and discourage participation in BWC. Yet, about 37% of the videos violated six or more of the SPRC messaging guidelines.

**Conclusions:** These posts might have the problematic effect of normalizing BWC through repeated exposure, modeling, and reinforcement of self-harming and suicidal behavior, especially among vulnerable adolescents. Greater efforts are needed to educate social media users and content generators on safe messaging guidelines and factors that encourage versus discourage contagion effects.

**Keywords:** Social media challenges; Blue Whale Challenge; Self-harm; Digital self-harm; Suicide;

Mental health and wellbeing; Problematic internet behaviors

## Introduction

Adolescents and young adults are the largest population using the internet, as it has become an essential instrument for their schoolwork, information collection, and socializing [1-4]. Communication tools on the internet, including emails, direct texting, and blogging have become fundamental in adolescents' social development [5,6]. More recently, online social networks, also known as social media, such as Facebook, Instagram, Twitter, and YouTube have become more popular and common station among adolescents where they can develop public accounts or profiles to connect with other individuals and see their list of connections and posts [5-7].

While social media has created a number of opportunities for individuals to garner social support online [1,2,8-10], it also has the potential to further marginalize particularly vulnerable individuals. Recent studies have highlighted how social media can be used to harass, discriminate [11,12], dox [13], and socially disenfranchise those individuals [14,15,16]. Some research even suggests that social media use may be a contributing factor to the significant increase in suicide rates and depressive symptoms among adolescents and young adults in the past decade [1,17,18]. Evidence suggests that suicidal behavior can be propagated through social contagion effects, which model, normalize, and reinforce self-harming behavior [19-21]. These harmful behaviors and social contagion effects may occur more frequently through repetitive exposure and modelling via social media, especially when such content goes "viral" [19,20,22]. Such widespread exposure to harmful or suggestive content is particularly detrimental to vulnerable individuals, including adolescents and young adults [23,24].

Most of the literature regarding adolescent online safety focuses on sexual or aggressive behaviors that put youth at risk [25]. Unfortunately, less research has examined how social media can influence adolescents and young adults to engage in self-harming behavior [20,26,27]. Thus, interdisciplinary researchers seek to shed light on the urgent need to form a cohesive research agenda around digital self-harm [26], online non-suicidal self-injury [28], the use of social media to discuss deliberate acts of self-harm, and acts of cyber-suicide [29,30]. Most of this work, however, is in its early stages with few drawings upon the developmental aspect of vulnerable adolescents' social media use.

One example of viral self-harming behavior that has generated significant media attention is the Blue Whale Challenge (BWC). Allegedly in this challenge, adolescents and young adults are encouraged to engage in self-harm and eventually kill themselves [31]. Even though there is a possibility that BWC might be a hoax, we believe its portrayal on social media (or any similar phenomenon) still has an impact on these vulnerable individuals. Research is needed specifically regarding BWC ethical concerns, the effects the game may have on adolescents, and potential governmental interventions. To address this gap in the literature, the current study uses qualitative and content analysis research techniques to illustrate the risk of self-harm and suicide contagion through the portrayal of BWC on YouTube and Twitter Posts. The purpose of this study is to analyze the portrayal of BWC on YouTube and Twitter to answer the following questions:

RQ1: What themes are present on YouTube and Twitter posts that share and discuss BWC?

RQ2: To what extent are YouTube videos compliant with safe and effective suicide messaging guidelines?

## Methods

In this study we selected two social media platforms for data collection, YouTube and Twitter. We identified the common themes/categories of YouTube videos, comments on those videos, and Twitter posts through conducting grounded, thematic content analysis on data extracted from these platforms [32]. Due to the different policies, technical affordances, and data types of these platforms [26,33,34], we decided to build different codebooks for each type of data, analyze the data for each site, and then compare the codebooks across platforms. We deductively coded the YouTube videos based on the Suicide Prevention Resource Center (SPRC) Safe and Effective messaging guidelines to explore how many of these videos violated these guidelines, the extent to which they violated guidelines, and which guidelines tend to be violated the most frequently. SPRC developed nine Safe and Effective Messaging guidelines based on best practices from research to reduce the risk of inducing self-harm or suicide-related behavior in those who view a message, see (Table 1) [35–37]. Either following or violating these guidelines represent a metric for the contagion risk associated with any social media post.

Table 1. SPRC Safe and Effective Messaging for Suicide Prevention [36].

Guideline	Description
<b>Emphasize seeking help and provide information on where to find it.</b>	Provide steps on finding mental health treatment. Advise that help is available through local service providers and crisis centers, and through the National Suicide Prevention Lifeline (1-800-273-TALK [8255]).
<b>Emphasize prevention.</b>	Highlight that suicide is avoidable and preventable and that there are actions for individuals who have suicidal thoughts to prevent them from acting on those thoughts [58].
<b>List the warning signs, as well as risk and protective factors of suicide.</b>	List the warning signs like the ones developed by the American Association of Suicidology (AAS). List what could reduce risk of suicide and what could increase the risk of suicide; these can be found on pages 35-36 in the National Strategy for Suicide Prevention. Educate people how to identify a person with self-harming thoughts.
<b>Highlight effective treatments for underlying mental health problems.</b>	46% of people who died by suicide from 2014-2016 had a known mental health condition, with 67% of these decedents having a history of substance use treatment. Among the descendants, 46.7% had recently been released from a psychiatric facility [59], a statistic which can be reduced by having access to effective treatments and better social support [60].
<b>Avoid glorifying or romanticizing suicide or people who died by suicide.</b>	Vulnerable younger adults may relate to the attention given to and sympathy for a person who died by suicide [47].
<b>Avoid normalizing suicide by presenting it as a common event.</b>	Suicide ideation is not normal to most people, and they do not consider it an option. However, presenting suicide as a common event may remove this bias [50].

<b>Avoid presenting suicide as an inexplicable act or explain it as a result of stress only.</b>	Doing so may encourage identification with the victim as well as neglecting the complexity and preventability of suicide. Presenting suicide as an explainable or a result of stress only misleads vulnerable individuals to believe that it is a normal response to common life situation [36,48,50–52].
<b>Avoid focusing on personal details of people who died by suicide.</b>	Vulnerable younger adults may feel they are like with the person who died by suicide, eventually leading them to consider taking their lives in the same way [48].
<b>Avoid presenting overly detailed descriptions of suicide victims or methods of suicide.</b>	Including pictures and/or descriptions of where and how an individual died by suicide may lead a vulnerable person to imitate the act [48,51].

## Data Collection

We chose YouTube and Twitter for two reasons: both are ranked among the top 20 most popular social media sites and the posts on these platforms and their posts are normally open to the public [33,34,38]. The videos were collected using the YouTube search engine using “Blue Whale Challenge” as key words, then we sorted the results by relevance to collect information from the first 60 videos on the list, which combined, has a length of ~12 hours. The process of collecting and coding the data was iterative to assure we achieve data saturation [39]. Interestingly, only 5 out of these 60 videos required age verification. The following information was collected for each video: the link, number of views, and the first 30 comments sorted by “Top comments” if there were any, since this was when data saturation was achieved as found by iteratively collecting and coding data [39]. A total of 1,112 comments were collected for coding. Inclusion criteria for the videos were relation to the BWC and in English, translated into English, or contained English subtitles. Inclusion criteria for the comments were the comment had to be in English and include words. This kind of data collection strategy was chosen to mimic the typical user behavior search strategy [40–42]. One hundred and fifty Twitter posts were randomly collected using the social media analytics tool Radian6 [43] from the period between February 2012 - February 2018 using keywords and inclusion criteria provided in (Table 2). We iteratively collected and coded the Twitter posts as well starting with 100 posts, and ended up with 150 posts. This iterative process was used to assure data saturation [39], and the timeline for collecting the data was chosen to cover the time period where it is believed that BWC was most active [44].

Table 2. Keywords, Hashtags, and Inclusion Criteria for Scraping Twitter Posts using Radian6.

Keywords	Hashtags	Inclusion criteria
“Blue Whale Challenge”	“#BlueWhaleChallenge”	The post had to be in English and related to the BWC. No duplicates
“Blue Whale Game”	“#BlueWhaleGame”	
“BWC”	“#Blue_Whale_Challenge”	
	“#Blue_Whale_Game”	
	“#BWC”	

## Data Analysis Approach

To identify the common themes and categories of the YouTube videos, YouTube comments, and Twitter posts, codebooks were developed by dividing the top 40 videos, the top 40 videos’ related comments, and 100 Twitter posts between two raters. Each rater then developed a codebook



before convening to discuss the common themes/categories found in the videos, comments, and Twitter posts. This conversation resulted in three first-round codebooks. We conducted pilot analysis on the top 40 YouTube videos, 695 YouTube comments under those videos, and 100 Twitter posts to ensure that the initial codebooks sufficiently summarized the data. Each of the videos was then evaluated on the nine SPRC safe messaging guidelines. Following the completion of the pilot analysis, the codebooks were modified by the raters together to more accurately summarize the data.

Once the codebooks were finalized, the full analysis of 60 YouTube videos, 1112 YouTube comments under those videos, and 150 Twitter posts was conducted. The two raters individually coded without collaboration to negate bias. The inter-rater reliability measured by Cohen's kappa ranged between 0.61 - 0.88. When the two raters disagreed on a code, the video, comment, or tweet was set aside for consensus coding, a process where the two raters discussed the difference in coding until they agreed on a category [45]. The final codebooks and results are presented in (Tables 3, 4 and 5). Each of the videos was compared to the nine SPRC safe messaging guidelines. The inter-rater reliability measured by Cohen's Kappa was 0.66. If the two raters disagreed on a code, the video was set aside for further consensus coding.

Table 3. BWC's YouTube Videos Codebook and Coding Results in Descending Order.

Major code	Subcategory	Description	Percentage (%)
<b>Media Type</b>	Digital Media	Slideshows, animations, or videos without people	36.67
	Blog/Short Film	Bloggers or short film depicting a victim doing various tasks	35
	News	News videos, interviews, and chat show panels	23.33
	Personal Video	Recorded using personal phones or other personal devices	5
<b>Purpose of the Video: based on what mentioned in the video</b>	Inform or Raise Awareness/Warning	Videos that explained the challenge, how teens were convinced to participate, and provided information about the tasks and the targeted populations	83.33
	Sarcastic, Funny, or Prank	Funny videos about the challenge or the video maker sarcastically pretended to participate in the BWC	10
	Remembering the Victims	Pictured slideshows showing teenagers who allegedly died by playing the game and how	3.33
<b>Tone of the Speaker</b>	Neutral	Videos did not encourage nor discourage participation or	46.67

		there was no speaker throughout the video	
	Negative	Videos expressed sorrow or discouraged people from participating in the challenge	45.00
	Positive	Videos encouraged the victims or families of victims	8.33
<b>Video Content: media used in the video</b>	Victims Related	Included photos before, during, or after self-harm as well as quotes from victims	44.94
	Curator	Videos containing pictures or quotes from the curator	14.61
	Parents	Contained quotes or videos of the parents of victims of the BWC	13.48
	Interaction Between Admin and Victim	Mock interactions of how teens are approached to enter the game	11.24
<b>Video Topics: topics they talked about in the video</b>	Facts about the Challenge	Videos talked about how many people have died due to the BWC, who created the BWC, which countries have been affected by it, full descriptions of the tasks, and/or the different names for the game	46.67
	Social Media	Videos talked about suicide culture or suicide groups online	24.76
	Recommendations	Mentioned some of the interventions by authorities and provided support hotlines as well as recommendations for adults and teens concerning the BWC	21.90
<b>Parties Participating in the Video: who was involved in the process of making the video</b>	Anonymous	Videos had either informists/bloggers, news anchors, or no humans in the video	67.11
	Victims Related	Videos included parents of victims, people who saw the victim, or the actual victim themselves	17.11
	Specialists	Included psychologists, investigators, or other types of specialists commenting on the BWC	15.79

Table 4. YouTube Comments Codebook and Coding Results in Descending Order.

Code	Description	Example	Percentage (%)
<b>Criticizing the Game</b>	Comments which disagreed with the game, asked people not to participate, or asked for the game to be banned	<i>"It seems like every day I find a new thing to hate about social media (sic)"</i>	22.66
<b>Sarcastic, Funny, or Prank</b>	Included jokes and sarcastic comments	<i>"I played this game but I stopped when I was instructed to delete my Minecraft account"</i>	15.74
<b>Encouraging the Video/Video Maker</b>	Comments praising the person who posted the video or applauding the video maker for warning the public about the BWC	<i>"nice,good,perfect bro!!! (sic)"</i>	11.06
<b>Participating</b>	Comments in which users expressed their desire to play the game, agreed with the comments made by the creator of the game, agreed with the game, or asked for links to participate themselves	<i>"I want to play blue whale game plz give me link (sic)"</i>	8.63
<b>Criticizing the Victims</b>	Comments berating those who had played or said they would like to play the game. These comments were often condescending and insensitive to posters who expressed symptoms of	<i>"Your the most sadest human beeing if you kys to 'win the challenge' (sic)"</i>	8.27

	depression		
<b>Expressing Sorrow/Concern</b>	Typically pertained to the victims, the victims' families, or the user's own family and friends	<i>"My heart started crying watching this (sic)"</i>	8.18
<b>Personal Experience</b>	Included comments in which users revealed that they had played in the past or were currently playing or that someone they knew was playing the game or had played the game but did or did not survive	<i>"I already attempted this and I bled and had to go to hospital (sic)"</i>	7.37
<b>Intervention/Recommendation</b>	Included information for a support toll or other forms of SOCIAL support for those affected by the BWC. This code also included encouraging other users to report participation in the BWC to parents or to the authorities	<i>"Let's do a Get Rich Challenge. We all make each others bank accounts happy, one person is chosen, a bunch of people send them money, and it goes on and on. Something like that....someone more intelligent should come up with a formula of some sort that'll work. Let's just send each other money, get rich and help each other grow instead? (sic)"</i>	5.67
<b>Criticizing the Video/Video Maker</b>	Contained comments saying anything negative about the video quality, content or the video maker	<i>"Someone should report these outrageous videos to the authorities."</i>	5.40
<b>Other</b>	Related to the BWC but was not	<i>"Secret task? That means that they</i>	3.96

	summarized by any of the other codes	<i>send you something original or what :) im just courious (sic)</i>	
<b>Encouraging Teens/Parents</b>	The comment praised other teens and encouraged them not to participate or encouraged parents to play a role in their child's safety	<i>"Love you girl so many people care about you. Depression is a horrible :( stay strong girl!! (sic)"</i>	3.06

Table 5. Twitter Posts Codebook and Coding Results in Descending Order.

Post Content	Description	Example	Percentage (%)
<b>Warning/Awareness/Interventions</b>	Warnings about the challenge directed toward parents, teens, the government or the police. This code also included posts that asked for the game and smartphones to be banned and warned against social media use	<i>"Parents, you need to know about the deadly Blue Whale Challenge."</i>	34.00
<b>Information about the Challenge</b>	Posts about the targeted population, the curators, victims, and/or an estimated number of deaths	<i>"Blue Whale Challenge: Creator Budeikin has aides to help him in his absence."</i>	31.33
<b>Sarcastic/Funny/Jokes</b>	Posts that made fun of the challenge or the people who participated in it	<i>"Engineer downloads Blue whale game: Sorry! But no professionals allowed. *Uninstalls itself*"</i>	16.00
<b>Opinion</b>	Either someone voicing their thoughts about	<i>"Apex of misuse of psychology. The blue whale game"</i>	13.33

	the BWC or calling on others to respond to their comment	<i>would have been better if that game maker stressed on social welfare."</i>	
<b>Mental Health</b>	Informing others about depression and suicidal thoughts among teens	<i>"The Blue Whale Challenge: What is the psychology behind it?"</i>	3.33
<b>Personal Experience</b>	Posts in which the users expressed either their own experiences with the BWC or the experiences of someone they knew	<i>"We just had a meeting here at work and this lady told us that her 10-year-old niece committed suicide because of this other Blue Whale Challenge."</i>	0.67
<b>Other</b>	Related to the BWC but did not fit under any of the other codes	<i>"Understanding BLUE WHALE CHALLENGE can tell us how to deal with TERRORISM."</i>	0.67
<b>Miscellaneous</b>	Posts that had nothing with the BWC		0.67

## Results & Findings

This research focused on two main goals: exploring the types of messages social media users share about BWC and violation of SPRC guidelines in the YouTube videos. We examined the common themes among the different social media platforms, then used a deductive coding approach to explore how many of the YouTube videos violated the SPRC guidelines.

### Common themes among the social media platforms

We identified four common themes among all the posts in the three data types included in this study, (Table 6) presents the themes, the corresponding codes, and the percentages of posts from each type of data. The first theme included posts in which social media users were trying to raise awareness and warn parents about this dangerous phenomenon. YouTube videos in this context were either news or bloggers that started their videos by listing different names of the BWC, statistics on how many people died by suicide due to this game, and then talking about the tasks involved in BWC, including that the only way out of the game is to die by suicide. These videos tend to provide clips from interviews with victims' parents or provide pictures of the victims while describing BWC. The YouTube comments were completely against BWC and suggested that parents and authorities pay more attention to children's safety. For example, "This is getting ridiculous. The government or FBI or someone needs to do something". The Twitter posts were centered around awareness of this dangerous game: "Please understand & beware of Blue Whale Challenge -4 parents". This was the most common theme in the YouTube videos (~83%) and Twitter posts (~69%),

which shows that the majority of the users on these two platforms are trying to spread the word about the danger of BWC.

Table 6. Common Themes and Corresponding Codes Among the Three Data Types.

Theme	Codes	YouTube Videos	YouTube Comments	Twitter posts
<b>Trying to raise awareness about BWC and discouraging participation</b>	<b>YouTube Videos:</b> <i>Inform or Raise Awareness/Warning, Facts about the Challenge, Recommendations</i> <b>YouTube Comments:</b> <i>Criticizing the Game, Intervention/Recommendation</i> <b>Twitter:</b> <i>Warning/Awareness/Intervention, Information about the Challenge</i>	83.33%	28.33%	68.67%
<b>Expressing sorrow for participants or for people with mental health</b>	<b>YouTube Videos:</b> <i>Remembering the Victims, Negative Tone of speaker</i> <b>YouTube Comments:</b> <i>Expressing Sorrow/Concern, Encouraging Teens/Parents</i> <b>Twitter:</b> <i>Mental Health</i>	46.67%	11.14%	3.33%
<b>Criticizing or funny posts about the participants or BWC</b>	<b>YouTube Videos:</b> <i>Sarcastic, Funny, or Prank</i> <b>YouTube Comments:</b> <i>Criticizing the Game, Criticizing the Victims, Sarcastic, Funny, or Prank</i> <b>Twitter:</b> <i>Sarcastic/Funny/Jokes</i>	10%	46.67%	16%
<b>Talking about someone who participated in the challenge, saying the already participated or asking to participate</b>	<b>YouTube Videos:</b> <i>Victims Related (both in video content and parties participating), Parents</i> <b>YouTube comments:</b> <i>Participating, Personal experience</i> <b>Twitter:</b> <i>Personal experience</i>	60%	16%	0.67%

The second theme included posts in which the users felt bad for those who participated in BWC or people with mental illness. These videos mainly presented pictures of these people with sad music, and the YouTube comments included words of encouragement and support for people with mental health issues: "Love you girl so many people care about you. Depression is a horrible :( stay strong girl!! (sic)". The Twitter posts mainly mentioned that people are participating in BWC because they have mental health issues: "Blue Whale Challenge: Hey Mr. Scribe, Brush Up Your Archaic Knowledge of Mental Health". A high percent of the videos (~47%) fell in this theme, but fewer comments (~11%) and Twitter posts (~3%).

The third theme included posts of individuals that either criticized the teenagers who participated in

BWC or made fun of them. The YouTube videos that fell under this theme (10%) mentioned that teenagers participate in BWC just to “show off” and criticized them harshly or made fun of them. The YouTube comments that fell under this category (46%) were criticizing the BWC participants by saying things such as: “people who play this game are more stupid than the game itself. How can one lost his sense and manipulated by others...grow up guys...u have brains to think what is good and bad... (sic)”. The Twitter posts in this theme (16%) were mainly sarcastic. For example, “Husband silently downloaded the blue whale game on his #wife's phone. 50 days later the #bluewhale committed suicide (sic)”.

The last theme we identified included posts where users spoke in detail about someone who already participated. This theme was very common in the YouTube videos (60%), where the videos interviewed teenagers’ parents and provided pictures of the teenagers’ body parts showing instances of self-harm. This theme was slightly less common among the YouTube comments (16%), and these posts were mainly about experiences of acquaintances or users that were asking to participate in BWC in the comments: “I want to play blue whale game plz give me link (sic)”. Very few Twitter posts were in this theme (~1%), and they were about acquaintances stories: “We just had a meeting here at work and this lady told us that her 10-year-old niece committed suicide because of this other Blue Whale Challenge.”

### Safe and Effective Messaging Guidelines

Of the 60 videos evaluated based on the SPRC safe messaging guidelines, over one-third (37%) of the YouTube videos met fewer than 3 of the 9 safe messaging guidelines, meaning the videos were considered more unsafe than safe. Approximately 50% of the videos were considered neutral, meaning they met only 4 to 6 of the guidelines, while the remaining 13% were considered more safe than unsafe because they met 7 or more of the guidelines.

When compared to the number of views each video had, 50% of the top 20 viewed videos were more unsafe than safe, meaning the videos violated at least 6 of the 9 criteria for safe and effective messaging for suicide. Only 10% of these 20 videos were considered more safe than unsafe and 40% were considered neutral. The top 20 viewed videos had 46,099,923 views in total. Of the middle 20 most viewed videos, 30% were considered more unsafe than safe, 25% more safe than unsafe, and 45% neutral. The middle 20 viewed videos had 1,169,054 views in total. Of the 20 least viewed videos, 30% were considered more unsafe than safe, 5% considered more safe than unsafe, and 65% neutral. The 20 least viewed videos had only 123,450 views total. All videos met at least one of the nine guidelines, and only one video met all 9 criteria. The total number of views for all 60 videos was 47,392,427, with the top 20 videos having 97.27% of the total views. To better understand which guidelines were most frequently violated, we include the number of videos that violated each guideline in (Table 7). More than 75% of the videos violated the guidelines “Highlight effective treatments for underlying mental health problems” and “Avoid presenting overly detailed descriptions of suicide victims or methods of suicide”.

Table 7. Percentage of Videos Violated Each of the SPRC Safe Messaging Guidelines.

Guideline	Percentage (%)
Highlight effective treatments for underlying mental health problems.	90
Emphasize seeking help and provide information on where to find it.	78.33



Avoid presenting overly detailed descriptions of suicide victims or methods of suicide.	75
List the warning signs, as well as risk and protective factors of suicide.	71.67
Avoid normalizing suicide by presenting it as a common event.	66.67
Avoid glorifying or romanticizing suicide or people who died by suicide.	55
Emphasize prevention.	50
Avoid focusing on personal details of people who died by suicide.	50
Avoid presenting suicide as an inexplicable act or explain it as a result of stress only.	33.33

## Discussion

To the best of our knowledge, our study is the first to systematically document the quality, portrayal, and reach of BWC on social media. In this study, we found that it is easy for adolescents to access almost any post about BWC on social media since only five of the YouTube videos blocked minors from viewing the content. We assessed the portrayal of BWC on social media by investigating the common themes of the videos and posts and the videos' adherence to the safe and effective messaging guidelines.

### Common themes among the social media platforms

Posts in the first theme were meant to raise awareness and warn parents and society about the BWC. These posts were mainly anti-BWC, as opposed to pro-BWC. This finding implies that it is harder to find information on how to participate in BWC in comparison to obtaining pro-BWC or pro-suicide information on the internet as found by a previous study [46]. It also implies that it is hard to find videos from actual participants of the BWC. This could be partially due to the nature of the challenge, in that it encourages participants to conduct self-harm secretly.

Another highlighted topic in BWC related social media posts was the users felt sympathetic towards people who participated in BWC or people with mental illness. This also parallels our first two implications that it is not easy to find pro-BWC information on social media and it is even harder to find posts from actual participants. However, posts of this kind actually make the viewer, potentially an adolescent, think that there are actually other adolescents participating in BWC. This could lead them to neglect the preventability of suicide or self-harm and to believe they are normal responses to a common life situation (e.g. stress).

On the other hand, there were many posts in which people either criticized the adolescents who participated in BWC or made fun of them agreeing with the purpose of BWC of "cleaning the society of people with mental issues," or showing a misunderstanding on why teenagers participate in these kinds of challenges. This indicates that some community members misunderstand why youth participate in self-harming activities. Therefore, greater effort is needed to educate the community about mental health and factors that could lead to self-harm, in addition to educate them on how to respond to unsafe posts.

Lastly, we found a vast amount of social media users tend to speak in detail about someone who participated in BWC, such as providing their demographics or interviewing their parents or acquaintances, who also provide more details about the participant's personal life. This theme was also found by other researchers that looked at traditional media posts about suicide[47]. It is

possible that reporting this level of detail might lead vulnerable adolescents to feeling that they are similar to the adolescent who participated in BWC and make them more likely to participate themselves [48].

### **Safe and Effective Messaging Guidelines**

We suggest that videos like those we examined could contribute to the spread of these deadly challenges instead of their intention of raising awareness. Most posts romanticized people who have died by following this challenge, and younger vulnerable adolescents may see the victims as role models, possibly leading them to end their lives in the same way [47,49]. In violation of guidelines, the videos presented statistics about the number of suicides believed to be related to this challenge in a way that made suicide seem common [50]. In addition, the videos presented extensive personal information about the people who have died by suicide while playing the BWC. They also provided detailed descriptions of the final task, including pictures of self-harm, material that may encourage vulnerable adolescents to consider ending their lives and provide them with methods on how to do so [48,51]. On the other hand, these videos both failed to emphasize prevention by highlighting effective treatments for mental health problems and failed to encourage individuals with mental health problems to seek help and providing information on where to find it.

The SPRC Safe and Effective Messaging guidelines are provided to help people working in suicide prevention, mental health promotion, or any form of media ensure that messages about suicide are safe, positive, and strategic. This makes the guidelines applicable for assessing the appropriateness and safety of the content in messages in suicide campaigns and those discussing suicide across a variety of platforms [36,49]. Since adolescents often look for emotional support from the internet, it is critical that internet-based resources follow these guidelines, as not doing so may contribute to and/or increase the likelihood of someone with suicidal thoughts attempting suicide [36,48,50–52].

As one of the most popular websites and social media platforms in the world, YouTube and Twitter are potentially capable of influencing countless adolescents [2,34,38,42,53,54]. With most of the BWC related posts on these platforms found to be potentially harmful to vulnerable populations, our findings suggest that it is urgent to monitor the social media posts related to BWC and similar self-harming challenges. The SPRC should appropriately inform social media users, particularly those with greater influence (e.g., celebrities, news anchors, etc.), how to address suicide in a safe way in order to reduce contagion. Additionally, there is a possibility that these challenges are hoaxes and by posting about them in a harmful way, the poster is unintentional contributing to self-harm contagion. Therefore, it is critical for social media users to evaluate the sources before sharing any information. They should also be educated on how to respond to unsafe posts and how to report unsafe posts to social media administrators.

Though there are no studies that investigate how BWC (or any other similar challenges, such as Momo Challenge), messages on social media may affect an individual's perception, belief, or behavior regarding self-harm, other studies have showed dramatic effects of movie and TV show portrayals that have led to increased rates of suicide and self-harm using the same methods displayed in the shows [37,55,56]. We expect that social media posts could have a similar effect on self-harm and suicide contagion.

### **Limitations and Comparison with Prior Work**

There are several limitations to our study, many related to the data used. We only studied videos and posts that are publicly available about BWC. It is possible that more harmful videos could be

posted on private pages or personal accounts. The YouTube videos were selected by relevance to the topic based on one keyword. Ideally, videos should be randomly selected using several keywords. Moreover, they may differ from videos sampled at a subsequent time, since YouTube is rapidly changing by nature. Further, this study focused on two social media platforms, while people could be posting about the challenge on other platforms including, but not limited to, Facebook, VKontakte, Snapchat, and Instagram. Our study focused on one self-harm and suicide challenge, the BWC; there are many other self-harm challenges such as the Tide Pod Challenge, the Cinnamon Challenge, and other suicide challenges such as the Momo Challenge.

Future research could include a comparison of the characteristics of the posts about different challenges at various self-harm levels to better understand the reasons for their viral spread. Understanding these factors will help build simulation models, such as Agent Based simulation models to visualize the spread of these challenges. In addition, understanding these factors will help develop and adopt improved policies and interventions for eliminating the spread of harmful challenges among teenagers, such as those found in [57]. Integrating these policies in simulation models will help to identify the policies most effective in reducing these challenges with minimal cost.

## Conclusion

We investigated the characteristics of YouTube and Twitter posts focusing on the BWC and the characteristics of these posts that make them potentially harmful to vulnerable populations. Through our qualitative analysis, we found that while most videos and Twitter posts attempt to raise awareness about the challenge and inform parents, yet they may have unintentional harmful effects since most of the posts violated the SPRC safe messaging guidelines. We conclude that safe messaging guidelines should be more widely disseminated. Our data show that the majority of posters were not professionals, and that these individuals were likely to violate safe messaging guidelines. Greater efforts are needed to disseminate and educate community members on messaging guidelines, as well as factors that encourage contagion effects.

## Acknowledgments

This work was supported by a grant from the National Science Foundation under grant IIS:CHS #1832904.

## Conflicts of Interest

No competing financial interests exist.

## References

1. Lerman BI, Lewis SP, Lumley M, Grogan GJ, Hudson CC, Johnson E. Teen Depression Groups on Facebook: A Content Analysis. *J Adolesc Res* SAGE Publications Inc; 2017 Nov 1;32(6):719–741.
2. Gutowski E, White AE, Liang B, Diamonti A-J, Berado D. How Stress Influences Purpose Development: The Importance of Social Support. *J Adolesc Res* SAGE Publications Inc; 2018 Sep 1;33(5):571–597.
3. Subrahmanyam K, Lin G. Adolescents on the Net: Internet use and well-being. *Adolescence* [Internet] pdfs.semanticscholar.org; 2007;42(168). Available from: <https://pdfs.semanticscholar.org/7d79/f55b9c2d6ab4b437300f4c8e1849f58cc89b.pdf>
4. Baker RK, White KM. Predicting adolescents' use of social networking sites from an extended theory of planned behaviour perspective. *Comput Human Behav Elsevier*; 2010;26(6):1591–

- 1597.
5. Lenhart A, Madden M, Smith A, Macgill AR. Teens and social media: An overview. Washington, DC: Pew Internet and American Life academia.edu; 2009;97-119.
  6. Lenhart A, Duggan M, Perrin A, Stepler R, Rainie H, Parker K, Others. Teens, social media & technology overview 2015. Pew Research Center [Internet & American Life Project]; 2015.
  7. Boyd D, Ellison N. Social network sites: definition, history, and scholarship. *IEEE Eng Manage Rev infona.pl*; 2010;3(38):16-31.
  8. Jackson SJ, Bailey M, Foucault Welles B. # GirlsLikeUs: Trans advocacy and community building online. *New Media & Society SAGE Publications Sage UK: London, England*; 2018;20(5):1868-1888.
  9. Blackwell L, Hardy J, Ammari T, Veinot T, Lampe C, Schoenebeck S. LGBT Parents and Social Media: Advocacy, Privacy, and Disclosure During Shifting Social Movements. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems New York, NY, USA: ACM*; 2016. p. 610-622.
  10. Gonzales AL. Disadvantaged minorities' use of the Internet to expand their social networks. *Communic Res SAGE Publications Sage CA: Los Angeles, CA*; 2017;44(4):467-486.
  11. Lawson CE. Platform vulnerabilities: harassment and misogyny in the digital attack on Leslie Jones. *Inf Commun Soc Routledge*; 2018 Jun 3;21(6):818-833.
  12. Fritz N, Gonzales A. Privacy at the Margins| Not the Normal Trans Story: Negotiating Trans Narratives While Crowdfunding at the Margins. *Int J Commun Syst ijoc.org*; 2018 Mar 1;12(0):20.
  13. Wood M, Rose E, Thompson C. Viral justice? Online justice-seeking, intimate partner violence and affective contagion. *Theoretical Criminology SAGE Publications Ltd*; 2018 Jan 1;1362480617750507.
  14. Page X, Wisniewski P, Knijnenburg BP, Namara M. Social media's have-nots: an era of social disenfranchisement. *Internet Research Emerald Publishing Limited*; 2018;28(5):1253-1274.
  15. Linabary JR, Corple DJ. Privacy for whom?: a feminist intervention in online research practice. *Inf Commun Soc Routledge*; 2018 Feb 19;1-17.
  16. Flores-Yeffal NY, Vidales G, Martinez G. # WakeUpAmerica,# IllegalsAreCriminals: the role of the cyber public sphere in the perpetuation of the Latino cyber-moral panic in the US. *Inf Commun Soc Taylor & Francis*; 2017;1-18.
  17. Mitchell KJ, Wells M, Priebe G, Ybarra ML. Exposure to websites that encourage self-harm and suicide: prevalence rates and association with actual thoughts of self-harm and thoughts of suicide in the United States. *J Adolesc* 2014 Dec;37(8):1335-1344. PMID:25313930
  18. Guan L, Hao B, Cheng Q, Yip PS, Zhu T. Identifying Chinese Microblog Users With High Suicide Probability Using Internet-Based Profile and Linguistic Features: Classification Model. *JMIR Ment Health mental.jmir.org*; 2015 Apr;2(2):e17. PMID:26543921
  19. Emma Hilton C. Unveiling self-harm behaviour: what can social media site Twitter tell us about self-harm? A qualitative exploration. *J Clin Nurs Wiley Online Library*; 2017;26(11-12):1690-1704.
  20. Zhu L, Westers NJ, Horton SE, King JD, Diederich A, Stewart SM, Kennard BD. Frequency of Exposure to and Engagement in Nonsuicidal Self-Injury Among Inpatient Adolescents. *Arch Suicide Res* 2016 Oct;20(4):580-590. PMID:26983456
  21. Berry N, Bucci S, Lobban F. Use of the Internet and Mobile Phones for Self-Management of Severe Mental Health Problems: Qualitative Study of Staff Views. *JMIR Mental Health mental.jmir.org*; 2017 Nov 1;4(4):e52. PMID:29092809
  22. Insel BJ, Gould MS. Impact of modeling on adolescent suicidal behavior. *Psychiatr Clin North*

- Am Elsevier; 2008 Jun;31(2):293–316. PMID:18439450
23. Purington A, Whitlock J. Non-suicidal self-injury in the media. *Prev Res Integrated Research Services, Inc.*; 2010;17(1):11–14.
  24. Klingle KE, Van Vliet KJ. *Self-Compassion From the Adolescent Perspective: A Qualitative Study*. J Adolesc Res SAGE Publications Sage CA: Los Angeles, CA; 2017;0743558417722768.
  25. Livingstone S, Smith PK. Annual research review: Harms experienced by child users of online and mobile technologies: The nature, prevalence and management of sexual and aggressive risks in the digital age. *J Child Psychol Psychiatry Wiley Online Library*; 2014;55(6):635–654.
  26. Pater J, Mynatt ED. *Defining Digital Self-Harm*. CSCW 2017. p. 1501–1513.
  27. Robert A, Suelves JM, Armayones M, Ashley S. Internet use and suicidal behaviors: internet as a threat or opportunity? *Telemed J E Health liebertpub.com*; 2015 Apr;21(4):306–311. PMID:25654489
  28. Lewis SP, Seko Y. A double-edged sword: A review of benefits and risks of online nonsuicidal self-injury activities. *J Clin Psychol Wiley Online Library*; 2016;72(3):249–262.
  29. Patton GC, Coffey C, Romaniuk H, Mackinnon A, Carlin JB, Degenhardt L, Olsson CA, Moran P. The prognosis of common mental disorders in adolescents: a 14-year prospective cohort study. *Lancet Elsevier*; 2014 Apr 19;383(9926):1404–1411. PMID:24439298
  30. Muñoz-Sánchez J-L, Delgado C, Parra-Vidales E, Franco-Martín M. Facilitating Factors and Barriers to the Use of Emerging Technologies for Suicide Prevention in Europe: Multicountry Exploratory Study. *JMIR Mental Health mental.jmir.org*; 2018 Jan 24;5(1):e7. PMID:29367183
  31. Mukhra R, Baryah N, Krishan K, Kanchan T. “Blue Whale Challenge”: A Game or Crime? *Sci Eng Ethics [Internet]* 2017 Nov 11; PMID:29129013
  32. Padgett DK. *Qualitative and Mixed Methods in Public Health*. SAGE Publications; 2011. ISBN:9781483342313
  33. Pater JA, Haimson OL, Andalibi N, Mynatt ED. “Hunger Hurts but Starving Works”: Characterizing the presentation of eating disorders online. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing ACM*; 2016. p. 1185–1200.
  34. Pater JA, Miller AD, Mynatt ED. *This Digital Life: A Neighborhood-Based Study of Adolescents’ Lives Online*. CHI Conf Proc [Internet] dl.acm.org; 2015; Available from: <https://dl.acm.org/citation.cfm?id=2702534>
  35. Chambers DA, Pearson JL, Lubell K, Brandon S, O’Brien K, Zinn J. The science of public messages for suicide prevention: a workshop summary. *Suicide Life Threat Behav Guilford Press*; 2005 Apr;35(2):134–145. PMID:15843331
  36. Suicide Prevention Resource Center. *Safe and effective messaging for suicide prevention*. Suicide Prevention Resource Center Newton, MA; 2006.
  37. Gould M, Jamieson P, Romer D. *Media Contagion and Suicide Among the Young*. Am Behav Sci SAGE Publications Inc; 2003 May 1;46(9):1269–1284.
  38. Maina A. 20 Popular Social Media Sites Right Now - Small Business Trends [Internet]. *Small Business Trends*. 2016 [cited 2019 Aug 18]. Available from: <https://smallbiztrends.com/2016/05/popular-social-media-sites.html>
  39. Fusch PI, Ness LR. Are We There Yet? Data Saturation in Qualitative Research. *The Qualitative Report nsuworks.nova.edu*; 2015;20(9):1408–1416.
  40. Bromberg JE, Augustson EM, Backinger CL. Portrayal of smokeless tobacco in YouTube videos. *Nicotine Tob Res academic.oup.com*; 2012 Apr;14(4):455–462. PMID:22080585
  41. Richardson A, Vallone DM. YouTube: a promotional vehicle for little cigars and cigarillos? *Tob Control tobaccocontrol.bmj.com*; 2014 Jan;23(1):21–26. PMID:23047887

42. Luo C, Zheng X, Zeng DD, Leischow S. Portrayal of electronic cigarettes on YouTube. *BMC Public Health* [bmcpublichealth.biomedcentral.com](http://bmcpublichealth.biomedcentral.com); 2014 Oct 3;14:1028. PMID:25277872
43. Radian 6 Overview [Internet]. [cited 2019 Aug 18]. Available from: <https://www.webanalyticsworld.net/analytics-measurement-and-management-tools/radian-6-overview>
44. Khattar A, Dabas K, Gupta K, Chopra S, Kumaraguru P. White or Blue, the Whale gets its Vengeance: A Social Media Analysis of the Blue Whale Challenge [Internet]. arXiv [csSI]. 2018. Available from: <http://arxiv.org/abs/1801.05588>
45. Saldana J. *The Coding Manual for Qualitative Researchers*. SAGE; 2015. ISBN:9781473943599
46. Recupero PR, Harms SE, Noble JM. Googling suicide: surfing for suicide information on the Internet. *J Clin Psychiatry* [psycnet.apa.org](http://psycnet.apa.org); 2008 Jun;69(6):878–888. PMID:18494533
47. Fekete S, Schmidtke A. PsycNET [Internet]. [psycnet.apa.org](http://psycnet.apa.org); 1995 [cited 2019 Aug 18]. Available from: <http://psycnet.apa.org/record/1995-98656-010>
48. Fekete S, Macsai E. Hungarian suicide models, past and present. *Suicidal behavior and risk factors* 1990;149–156.
49. Noble-Carr D, Woodman E. Considering Identity and Meaning Constructions for Vulnerable Young People. *J Adolesc Res* SAGE Publications Inc; 2018 Nov 1;33(6):672–698.
50. Cialdini RB. Crafting Normative Messages to Protect the Environment. *Curr Dir Psychol Sci* SAGE Publications Inc; 2003 Aug 1;12(4):105–109.
51. Sonneck G, Etzersdorfer E, Nagel-Kuess S. Imitative suicide on the Viennese subway. *Soc Sci Med Elsevier*; 1994 Feb;38(3):453–457. PMID:8153751
52. Jacobs DG. PsycNET [Internet]. [psycnet.apa.org](http://psycnet.apa.org); 1999 [cited 2019 Aug 18]. Available from: <http://psycnet.apa.org/record/1998-06811-000>
53. Vaterlaus JM, Tulane S, Porter BD, Beckert TE. The Perceived Influence of Media and Technology on Adolescent Romantic Relationships. *J Adolesc Res* SAGE Publications Inc; 2018 Nov 1;33(6):651–671.
54. Swendeman D, Arnold EM, Harris D, Fournier J, Comulada WS, Reback C, Koussa M, Ocasio M, Lee S-J, Kozina L, Others. Text-Messaging, Online Peer Support Group, and Coaching Strategies to Optimize the HIV Prevention Continuum for Youth: Protocol for a Randomized Controlled Trial. *JMIR Res Protoc* JMIR Publications Inc., Toronto, Canada; 2019;8(8):e11165.
55. Gould MS, Shaffer D. The Impact of Suicide in Television Movies. *N Engl J Med* Massachusetts Medical Society; 1986 Sep 11;315(11):690–694.
56. Tanner MA, Murray JA, Phillips DP. The impact of televised movies about suicide. *N Engl J Med* [scholars.northwestern.edu](http://scholars.northwestern.edu); 1988 Mar 17;318(11):707–708. PMID:3344023
57. Luxton DD, June JD, Fairall JM. Social media and suicide: a public health perspective. *Am J Public Health* [ajph.aphapublications.org](http://ajph.aphapublications.org); 2012 May;102 Suppl 2:S195–200. PMID:22401525
58. Center for Mental Health Services (US), Office of the Surgeon General (US). *National Strategy for Suicide Prevention: Goals and Objectives for Action*. Rockville (MD): US Public Health Service; 2010.
59. Stone DM, Simon TR, Fowler KA, Kegler SR, Yuan K, Holland KM, Ivey-Stephenson AZ, Crosby AE. Vital Signs: Trends in State Suicide Rates - United States, 1999-2016 and Circumstances Contributing to Suicide - 27 States, 2015. *MMWR Morb Mortal Wkly Rep* [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov); 2018 Jun 8;67(22):617–624. PMID:29879094
60. Baldessarini RJ, Tondo L, Hennen J. Effects of lithium treatment and its discontinuation on suicidal behavior in bipolar manic-depressive disorders. *J Clin Psychiatry* [europepmc.org](http://europepmc.org); 1999;60 Suppl 2:77–84; discussion 111–6. PMID:10073392