Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults

Florian Arendt
University of Vienna, Austria

Sebastian Scherr
KU Leuven, Belgium

Daniel Romer
University of Pennsylvania, USA

Abstract
Suicide is a leading cause of death among youth, and media depictions of suicidal behavior can be a contributing risk factor. Of interest, Instagram recently received more scholarly attention due to its large number of publicly available, explicit, and graphic depictions of self-harm. Importantly, researchers have hypothesized that exposure to this content could be a risk for self-harm and suicide in vulnerable audiences. We tested this hypothesis using a two-wave US panel survey among young adults (N = 729). Analyses indicated that exposure to self-harm on Instagram was associated with suicidal ideation, self-harm, and emotional disturbance even controlling for exposure to other sources with similar content. As hypothesized, exposure to self-harm on Instagram at the first wave prospectively predicted self-harm and suicidality-related outcomes at the second wave 1 month later. These findings provide evidence that such exposure can lead to contagion in vulnerable users. Implications are discussed.

Keywords
Instagram, self-harm, social media, suicide, suicide contagion, Werther effect, young adults

Corresponding author:
Florian Arendt, University of Vienna, Währinger Straße 29, Vienna 1090, Austria.
Email: florian.arendt@univie.ac.at
Deliberate self-harm tends to be repetitive and largely non-suicidal (Klonsky et al., 2014), but it increases the risk of future suicide (Chandler et al., 2015; Dyson et al., 2016). Among youth, suicide is especially problematic and a relevant global public health threat given that suicide has been the second-leading cause of death among 15- to 29-year-olds according to recent World Health Organization (WHO, 2018) estimates. Media depictions are deemed to play a key role (Niederkrotenthaler and Stack, 2017). In fact, exposure to media depictions of suicide can elicit imitative effects that may lead to additional suicides (Romer et al., 2006; Stack, 2005), called the “Werther effect” (Phillips, 1974), including imitational effects among adolescents and young adults (Schmidtke and Häfner, 1988). Empirical evidence for media-induced effects on suicide rates continues to grow (Niederkrotenthaler et al., 2009; Pirkis and Blood, 2001; Sisask and Värnik, 2012; Stack, 2003, 2005).

Importantly, our knowledge about media effects in the suicide domain still largely traces back to traditional, legacy media studies, largely dominated by televised depictions and newspaper stories (see, for example, Wasserman et al., 1994). However, the Internet has dramatically changed the way suicide and self-harm are depicted and publicly accessible today (Pirkis et al., 2017; Reid-Chassiakos et al., 2016). Social networking sites have become increasingly popular during recent years, especially among youth (Pew Research Center, 2015). Notably, Instagram has over 1 billion monthly users and over 500 million daily active users, and it is one of the most popular social networking platforms among young people (Instagram, 2018). As we will outline below, Instagram represents a publicly available, unrivaled source for large amounts of regularly uploaded explicit and graphic self-harm depictions.

Previous Instagram-related research on self-harm focused almost exclusively on use patterns (e.g. Mitchell and Ybarra, 2007; Vente et al., 2017) and the prevalence of this content (see below). Importantly, we lack knowledge on the effects of exposure. There may be helpful effects, such as reduction of social isolation by providing a sense of community, encouragement for recovery, and reduction of deliberate self-harm urges. However, harmful effects may include a triggering effect on imitative behaviors (Brown et al., 2018). Although there are repeatedly expressed speculations about effects that can theoretically range from “contagion” to “caring” (Carlyle et al., 2018), the detrimental consequences are typically emphasized by previous research (see, for an exception, Scherr and Reinemann, 2011 for the beneficial longitudinal effects of online health forums on suicidality). Taken together, it has been speculated that exposure to self-harm on Instagram is related to self-harm and suicidality-related outcomes, possibly eliciting a triggering effect which may lead to increases in imitative self-harm behaviors and increased levels of suicidality (Arendt, 2019; Baker and Lewis, 2013; Brown et al., 2018).

In the present article, we present the results of a longitudinal two-wave panel survey study among young adults living in the United States. To the best of our knowledge, the present study is the first to test for the hypothesized consequences of exposure to self-harm on Instagram among young people.

**Suicide and social media**

The importance of social media, including Instagram, for suicide prevention has been acknowledged in recent scholarship: In a seminal study, Brown et al. (2018) presented a
content analysis of Instagram posts and pictures. Analyses indicated that pictures and communication about self-harm can frequently be found on this social networking site. Most commonly those were of mild or moderate injuries, but pictures of bleeding flesh wounds caused by cuttings on extremities, frequently on arms or thighs, could also be identified. Interestingly, the deeper the cuts, the more reactions in the comments section below the posts were observed. Most elicited comments were neutral or empathic in nature, some even offered help, and only a few comments were hostile.

Arendt (2019) provides a content analysis of Instagram posts, including pictures and video clips, that were publicly available via the German speaking hashtag #selbstmord (i.e. #suicide). Consistent with the study by Brown and colleagues, he showed that cutting was far the most frequently depicted method. Emotional expressions were frequently depicted. Although sadness had been identified as the most frequently depicted emotion, self-hate and loneliness also appeared regularly. Moreover, many posts depicted content pointing to a gap between one’s inner mental state (e.g. sadness) and one’s overtly expressed behavior (e.g. smiling). Expressed death wishes and awareness–intervention material from professional organizations were very rare.

Carlyle et al. (2018) conducted a content analysis of suicide-themed posts on Instagram that were tagged on #suicide and/or #suicidal. Analyses indicate that self-harm was present in the majority of the posts. Posts about suicide ideation elicited higher engagement than posts that did not. Consistent with the findings provided by Arendt (2019), potentially helpful materials from professional public health voices were largely absent. However, the effects of the predominantly visual content on Instagram remain largely unknown. There is some limited evidence from qualitative research about the normalizing effects of pictures shared on the Tumblr platform on 16- to 24-year-olds (Jacob et al., 2017). In 21 interviews, the study showed that anonymously posted pictures can invoke a physical reaction and provide an inspiration for own self-harm behaviors, particularly among young participants who visited the platform regularly. The visuals can normalize the act of self-harm in the young people’s eyes.

Above and beyond this, Moreno et al. (2016) recently stressed a new, platform-dependent phenomenon that we may call “accidental exposure.” Some non-suicide-related hashtags such as #cat overlap with suicide-related hashtags such as #cutting. Some users may perceive an overlap of cat-related concepts, such as scratching or clawing, and suicide-related concepts, such as cutting. This overlap may be due to the fact that both cats and (self–harm-related) cutting yield injuries that look similar (i.e. red lines) and that individuals who engage in self-harm might as well use cats as an explanation for their wounds or scars (“My cat scrapped me”). Moreno et al. (2016) raised awareness to the fact that users who may look for pictures of cute cats by intentionally viewing the content tagged on #cat may be inadvertently exposed to posts with cutting-related self-harm content. Thus, as a supplement to intentional exposure, some users may get accidentally exposed to self-harm.

The present study

This study presents the first investigation of the relationship between exposure to self-harm on Instagram and self-harm and suicidality-related outcomes. We sought
to determine whether existing risks for suicide and self-harm are related to exposure to self-harm on Instagram. Furthermore, we probed whether exposure was likely to increase or maintain such outcomes. To answer these questions, we conducted a two-wave panel study with young people aged between 18 and 29 years.

Based on the research described above, we hypothesized that exposure to self-harm on Instagram will be positively related to self-harm and suicidality-related outcomes (General Hypothesis). We focused on a broad range of established self-harm and suicidality-related outcomes to reduce the likelihood of overlooking important relationships. These ranged from a user’s own self-harming behavior (H1), suicidal ideation (H2), and specific suicide plans (H3) to more general suicide risk (H4) and established predictors of suicidality (see Klonsky et al., 2012; Luo et al., 2016), such as hopelessness (H5) and reasons for living (H6). To reduce the possibility that other sources of suicidal content might be responsible for the relation, we also assessed other sources of exposure to suicidal behavior either in media or face-to-face relationships.

Regarding the second wave of the panel study, we hypothesized that exposure to self-harm on Instagram as reported in the first wave would be related to subsequent reports of the same set of outcomes holding constant prior levels of those outcomes. We reasoned that with the second wave occurring only 1 month after the first, the effects of prior exposure could still be evident even after controlling for wave 1 outcomes. On the other hand, if any relation between the suicidal outcomes and exposure merely reflected the effects of selection by those who were already at risk of suicide and self-harm, then controlling for those risks at wave 1 should remove any subsequent relation between prior exposure and those risks at wave 2.

We applied a regression-based approach to a two-wave panel data set to test these hypotheses. Using a longitudinal study design in which we collected indicators of self-harm and suicidality-related outcomes from the same participants at two points in time approximately 1 month apart, we drew conclusions about short-term, within-subject changes. Of interest, we investigated whether exposure to self-harm on Instagram is cross-sectionally and prospectively related to self-harm and suicidality-related outcomes. We acknowledge that the relationship between individual suicidality, exposure to mediated self-harm, and other individual- and societal-level factors is complex, intertwined, and represents a process of development over time. Thus, a panel perspective may allow for a deeper understanding of how a social media site such as Instagram influences how self-harm and suicidality-related outcomes develop (i.e. change) over time.

**Method**

We conducted a two-wave panel survey relying on a convenience sample of young adults aged between 18 and 29 years living in the United States. Exposure to self-harm content on Instagram was measured in the first wave and all self-harm and suicidality-related concepts were measured in both panel waves. We focused on cross-sectional relationships between exposure to self-harm content on Instagram and a range of outcomes (i.e. wave 1 measures) and also tested for panel effects looking at changes in these outcomes from wave 1 to wave 2 again in relation to Instagram self-harm exposure.
Participants

As part of a larger project on the role of new media in the self-harm domain, we contracted the online survey firm Qualtrics to recruit young adults aged between 18 and 29 years and reported having access to Netflix. The criterion of Netflix access is relevant for another module of the larger project in which we focus on suicide depictions on this streaming platform (Arendt et al., 2019). Qualtrics obtained the participants from a survey vendor that maintains a panel of respondents recruited from various Internet gaming sites. Approximately 10% of those who were invited began the survey, a typical response rate in comparable studies that has not been found to adversely affect disclosure rates (Pew Research Center, 2012). This (non-representative) convenience sample does not represent a high-risk sample of highly vulnerable individuals. However, the fact that respondents were recruited from various Internet gaming sites enhanced the likelihood that they were users of social media. At the outset of the survey, participants were told that the study involved a two-part anonymous survey regarding “media content that features suicide.” We told them that they would be invited to a second survey in a month’s time to “see if your experiences around this topic have changed.” They were invited again approximately 1 month after the first wave to complete the follow-up panel wave two survey.

A total of \( N=1262 \) individuals participated in the first wave and \( N=729 \) individuals participated in both waves and represent the sample for all statistical analyses. With the exception of a slight increase in age at the second wave, there were no significant demographic differences between those who participated at waves 1 and 2. Of these respondents, 81.6% were female, 16.7% were male, and 1.7% indicated “other” or “refused to answer.” The sample ranged in age between 18 and 29 (\( M=24.15, \text{ standard deviation } [SD]=3.23 \) years. Less than 9% did not graduate from high school, 30% graduated high school, 25% had some college experience, and the rest had a college/university degree (36%). The majority indicated that they are “white” (72%).

Instagram exposure measures

All Instagram-related variables were measured in the panel wave 1 survey.

Exposure to self-harm on Instagram. We asked the following question: Please think about the social networking site Instagram: How often, if ever, have you seen a post on Instagram showing someone who intentionally harms him- or herself, for example, by cutting? Was it more than once, just once, or never? A substantial part of the sample had never seen self-harm content on Instagram (\( n=325, 45\% \)). A comparable number of individuals had seen self-harm content on Instagram just once (\( n=129, 18\% \)) or more than once (\( n=184, 25\% \)). Some participants refused to answer or did not know (\( n=91, 12\% \)). We did not use the latter group for the regression-based statistical analyses reported here. For these regression analyses, we dummy-coded this exposure measure: dummy 1 (just once) and dummy 2 (more than once); reference category = never.

Intentional exposure. For those who had seen self-harm content on Instagram (\( n=313 \)), we asked whether this exposure happened by accident or did they intentionally look for
self-harm-related content on Instagram? Only the minority indicated that they intentionally searched for this content \((n=63, 20\%)\).

**Emotional disturbance.** We used a single item to measure whether or not users who were exposed to self-harm got emotionally disturbed from the content: *Did this exposure emotionally disturb you?* The majority indicated *Yes* \((n=184, 59\%)\). Other participants indicated *No* \((n=104, 33\%)\) or refused to answer or did not know \((n=25, 8\%)\).

**Emotional perspective-taking.** We used a single item to measure whether or not users took the perspective and tried to imagine how it would be to emotionally feel that way: *Did you ever think about how it would feel if you would do the same thing to yourself?* The majority said *No* \((n=180, 58\%)\). However, a substantial part of the sample indicated *Yes* \((n=122, 39\%)\). A minority refused to answer \((n=11, 4\%)\).

**Presumed copycat influence.** We used a single item to measure whether or not users thought that exposure to self-harm content elicited imitative effects in themselves: *As a consequence of seeing this self-harm content on Instagram, did you ever perform the same (or a very similar) self-harming behavior?* The majority indicated *No* \((n=206, 66\%)\). However, a somewhat surprisingly high number of participants indicated *Yes* \((n=99, 32\%)\). Some refused to answer \((n=7, 2\%)\).

**Self-harm and suicidality-related target outcomes**

We focused on target outcomes that represent established concepts. All of these outcomes were measured both at wave 1 \((n_1)\) and at wave 2 \((n_2)\).

**Self-harming behavior.** We asked whether the respondents had ever (wave 1: *Have you ever engaged in self-harming behavior, such as cutting your wrists?*) or in the past month (wave 2: *Since the first survey, have you ever engaged in self-harming behavior, such as cutting your wrists?*) engaged in self-harming behaviors in line with research on non-suicidal self-injury (Muehlenkamp et al., 2012). The majority indicated *No* \((n_1=500, n_2=562)\), but a substantial number of participants indicated *Yes* \((n_1=195, n_2=132)\). Some participants refused to answer \((n_1=33, n_2=35)\).

**Suicidal ideation.** We used one item from the Youth Risk Behavior Survey (Kann et al., 2018) that measures suicidal ideation: *Have you ever (wave 1) or in the past month (wave 2) seriously considered attempting suicide?* This measure had four relevant response options (coded 1–4): *never* \((n_1=366, n_2=498)\), *once* \((n_1=130, n_2=72)\), *twice* \((n_1=89, n_2=60)\), and *three or more times* \((n_1=75, n_2=38)\). We used additional response options: Some refused to answer or indicated that they did not know \((n_1=69, n_2=61)\). We excluded the latter category for regression-based analyses.

**Suicide plan.** For those who considered attempting suicide at least once, we asked whether the respondents had ever (wave 1) or in the past month (wave 2) made a precise plan how to die by suicide: *Have you ever made a precise plan about how you would attempt
The item, stemming from the Youth Risk Behavior Survey (Kann et al., 2018), had the response options: never \((n_1=81, n_2=37)\), once \((n_1=96, n_2=51)\), twice \((n_1=55, n_2=42)\), and three or more times \((n_1=40, n_2=31)\). Some \((n=22, n_2=8)\) refused to answer or indicated that they did not know.

**Suicide risk.** We used items from the Suicide Probability Scale of Eltz et al. (2007). This scale has been found to predict repeated suicide attempts in an adolescent sample. Respondents were asked, How often, if ever, are the following statements true of yourself? (wave 1) and Since the first survey, how often do you feel the following statements are true? (wave 2). They were asked to rate six items (I think of suicide; I have thoughts about how to end my life; I feel it would be less painful to die than to keep living, given the way things are; I feel the world is not worth continuing to live in; I feel people would be better off if I were dead; In order to punish others, I think of suicide) on a five-point scale ranging from 1 (none of the time) through 3 (sometimes) to 5 (most of the time). The sample showed moderate to low levels on suicide risk, both in wave 1 \((M=1.97, SD=1.08, \alpha=0.94)\) and wave 2 \((M=2.10, SD=1.16, \alpha=0.95)\).

**Hopelessness.** As an important suicide-related concept, we measured hopelessness using the validated brief version (Yip and Cheung, 2006) of Beck’s Hopelessness Scale (Beck et al., 1985). Participants rated four items: In the future, I expect to succeed in what concerns me most; My future seems dark to me; I just don’t get the breaks, and there is no reason to believe I will in the future; I have great faith in the future) on a seven-point scale ranging from 1 (strongly disagree) to 7 (strongly agree). The sample showed low to moderate levels of hopelessness, both in wave 1 \((M=3.01, SD=1.15, \alpha=0.63)\) and wave 2 \((M=3.16, SD=1.19, \alpha=0.66)\). Unfortunately, the internal consistencies are low but just acceptable.

**Reasons for living.** We used the “survival and coping beliefs” subscale of the Brief Reasons for Living Scale (Cwik et al., 2017). This concept has been shown to be responsive for media influence (Till et al., 2017). Participants rated two items: I believe I can find a purpose in life, a reason to live; I do not want to die on a 7-point scale ranging from 1 (strongly disagree) to 7 (strongly agree). The sample showed moderate to strong reasons for living, both in wave 1 \((M=5.43, SD=1.55, \alpha=0.76)\) and wave 2 \((M=5.38, SD=1.44, \alpha=0.80)\).

**Controls**

**Exposure to self-harm via other sources.** We also measured exposure to suicide via other media channels in the wave 1 survey. This is important to separate out the genuine influence of exposure to self-harm on Instagram. In fact, we used exposure to six alternative sources for exposure to self-harm (How often, if ever, have you heard or seen a story about someone who attempted to or actually died by suicide ...): (1) in a newspaper story (print or online) (56.7%), (2) from a face-to-face conversation with a friend or relative (56.5%), (3) on a video-streaming site, such as Netflix (50.9%), (4) on an Internet news site, such as CNN or Yahoo News (46.6%), (5) when chatting online (35.0%), or (6) on
an online forum, discussion board, or self-help website (37.6%). We coded whether participants have heard or seen a suicide story (coded as 1; for each of the six sources). We summed up all items (\(M=2.83, SD=1.99\)).

**Socio-demographics.** We also measured age, gender, education, and race and used these measures as control variables in the regression analyses.

**Statistical analysis**

We used a regression-based approach to test the hypotheses. For cross-sectional analyses, we put age, gender (dummy 1 = male, dummy 2 = female, reference category = other or refused to answer), education, and race (0 = non-white, 1 = white), and the exposure to self-harm via other sources into the first step of a hierarchical regression model. In the second step, we included the dummy-coded exposure to self-harm on Instagram (see above). We used the target outcomes measured at wave 1 as the dependent variables. This approach tests whether there is a relationship between lifetime prevalence of self-harm exposure on Instagram and the outcomes.

For the panel analyses, we relied on a similar approach. Again, we put age, gender, education, race, and the exposure to self-harm via other sources into the first step. We also included the target outcome’s baseline measure (wave 1). In the second step, we included the dummy-coded exposure to self-harm on Instagram. We used the target outcomes measured at wave 2 as the dependent variables. This approach allowed us to test whether exposure to self-harm on Instagram (wave 1) predicted a change in target outcomes between both panel waves.

We relied on both the cross-sectional and panel approaches because each has its limitations. However, when applied together, they allow for a more thorough test of the hypotheses. We report the full models (step 2) in the tables and provide the statistics for changes in \(R^2\) and \(\chi^2\) (step 2) to assess whether Instagram exposure adds explanatory power. Furthermore, we report (unplanned) additional analyses regarding the role of accidental and intentional exposure at the end of the “Results” section.

**Ethics statement**

The institutional review board of the University of Pennsylvania approved the study.

**Results**

**Preliminary analysis**

We found that among those who were exposed to self-harm content on Instagram (\(N=313, 43\%\)), only 20.1\% indicated to have intentionally searched for this content. Thus, accidental exposure to self-harm content on Instagram was of central relevance for the majority of participants who were exposed to self-harm. Indeed, the majority of users seem to make contact with Instagram’s self-harm content unintentionally.
Among those who could remember and did not refuse to tell, the majority (63.9%) indicated that their exposure to self-harm content on Instagram had emotionally disturbed them. Interestingly, emotional disturbance did not show correlations with own self-harming behavior, suicidal ideation, suicide plans, suicide risk, hopelessness (all \( ps > .12 \)), except with reasons for living, \( r(286) = .20, p = .001 \). Furthermore, among those who were exposed to self-harm content on Instagram, the majority (59.6%) indicated that they have thought about how it would feel if they would do the same thing to themselves (perspective-taking). Interestingly, this measure did not show a correlation with reasons for living, \( r(300) = -.07, p = .252 \), but showed positive relationships with all other target outcomes, that is, own self-harming behavior, \( r(286) = .34, p < .001 \); suicidal ideation, \( r(279) = .32, p < .001 \); suicide plans, \( r(144) = .26, p = .002 \); suicide risk, \( r(299) = .30, p < .001 \); and hopelessness, \( r(299) = .18, p = .001 \).

Similarly, 32.5% indicated that they have performed the same (or very similar) self-harming behavior as a consequence of seeing self-harm content on Instagram, indicating a substantial number of participants reporting a (presumed) copycat influence. This analysis speaks to self-reported and not copycat effects and, therefore, may not indicate actual media effects. Nevertheless, this measure showed substantial correlations with all self-harm and suicidality-related outcomes, that is, own self-harming behavior \( r(302) = .40, p < .001 \); suicidal ideation \( r(280) = .27, p < .001 \); suicide plans \( r(143) = .16, p = .063 \); suicide risk \( r(302) = .40, p < .001 \); hopelessness \( r(302) = .26, p < .001 \); and reasons for living \( r(303) = -.17, p = .002 \).

**Test of hypotheses**

We hypothesized that exposure to self-harm on Instagram would be positively related to self-harm and suicidality-related outcomes. We used six operational hypotheses to test this general hypothesis: effects on a user’s own self-harming behavior \((H1)\), suicidal ideation \((H2)\), specific suicide plans \((H3)\), general suicide risk \((H4)\), and predictors of suicidal behavior, namely, hopelessness \((H5)\), and reasons for living \((H6)\). In a first step, we provide cross-sectional evidence. In a second step, we report a panel analysis predicting the change in outcomes between both panel waves by exposure to self-harm on Instagram.

**Cross-sectional evidence.** We now report whether lifetime exposure to self-harm on Instagram (wave 1) is related to self-harm and suicidality-related outcomes (wave 1). Regarding the latter, self-harming behavior, suicidal ideation, and suicide plans were also measured as lifetime prevalence, focusing on the same time period as the exposure measure. Conversely, hopelessness, reasons for living, and suicide risk are measured with scales that focus on the very moment when filling out the questionnaire (see “Methods” section).

As Tables 1 to 6 (column: “Cross-sectional evidence”) show, exposure to self-harm was statistically related to five outcomes, indicated by a significant change in explained variance of the second step (i.e. dummy-coded exposure to self-harm on Instagram) in five regression models: self-harming behavior, \( \Delta \chi^2_{(2)} = 20.30, \Delta R^2 = .044, p < .001 \); suicidal ideation, \( \Delta F(2, 583) = 5.78, \Delta R^2 = .018, p = .003 \); suicide risk, \( \Delta F(2, 625) = 26.69, \)
Table 1. Predicting self-harming behavior: hierarchical binary logistic regression (step 2: full model).

| Outcome = self-harming behavior | Cross-sectional evidence (predicting wave 1 outcome) | Panel evidence (predicting wave 2 outcome) |
|--------------------------------|-----------------------------------------------------|---------------------------------------------|
|                                | B        | SE   | OR  | p        | B        | SE   | OR  | p        |
| Step 2 (effect of Instagram exposure) |                      |                               |                               |
| \(\Delta \chi^2 = 20.30, \Delta R^2 = .044, p < .001\) | \(\Delta \chi^2 = 28.10, \Delta R^2 = .064, p < .001\) |
| Age                            | 0.02     | 0.03 | 1.02 | .515     | -0.08    | 0.04 | 0.93 | .046     |
| Gender D1 (male)               | -1.57    | 0.78 | 0.21 | .045     | 1.47     | 1.17 | 4.37 | .209     |
| Gender D2 (female)             | -1.00    | 0.75 | 0.37 | .179     | 1.48     | 1.14 | 4.40 | .195     |
| Education                      | -0.13    | 0.06 | 0.88 | .035     | -0.08    | 0.07 | 0.92 | .281     |
| Race                           | 0.26     | 0.22 | 1.30 | .236     | -0.56    | 0.25 | 0.57 | .028     |
| Self-harm exposure via other sources | 0.16     | 0.05 | 1.18 | .003     | 0.02     | 0.07 | 1.02 | .787     |
| Self-harming behavior (wave 1) |         |      |      |          | 1.64     | 0.24 | 5.17 | <.001    |
| Instagram exposure D1 (just once) | 0.68     | 0.25 | 1.97 | .007     | 1.58     | 0.30 | 4.87 | <.001    |
| Instagram exposure D2 (more than once) | 1.05     | 0.24 | 2.87 | <.001    | 0.79     | 0.32 | 2.20 | .013     |

SE: standard error; OR: odds ratio; \(R^2\): Nagelkerke’s \(R^2\); D: dummy variable; cross-sectional evidence full model: \(\Delta \chi^2 = 55.52, R^2 = .124, p < .001\); panel evidence full model: \(\Delta \chi^2 = 110.74, R^2 = .270, p < .001\). This table only reports step 2 of the hierarchical model (i.e. full model). Both Instagram dummies were added in step 2. Therefore, the change in \(\chi^2\) of step 2 indicates whether Instagram exposure explains additional variance.

\(\Delta R^2 = .071, p < .001\); hopelessness, \(\Delta F(2, 626) = 8.92, \Delta R^2 = .027, p < .001\); and reasons for living, \(\Delta F(2, 627) = 5.96, \Delta R^2 = .018, p = .003\). Exposure showed no statistical relationship with suicide plan, \(\Delta F(2, 237) = 1.88, \Delta R^2 = .015, p = .154\). This indicates that there were cross-sectional relationships, even after controlling for age, gender, education, race, and exposure to self-harm via other media. This cross-sectional evidence is consistent with \(H1, H2, H4, H5,\) and \(H6\). However, causal claims cannot be made. It is possible, for example, that (1) vulnerable individuals tend to expose themselves to self-harm, that (2) those who view self-harm content (due to whatever reasons) show an increase in self-harm as a consequence of exposure, or that (3) a more complex, interactive model had been operating.

Panel evidence. We now extend this analysis by including the wave 2 measures. Here we predicted each outcome (wave 2) from its baseline score (wave 1), exposure to self-harm on Instagram (wave 1), and control variables. All outcomes measured at wave 2 measured “previous month” or “since the first survey” prevalence. Thus, lifetime exposure to self-harm was used to predict prevalence of self-harming behavior, suicidal ideation, and suicide plan in the time period between both panel waves (while simultaneously controlling for lifetime prevalence of the same outcomes measured at wave 1). As already noted,
hopelessness, reasons for living, and suicide risk are measured with scales that focus on the very moment when filling out the wave 2 questionnaire.

As Tables 1 to 6 (column: “Panel evidence”) show, exposure to self-harm on Instagram was positively related to an increase in self-harming behavior $\Delta \chi^2 = 28.10, \Delta R^2 = .064, p < .001$; suicidal ideation $\Delta F(2, 550) = 12.22, \Delta R^2 = .037, p < .001$; suicide risk $\Delta F(2, 622) = 3.77, \Delta R^2 = .007, p = .024$; and hopelessness $\Delta F(2, 623) = 5.74, \Delta R^2 = .011, p = .003$. Furthermore, exposure to self-harm on Instagram was related to a decrease in reasons for living $\Delta F(2, 625) = 6.35, \Delta R^2 = .015, p = .002$. Conversely, exposure showed no effect on suicide plan $\Delta F(2, 84) = 0.27, \Delta R^2 = .005, p = .767$.

It is possible that vulnerable individuals generally showed an increase in self-harm and suicidality-related outcomes and exposure to self-harm on Instagram. According to this idea, prior vulnerability (indicated by an individual’s suicide risk) would be causally responsible for an increase in both, exposure and other target outcomes. In order to address this alternative model, we reran all “panel evidence” regression models (see Tables 1 to 6, right column) additionally controlling for initial suicide risk (wave 1) as an indicator for prior vulnerability in the first step of the hierarchical models. This analysis helps to separate the effects of prior vulnerability from the effects of exposure to self-harm on Instagram. Consistent with the results reported above, exposure was again positively related to increases in self-harming behavior $\Delta \chi^2 = 20.80, \Delta R^2 = .045, p < .001$;
suicidal ideation $\Delta F(2, 547) = 9.23, \Delta R^2 = 0.025, p < .001$; and hopelessness $\Delta F(2, 620) = 4.08, \Delta R^2 = 0.007, p = .017$, as well as negatively related to reasons for living $\Delta F(2, 622) = 5.67, \Delta R^2 = 0.013, p = .004$. Exposure was unrelated to suicide plan, $\Delta F(2, 82) = 0.04, \Delta R^2 = 0.001, p = .966$.

**Additional analysis: accidental versus intentional exposure**

Analyses showed that a high number of those who get exposed to self-harm on Instagram came across this content by accident. To get a deeper understanding of this finding, we now report (unplanned) additional analyses: Although previous research has already emphasized the problematic phenomenon that exposure to self-harm and suicide on Instagram can happen accidentally (e.g. see the “cat” example given earlier, Moreno et al., 2016), the minority of participants among those who viewed self-harm content indicated intentional exposure. Thus, accidental exposure seems to play an important role, for which we now provide additional analyses.

In a first step, we tested bivariate correlations between intentional exposure and self-harm and suicidality-related outcomes (all from wave 1). We found that intentional exposure (i.e. dichotomous variable; intentional exposure: no, yes) was positively related to hopelessness $r(310) = .22, p < .001$ and suicide risk $r(310) = .17, p < .001$ and negatively

---

**Table 3. Predicting suicide plan: hierarchical multiple regression (step 2: full model).**

| Outcome = suicide plan | Cross-sectional evidence (predicting wave 1 outcome) | Panel evidence (predicting wave 2 outcome) |
|------------------------|-------------------------------------------------------|--------------------------------------------|
|                        | $B$ $SE$ $\beta$ $p$                                  | $B$ $SE$ $\beta$ $p$                        |
| **Step 2 (effect of Instagram exposure)** | $\Delta F(2, 237) = 1.88, \Delta R^2 = .015, p = .154$ | $\Delta F(2, 84) = 0.27, \Delta R^2 = .005, p = .767$ |
| Age                    | $-0.03$ $0.02$ $-0.10$ $.144$                        | $-0.03$ $0.04$ $-0.09$ $.383$               |
| Gender $D1$ (male)     | $0.13$ $0.53$ $0.05$ $.808$                         | $-0.69$ $0.79$ $-0.25$ $.388$              |
| Gender $D2$ (female)   | $0.14$ $0.51$ $0.05$ $.789$                         | $-0.28$ $0.75$ $-0.11$ $.710$              |
| Education              | $-0.05$ $0.04$ $-0.09$ $.184$                        | $0.02$ $0.07$ $0.03$ $.767$                |
| Race                   | $0.05$ $0.14$ $0.02$ $.708$                         | $-0.71$ $0.23$ $-0.31$ $.003$              |
| Self-harm exposure via other sources | $0.08$ $0.04$ $0.15$ $.032$ | $0.02$ $0.07$ $0.04$ $.746$ |
| Suicide plan (wave 1)  | — — — —                                              | $0.18$ $0.11$ $0.18$ $.095$               |
| Instagram exposure $D1$ (just once) | $0.33$ $0.17$ $0.14$ $.054$ | $0.09$ $0.28$ $0.04$ $.744$ |
| Instagram exposure $D2$ (more than once) | $0.17$ $0.17$ $0.08$ $.322$ | $0.22$ $0.31$ $0.10$ $.471$ |

$SE$: standard error; $D$: dummy variable; cross-sectional evidence full model: $F(8, 237) = 2.31, R^2 = .072, p = .021$; panel evidence full model: $F(9, 84) = 1.99, R^2 = .175, p = .051$.

This table only reports step 2 of the hierarchical model (i.e. full model). Both Instagram dummies were added in step 2. Therefore, the change in $R^2$ of step 2 indicates whether Instagram exposure explains additional variance.
related to reasons for living $r(311) = -0.30$, $p < .001$. There were no substantial correlations with own self-harming behavior $r(301) = 0.03$, $p = .590$; suicidal ideation $r(287) = 0.05$, $p = .398$; and suicide plan $r(147) = .10$, $p = .218$. Although not all coefficients were significant, there is preliminary evidence that those scoring higher on self-harm and suicidality-related outcomes tend to intentionally search for self-harm content on Instagram to a higher extent. The pattern of (significant and non-significant) correlations may suggest the interpretation that at-risk individuals are more likely to intentionally access self-harm content on Instagram but those with more intentional exposure are not more likely to do anything immediately about it.

In a second step, we tested whether the hypothesized effects hold true among those who accidentally came across self-harm on Instagram. As an additional analysis, we reran the regression models (panel evidence). Conversely to the models reported above, we used two dummies with non-exposure as the referent and the two types of exposure (accidental, intentional) as separate predictors. Thus, this analysis tested the difference between those with no self-harm exposure versus those with accidental exposure versus those with intentional exposure. Both dummies (dummy 1: effect of accidental exposure; dummy 2: effect of intentional exposure) elicited similar significant effects among the target outcomes: self-harming behavior $\Delta \chi^2 = 23.44$, $\Delta R^2 = 0.047$, $p < .001$ ($p_{\text{acc}} = .003$, $p_{\text{int}} < .001$); suicidal ideation $\Delta F(2, 605) = 12.13$, $\Delta R^2 = 0.033$, $p < .001$ ($p_{\text{acc}} = .038$, $p_{\text{int}} < .001$).
Table 5. Predicting hopelessness: hierarchical multiple regression (step 2: full model).

| Outcome = hopelessness | Cross-sectional evidence (predicting wave 1 outcome) | Panel evidence (predicting wave 2 outcome) |
|------------------------|-----------------------------------------------------|------------------------------------------|
|                        | B         | SE       | SE       | Β       | p         | B         | SE       | Β       | p         |
| Age                    | -0.02     | 0.01     | -.06     | .125    | 0.02      | 0.01      | .05      | .143    |
| Gender D1 (male)       | -0.32     | 0.39     | -1.10    | .419    | 0.03      | 0.34      | .01      | .940    |
| Gender D2 (female)     | -0.46     | 0.38     | -1.16    | .225    | -0.14     | 0.33      | -0.04    | .675    |
| Education              | -0.09     | 0.10     | -1.12    | .003    | -0.01     | 0.08      | -0.02    | .569    |
| Race                   | 0.07      | 0.10     | .03      | .519    | 0.02      | 0.08      | .01      | .772    |
| Self-harm exposure via other sources | -0.01  | 0.03     | -0.01    | .814    | -0.02     | 0.02      | -0.03    | .448    |
| Hopelessness (wave 1)  | —         | —        | —        | —       | 0.65      | 0.03      | .62      | <.001   |
| Instagram exposure D1 (just once) | 0.31  | 0.12     | .11      | .011    | 0.30      | 0.10      | .10      | .003    |
| Instagram exposure D2 (more than once) | 0.48  | 0.12     | .19      | <.001   | 0.25      | 0.10      | .10      | .010    |

SE: standard error; D = dummy variable; cross-sectional evidence full model: $F(8, 626) = 5.21$, $R^2 = .062$, $p < .001$; panel evidence full model: $F(9, 623) = 50.79$, $R^2 = .423$, $p < .001$.

This table only reports step 2 of the hierarchical model (i.e. full model). Both Instagram dummies were added in step 2. Therefore, the change in $R^2$ of step 2 indicates whether Instagram exposure explains additional variance.

$p_{int} < .001$); suicide risk $ΔF(2, 709) = 4.88$, $ΔR^2 = 0.008$, $p = .008$ ($p_{acc} = .046$, $p_{int} < .004$); hopelessness $ΔF(2, 712) = 7.00$, $ΔR^2 = 0.012$, $p = .001$ ($p_{acc} = .001$, $p_{int} = .009$); and reasons for living $ΔF(2, 714) = 3.83$, $ΔR^2 = 0.008$, $p = .022$ ($p_{acc} = .009$, $p_{int} = .111$). Consistent with the analysis reported above, there was no significant effect for suicide plan $ΔF(2, 93) = 0.28$, $ΔR^2 = 0.005$, $p = .759$ ($p_{acc} = .635$, $p_{int} = .467$). Full models can be obtained upon request. Taken together, this additional analysis indicates that exposure to self-harm on Instagram was related to harmful outcomes even among those who accidentally came across self-harm content.

Discussion

Instagram is a very popular social networking site among youth with about a billion active users per month (Instagram, 2018). Scholars have expressed the concern of possible harmful effects (e.g. Arendt, 2019; Brown et al., 2018). We tested relationships between exposure to self-harm on Instagram with a total of six different outcomes. These outcomes ranged from a user’s own self-harming behavior, suicidal ideation, and specific suicide plans, to more general suicide risk, hopelessness, and reasons for living. We focused on this broad range of self-harm and suicidality-related outcomes to reduce the likelihood of overlooking important relationships. Of interest, we found that those who
viewed self-harm on Instagram during their lifetime tended to show more self-harm and suicidality-related outcomes. At a most basic level, the present study thus shows that young people who get exposed to self-harm on social media (either intentionally or by accident) are at higher risk of self-harm or suicide either because of the exposure itself or because they are at higher risk to begin with and thus more likely to run into this problematic content. Either way, findings suggest that the self-harm content on Instagram should be a source of concern. Clearly, the findings we provide are an advancement beyond what has been studied so far and emphasize the importance of considering Instagram in the suicide prevention domain.

We also provided panel evidence aiming to get a deeper understanding of the dynamic process of the complex and intertwined relationship between suicidality and media exposure. This analysis indicated that those who were exposed to self-harm on Instagram sometimes during their life (measured at wave 1) tended to show more self-harm and suicidality-related outcomes at the second wave compared to those who were not exposed at wave 1. If selection of content on Instagram were the only reason for the association at wave 1, there would be no reason to expect the relation to continue at wave 2. An alternative explanation is that our outcome measures lacked sufficient reliability to enable their control at wave 1. However, we found evidence of exposure effects even for measures with high reliability at both waves (suicide risk and reasons for living). It is

| Table 6. Predicting reasons for living: hierarchical multiple regression (step 2: full model). |
|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| Outcome = reasons for living                  | Cross-sectional evidence (predicting wave 1 outcome) | Panel evidence (predicting wave 2 outcome) |
|                                               | B     | SE   | β     | p     | B     | SE   | β     | p     |
| Step 2 (effect of Instagram exposure)        |       |      |       |       |       |      |       |       |
| ΔF(2, 627) = 5.96,                            |       |      |       |       | ΔF(2, 625) = 6.35,                           |       |      |       |       |
| ΔR² = .018, p = .003                          |       |      |       |       | ΔR² = .015, p = .002                          |       |      |       |       |
| Age                                           | 0.04  | 0.02 | .09   | .031  | −0.01 | 0.02 | −.01  | .754  |
| Gender D1 (male)                              | 0.22  | 0.53 | .05   | .679  | 0.32  | 0.44 | .08   | .472  |
| Gender D2 (female)                            | 0.38  | 0.51 | .10   | .454  | 0.57  | 0.43 | .15   | .178  |
| Education                                     | 0.11  | 0.04 | .11   | .005  | 0.06  | 0.03 | .06   | .072  |
| Race                                          | 0.50  | 0.14 | .14   | <.001 | 0.00  | 0.12 | 0.00  | .973  |
| Self-harm exposure via other sources          | 0.05  | 0.03 | .07   | .134  | 0.08  | 0.03 | .10   | .008  |
| Reasons for living (wave 1)                   |       |      |       |       |       |      |       |       |
| Instagram exposure D1 (just once)             | −0.52 | 0.16 | −.13  | .002  | −0.49 | 0.14 | −.14  | <.001 |
| Instagram exposure D2 (more than once)        | −0.39 | 0.16 | −.11  | .015  | −0.21 | 0.13 | −.07  | .107  |

SE: standard error; D: dummy variable; cross-sectional evidence full model: F(8, 627) = 6.18, R² = .073, p < .001; panel evidence full model: F(9, 625) = 24.10, R² = .258, p < .001. This table only reports step 2 of the hierarchical model (i.e. full model). Both Instagram dummies were added in step 2. Therefore, the change in R² of step 2 indicates whether Instagram exposure explains additional variance.
thus likely that some of the self-harm exposures (measured at wave 1) were quite recent, and these exposures did not result in changes until the time period after wave 1. Taken together, our findings indicate that exposure seems heightened among those at greater risk of suicide, and thus may add to the risk. This conceptual extension onto Instagram thus provides an important contribution to the “media and suicide” literature that speaks to the cross-platform stability of possible negative effects.

Of interest, exposure to self-harm content on Instagram often occurred accidentally (i.e. without the intention to view it). Analyses showed, however, that predicted statistical relationships could be observed even for those who accidentally came across self-harm content. This is important given scholarly warnings that exposure to self-harm and suicide on Instagram can happen accidentally (see the “cat” example given earlier, Moreno et al., 2016). Exposure to self-harm and suicide-related content on Instagram could be particularly dangerous even for those who are only peripherally involved with this topic, which underlines the relevance of the found effects of accidental exposure. However, more research is needed to get a more thorough theoretical understanding of the role of accidental and intentional exposure.

Instagram is a global company and content posted on publicly available self–harm-related hashtags is potentially accessible for a global audience. Of interest, Instagram has recently announced that it will put more effort in managing access to self-harm content. Albeit the announcement, no further details about how and when the platform will have achieved this goal are available (Instagram, 2019b). Adam Mosseri, Head of Instagram, noted recently that “over the past month we have seen that we are not where we need to be on self-harm and suicide, and that we need to do more to keep the most vulnerable people who use Instagram safe” (Instagram, 2019a). In fact, Instagram announced aiming to ban graphic and non-graphic images of self-harm (e.g. cutting); at the same time Instagram plans to include more resources for users who post or search self–harm-related content and try to direct them to professional mental health organizations. Importantly, as Instagram (2019a) has acknowledged, it is difficult to find the right balance between creating safe spaces for users to talk about their experiences—talking can be an important first step in overcoming a suicidal crisis (Sonneck et al., 2012)—while also reducing or even eliminating the likelihood of imitational effects of self-harm content (on others).

The company deserves credit for acknowledging the important role of self–harm-related content. The cross-sectional and panel evidence of the present study emphasizes the need for action as it “will take time” and this is “not as simple as just switching off a button” (Instagram, 2019a). Although Instagram has not yet publicly communicated their detailed intervention strategy, mere censoring of specific hashtags (such as #suicide) or banning all explicit cutting images is an important step. However, many users encounter self-harm content accidentally, as shown in the present study. To accomplish this goal, self-harm will have to be accurately identified, potentially using computer vision and machine learning techniques (see Scherr et al., 2019, for an automatic image recognition algorithm to discover cutting-related posts). However, tackling this complex and problematic issue with artificial intelligence will also require more research.
Limitations

This study has several limitations. First and of utmost importance, although a panel design can increase the confidence into the causal interpretations, panel analysis cannot prove causality. We controlled for a number of third variables, but other unconsidered variables may have influenced the target outcomes. In addition, the nature of the self-harm and suicidality-related outcomes—established measures that we have taken from literature—may have introduced problems related to synchronicity, a problem often observed in panel designs. Synchronicity means that self-harm exposure and the respective target outcome are measured at the same time. Although synchronicity seems to be an easily fulfilled assumption given that we measured exposure and all target outcomes within the same survey questionnaire at panel wave 1, problems of retrospection and aggregation may have resulted in a low level of synchronicity (see Kenny, 1975, for a detailed discussion of the problem of synchronicity in panel designs). In fact, participants had to retrospectively recall their exposure to self-harm on Instagram and assess their (aggregated) frequency of exposure (just once or more than once). In addition, we asked participants, for example, about the lifetime prevalence of suicidal ideation at wave 1. Importantly, we (and very likely lots of participants as well) do not exactly know when and how often such thoughts have occurred.

Second, the data set of the present study is part of a larger project on the role of the new media in the self-harm and suicide domain. This larger project required us to restrict the sample to participants with Netflix access. Although Netflix access has no conceptual relevance for the present study, this decreases generalizability. Third, we relied on a convenience sample to test our hypothesis. For example, the overrepresentation of women decreases the generalizability of our findings. Importantly, gender is an important epidemiological factor for suicide prevention. Although the majority of suicide attempters are women, the majority of suicide deaths are in men (Sonneck et al., 2012). The overrepresentation of women in our sample might therefore lead to a bias toward an increased vulnerability for (non-suicidal) self-harm. However, although both women and men regularly appear in posted Instagram images, there is a female dominance in the overall visuals posted on the platform (see Arendt, 2019). In addition, we restricted the sample to participants aged 18 years or older. It is important to consider that younger users not only use social media more but are also presumably especially vulnerable to negative effects of exposure. Nevertheless, suicide deaths are typically in older people (Sonneck et al., 2012). Thus, generalizing our findings to older populations should only be made cautiously, as self-harm-related content on social media may elicit different effects in these age groups. Similarly, we used a non-clinical sample from the general population. A clinical sample (e.g. major depression) may be even more vulnerable to social media influences. Fourth, although we did not find evidence for dose–response relationships (Arendt, 2013, 2015), indicated by more frequent exposure to self-harm content (i.e. none vs just once vs more than once), our exposure measure may not have been sensitive enough to reveal dose–response effects (see Etzersdorfer et al., 2004, for a study on dose–response relationships in the media and suicide domain). Fifth, only 20.1% indicated to have intentionally searched for self-harm content. This low
A proportion of participants could have led to unstable results in our additional analysis related to accidental exposure.

**Conclusion**

Despite these limitations, the present study shows that many users are accidentally or intentionally exposed to self-harm content on Instagram, that exposure to this content elicited emotional disturbance in some users, and that exposure was statistically related to (possibly harmful) self-harm and suicidality-related outcomes. More studies that further test causal interpretations are urgently needed. As research progresses, however, more and more evidence for detrimental (or even beneficial) effects will accumulate so that we will hopefully better understand—as society—how to adequately react on time to problematic content on social media.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the Annenberg Public Policy Center of the University of Pennsylvania.

**ORCID iDs**

Florian Arendt [ID](https://orcid.org/0000-0003-1107-8682)
Sebastian Scherr [ID](https://orcid.org/0000-0003-4730-1575)

**Note**

1. We provide a definition of key terms to increase clarity: Deliberate self-harm behaviors include self-damaging acts both with and without a suicidal intention (Muehlenkamp et al., 2012). Non-suicidal self-injury (NSSI) is used for deliberate self-harming behaviors without a suicidal intention; conversely, suicidal behavior refers to self-damaging acts with a suicidal intention (Nock and Favazza, 2009; Nock et al., 2008). Suicidality is a broad concept, including (more or less serious) suicide-related thoughts (termed suicidal ideation), (more or less serious) suicide-related behavioral intentions (termed suicide plan), and suicide attempt, that is, “the tendency to suicide—from thought to attempt” (Sonneck et al., 2012: 231).

**References**

Arendt F (2013) Dose-dependent media priming effects of stereotypic newspaper articles on implicit and explicit stereotypes. *Journal of Communication* 63: 830–851.

Arendt F (2015) Toward a dose-response account of media priming. *Communication Research* 42: 1089–1115.

Arendt F (2019) Suicide on Instagram: content analysis of a German suicide-related hashtag. *Crisis* 40: 36–41.

Arendt F, Scherr S, Pasek J, et al. (2019) Investigating harmful and helpful effects of watching season 2 of 13 Reasons Why: results of a two-wave U.S. panel survey. *Social Science & Medicine*. Advance online publication. Retrieved from https://doi.org/10.1016/j.socsci-med.2019.04.007

Baker T and Lewis S (2013) Responses to online photographs of non-suicidal self-injury: a thematic analysis. *Archives of Suicide Research* 17: 223–235.
Beck AT, Steer R, Kovacs M, et al. (1985) Hopelessness and eventual suicide: a 10-years prospective study of patients hospitalized with suicidal ideation. *American Journal of Psychiatry* 142: 559–563.

Brown R, Fischer T, Goldwich, et al. (2018) #cutting: non-suicidal self-injury (NSSI) on Instagram. *Psychological Medicine* 48: 337–346.

Carlyle K, Guidry J, Williams K, et al. (2018) Suicide conversations on Instagram™: contagion or caring? *Journal of Communication in Healthcare* 11(1): 12–18.

Chandler A, King C, Burton C, et al. (2015) General practitioners’ accounts of patients who have self-harmed: a qualitative, observational study. *Crisis* 37(1): 42–50.

Cwik J, Siegmann P and Willutzki U (2017) Brief reasons for living inventory: a psychometric investigation. *BMC Psychiatry* 17: 358.

Dyson MP, Hartling L, Shulhan J, et al. (2016) A systematic review of social media use to discuss and view deliberate self-harm acts. *PLoS ONE* 11(5): e0155813.

Eltz M, Evans AS, Celio M, et al. (2007) Suicide probability scale and its utility with adolescent psychiatric patients. *Child Psychiatry and Human Development* 38: 17–29.

Etzersdorfer E, Voracek M and Sonneck G (2004) A dose-response relationship between imitative suicides and newspaper distribution. *Archives of Suicide Research* 8: 137–145.

Instagram (2018) Instagram Statistics. Retrieved from https://instagram-press.com/our-story/Instagram (2019a) Changes We’re Making to Do More to Support and Protect the Most Vulnerable People Who Use Instagram. Retrieved from https://instagram-press.com/blog/2019/02/07/changes-were-making-to-do-more-to-support-and-protect-the-most-vulnerable-people-who-use-instagram/

Instagram (2019b) Partnering with Experts to Protect People from Self-harm and Suicide. Retrieved from https://newsroom.fb.com/news/2019/02/protecting-people-from-self-harm/

Jacob N, Evans R and Scourfield J (2017) The influence of online images on self-harm: a qualitative study of young people aged 16–24. *Journal of Adolescence* 60: 140–147.

Kann L, McManus T, Harris WA, et al. (2018) Youth risk behavior surveillance—United States, 2017. *MMWR* 67(8): 1–114.

Kenny D (1975) Cross-lagged panel correlation: a test for spuriousness. *Psychological Bulletin* 82: 887–903.

Klonsky E, Kotov R, Bakst S, et al. (2012) Hopelessness as a predictor of attempted suicide among first admission patients with psychosis: a 10-year cohort study. *Suicide & Life-threatening Behavior* 42(1): 1–10.

Klonsky E, Victor S and Saffer B (2014) Nonsuicidal self-injury: what we know, and what we need to know. *The Canadian Journal of Psychiatry* 59: 565–568.

Luo X, Wang Q, Wang X, et al. (2016) Reasons for living and hope as the protective factors against suicidality in Chinese patients with depression: a cross sectional study. *BMC Psychiatry* 16: 252.

Mitchell KJ and Ybarra ML (2007) Online behavior of youth who engage in self-harm provides clues for preventive intervention. *Preventive Medicine* 45(5): 392–396.

Moreno M, Ton A, Selkie A, et al. (2016) Secret society 123: understanding the language of self-harm on Instagram. *Journal of Adolescent Health* 58: 78–84.

Muehlenkamp JJ, Claes L, Havertape L, et al. (2012) International prevalence of adolescent nonsuicidal self-injury and deliberate self-harm. *Child and Adolescent Psychiatry and Mental Health* 6: 10.

Niederkrotenthaler T and Stack S (2017) *Media and Suicide: International Perspectives on Research, Theory, and Policy*. New York: Transaction Publishers.

Niederkrotenthaler T, Till B, Kapusta ND, et al. (2009) Copycat effects after media reports on suicide: a population-based ecologic study. *Social Science and Medicine* 69(7): 1085–1090.
Nock M and Favazza A (2009) Understanding nonsuicidal self-injury: origins, assessment and treatment. In: Nock M (ed.) Nonsuicidal Self-Injury: Definition and Classification. Washington, DC: American Psychological Association, pp. 9–18.

Nock M, Borges G, Bromet E, et al. (2008) Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. The British Journal of Psychiatry 192: 98–105.

Pew Research Center (2012) Assessing the representativeness of public opinion surveys. Available at: http://www.people-press.org/2012/05/15/assessing-the-representativeness-of-public-opinion-surveys/ (accessed 1 January 2019).

Pew Research Center (2015) Teens, Social Media and Technology Overview 2015. Available at: http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/ (accessed 1 January 2019).

Phillips DP (1974) The influence of suggestion on suicide: substantive and theoretical implications of the Werther effect. American Sociological Review 39: 340–354.

Pirkis J and Blood RW (2001) Suicide and the media: part I. Reportage in nonfictional media. Crisis 22(4): 146–154.

Pirkis J, Mok K and Robinson J (2017) Suicide and newer media: the good, the bad, and the Googly. In: Niederkrotenthaler T and Stack S (eds) Media and Suicide: International Perspectives on Research, Theory, and Policy. New York: Transaction Publishers, pp. 87–98.

Reid-Chassiakos Y, Radesky J, Christakis D, et al. (2016) Children and adolescents and digital media. Pediatrics 138: e30162593.

Romer D, Jamieson P and Jamieson KH (2006) Are news reports of suicide contagious? A stringent test in six US cities. Journal of Communication 56: 253–270.

Scherr S and Reinemann C (2011) Belief in a Werther effect: third-person effects in the perceptions of suicide risk for others and the moderating role of depression. Suicide and Life-threatening Behavior 41(6): 624–634.

Scherr S, Arendt F, Frissen T, et al. (2019) Detecting intentional self-harm on Instagram: development, testing, and validation of an automatic image-recognition algorithm to discover cutting-related posts. Social Science Computer Review. Epub ahead of print 18 March. DOI: 10.1177/0894439319836389.

Schmidtte A and Häfner H (1988) The Werther effect after television films: new evidence for an old hypothesis. Psychological Medicine 18(3): 665–676.

Sisask M and Värnik A (2012) Media roles in suicide prevention: a systematic review. International Journal of Environmental Research and Public Health 9: 123–138.

Sonneck G, Kapusta N, Tomandl G, et al. (2012) Krisenintervention Und Suizidverhütung [Crisis intervention and suicide prevention]. Vienna, Austria: UTB.

Stack S (2003) Media coverage as a risk factor in suicide. Journal of Epidemiology and Community Health 57: 238–240.

Stack S (2005) Suicide in the media: a quantitative review of studies based on nonfictional stories. Suicide and Life-threatening Behavior 35: 121–133.

Till B, Tran US, Voracek M, et al. (2017) Papageno vs. Werther effect online: randomized controlled trial of beneficial and harmful impacts of educative suicide prevention websites. British Journal of Psychiatry 211: 109–115.

Vente T, Daley M, Killmeyer E, et al. (2017) Social media as a vehicle for expression of self-harm and risk-taking behavior in adolescents. Journal of the American Academy of Child and Adolescent Psychiatry 56(10): S158.

Wasserman IM, Stack S and Reeves JL (1994) Suicide and the media: the New York Times’s presentation of front-page suicide stories between 1910 and 1920. Journal of Communication 44(2): 64–83.
World Health Organization (WHO) (2018) *Suicide: Key Facts*. Available at: https://www.who.int/en/news-room/fact-sheets/detail/suicide (accessed 1 January 2019).

Yip P and Cheung Y (2006) Quick assessment of hopelessness: a cross-sectional study. *Health and Quality of Life Outcomes* 4: 13–18.

**Author biographies**

Florian Arendt holds the tenure track professorship for Health Communication at the Department of Communication at the University of Vienna, Austria.

Sebastian Scherr is an assistant professor at the School for Mass Communication Research at the University of Leuven (KU Leuven), Belgium.

Daniel Romer is research director of the Annenberg Public Policy Center and director of its Adolescent Communication Institute (ACI) at the University of Pennsylvania, USA.