Investigating Google's suicide-prevention efforts in celebrity suicides using agent-based testing: A cross-national study in four European countries

Florian Arendt[a,*], Mario Haim[b], Sebastian Scherr[c]

[a] Department of Communication, University of Vienna, Austria
[b] Department of Media and Social Sciences, University of Stavanger, Norway
[c] School for Mass Communication Research, University of Leuven (KU), Belgium

Rationale: Google can act as a "gatekeeper" for individuals who seek suicide-related information online (e.g., "how to kill oneself"). The search engine displays a "suicide-prevention result" (SPR) at the very top of some suicide-related search results. This SPR comes as an info box and contains supposedly helpful crisis help information such as references to a telephone counseling service.
Objective: It remains unknown, however, how Google has implemented the SPR in the especially dangerous context of celebrity suicide for which imitational copycat suicides in vulnerable individuals are most likely.
Method: Relying on agent-based testing, a computational social science method, we emulated a total of 137,937 Google searches in April 2019, using both general suicide-related and specific celebrity suicide-related search terms. Given the recently discovered language-based differences in SPR display rates, we held the language constant and focused on German-speaking populations in four European countries.
Results: The SPR was never shown in searches for celebrities who died by suicide in all four countries. Furthermore, analyses indicated a digital divide in access to suicide-prevention information with moderately high SPR display rates in Germany and Switzerland, yet with no display in Austria and Belgium.
Conclusion: Higher SPR display rates could support global suicide-prevention efforts at virtually no cost by providing preventive information to vulnerable users precisely at the moment when it is apparently needed.

1. Introduction

Suicide is a severe threat to global public health (Wasserman, 2016; World Health Organization, 2019) and the media are perceived to be a key factor in eliciting suicide-protective as well as harmful imitation effects (Mann et al., 2005; Niederkrotenthaler et al., 2010). Although most of the research on media and suicide has focused on media in traditional media environments largely dominated by television and newspapers more recent studies have focused more on the implications of new digital media environments (e.g., Arendt and Scherr, 2017a, b; Arendt et al., 2019a, b; Fahey et al., 2018; Scherr and Reinemann, 2016; Till et al., 2017; see also Biddle et al., 2008). Regarding the latter, Internet search engines, such as global market leader Google, play an increasingly important role (Arendt and Scherr, 2017a; Scherr et al., 2019a, b).

Of note, Google acknowledges the importance of digital media environments for suicide prevention and has implemented a self proclaimed suicide-prevention result (SPR; Zeiger, 2010, see a screenshot in the online supplement). This info box, presented at the top of the list of search results, shows important online and offline help resources for individuals suffering from a suicidal crisis, including country-specific helpline telephone numbers or crisis chat rooms. Existing evidence demonstrates that such referrals are important as they can elicit beneficial effects on their users (Lester and Rogers, 2012); albeit more research in this regard is needed. Of particular interest, people seek professional help services more often when the media provide the sources (Kennedy et al., 2004; Sonneck et al., 2007; Zeiger, 2010). Therefore, high display frequencies of the SPR can be deemed beneficial from the perspective of suicide prevention.

Search engines are an important distributor of suicide-protective information, acting as a "gatekeeper" (Shoemaker and Vos, 2009) for preventive (mental) health information. Unfortunately, the display rates of Google's SPR seem to be disparate: the SPR is not equally presented to every user searching for suicide. The company does not disclose the underlying programming of the display algorithm, keeping it a closely guarded secret. Consequently, the most accurate and

* Corresponding author. Department of Communication, University of Vienna, Währinger Straße 29, 1090, Vienna, Austria.
E-mail address: florian.arendt@univie.ac.at (F. Arendt).

ethically justifiable research methodology to investigate SPR display rates is to emulate actual Google searches using "virtual agents" that perform Google searches just like real human beings do (Haim, in press). This practice allows researchers to measure the share of the results displaying the SPR [%] without having knowledge of Google's "algorithmic black box."

One study conducted in Germany (Haim et al., 2017) showed that the SPR was only displayed in one-in-ten Google searches for general suicide-related helpful search terms (e.g. "overcoming suicidal thoughts") and in one-in-four searches for general suicide-related harmful search terms (e.g. "best suicide method"). Another study that followed the same research paradigm (see Scherr et al., 2019a, b) found that the SPR display rates varied across 11 countries depending on the country and the language that was used. For example, in both the United States (92%) and the United Kingdom (92%), the SPR was shown at relatively high display rates after searches for general suicide-related harmful terms. Interestingly, the SPR display rate was substantially lower in Japan (36%), Germany (22%), and South Korea (20%). Strikingly, in a multi-lingual country such as India, the SPR was displayed after searches in some languages (English, Hindi) but not all (Telugu), even when the same search terms were used. Although Google deserves credit for including the SPR in its algorithm, the varying display rates can contribute to a digital divide in access to suicide-preventive information with possibly lethal consequences.

2. The present study

The present study focuses on a topic that has not yet been addressed by previous research. Research has revealed that media reporting about celebrity suicide is especially dangerous and can lead to substantial copycat effects (Chen et al., 2010; Fink et al., 2018; Stack, 1987, 2005). In fact, many users who want to learn more about a celebrity who has died by suicide will use Internet search engines (Arendt and Scherr, 2017b). Importantly, it is unknown whether the search engine presents the SPR in the context of a celebrity suicide. In fact, previous research has mainly focused on general suicide-related search terms (e.g., "how to kill oneself," "best suicide method"). Given that high display frequency rates in searches for celebrity suicide can be considered especially important due to possible strong copycat effects (Niederkrotenthaler et al., 2012), scholarly attention on this question is of utmost importance. Indeed, the SPR seems to be especially consequential in a celebrity-suicide context. Here, it can be understood as a tailored prevention message that appears exactly when vulnerable individuals search for suicide and presents preventive information presumably right at the moment when it is apparently needed—when suicidal thoughts are mentally activated in a suicidal individual's mind. From a global public health perspective, high display frequencies are thus desirable for celebrity-suicide search requests.

Although we already know that there are great country dependent disparities in display frequencies related to general suicide-related (harmful or helpful) terms, we lack knowledge on whether Google presents the SPR when users search for the name of a celebrity who has died by suicide. The investigation of this research question is the primary contribution of the present study, whose results may inform both our theoretical knowledge about online suicide prevention, and search-engine providers on how to adequately implement suicide-prevention efforts.

3. Method

We used an agent-based testing methodology (Haim, in press) to investigate display rates in four European countries with German-speaking populations. In fact, the obtained search results of ordinary users can be emulated with virtual agents. Thus, virtual agent-based testing is a computational social science methodology that can be used to explore the display frequencies of Google's SPR.

We used a four (country: Austria, Belgium, Germany, and Switzerland) × five (search-term category: suicide-unrelated control, general suicide-related harmful, general suicide-related helpful, names of celebrities who died by suicide, names of these celebrities plus suicide cue) design using day-by-day display frequency values as the target outcomes. While Austria, Germany, and Switzerland are countries where German is (one of) the majority language(s), we also included Belgium, where German is an official minority language only spoken by the German-speaking community of Belgium. Given prior findings of strong stability in SPR display rates (Haim et al., 2017), the main outcome is the daily share [%] of search results in which the SPR is displayed to a user. The observation period was one week (see below).

3.1. Five categories of search terms

The virtual agents were programmed to repeatedly search for terms from a list of either (1) suicide-unrelated control terms (e.g., "studieren," which is German for "to study"); (2) general suicide-related helpful terms (e.g., "hilfe bei selbstmordgedanken," which is German for "help in suicidal ideation"); (3) general suicide-related harmful terms (e.g., "wie man sich umbringt," which is German for "how to kill oneself"); (4) names of celebrities who died by suicide without a suicide reference (e.g., "chester bennington"); and (5) names of celebrities who died by suicide with a suicide reference (e.g., "chester bennington selbstmord," which is German for "chester bennington suicide"). The list of general suicide-related terms and the suicide-unrelated control terms were taken from previous studies (Biddle et al., 2008; Scherr et al., 2019a, b; Võ et al., 2009).

For the celebrity-related search term categories, we selected seven celebrities who had all died by suicide relatively recently. The selection was based on an expert discussion on celebrities who were known to the German-speaking populations of all four countries. Four celebrities can be broadly categorized as entertainers (Daniel Küblböck), musicians (Avicii; Chester Bennington), or actors (Robin Williams). We also included a prominent TV chef (Anthony Bourdain), a commercial airplane pilot (Andreas Lubitz), and a war criminal from the early 1990s Yugoslav Wars (Slobodan Praljak). All search terms and short descriptions of the celebrities can be found in the online supplement.

3.2. Software script and procedure

We set up a total of eight servers, two for each country. All eight servers were rented in Frankfurt, Germany, from Amazon Web Services. We then virtually moved each of the eight servers to one of the four countries under investigation using professional VPN tunneling. For multi-lingual countries, the VPN tunnels routed to German-speaking regions. To eliminate any methodological artifacts, we used two different, well-known VPN providers (ExpressVPN; NordVPN)—also for the German servers in order to be consistent. Technological issues prevented data generated through the NordVPN tunnel in Austria from being included in the final results. Of the remaining final seven servers, each had a country-specific geolocation (i.e., IP address) through which the virtual agents could emulate real user behavior from that respective country. We built on the established "ScrapeBot" (Haim, in press, Haim, 2019), which uses Selenium for Firefox to automatically emulate web requests for the agent-based testing task. Thereby, origin detection was further supported through ISO-based, country-specific German-language settings. All scripts have been made fully available online under https://github.com/MarHai/ScrapeBot.

Data collection took place over one week in April 2019 (i.e., from the 15th to the 21st). Each server (1) randomly chose one search-term category (i.e., one agent); (2) launched a fresh Firefox environment without cookies or a web history; then (3) navigated to the country's respective Google instance (e.g., google.at); (4) randomly chose one of the search terms; (5) searched for this term; and (6) logged whether the SPR was presented. Two precautionary measures were taken in order to

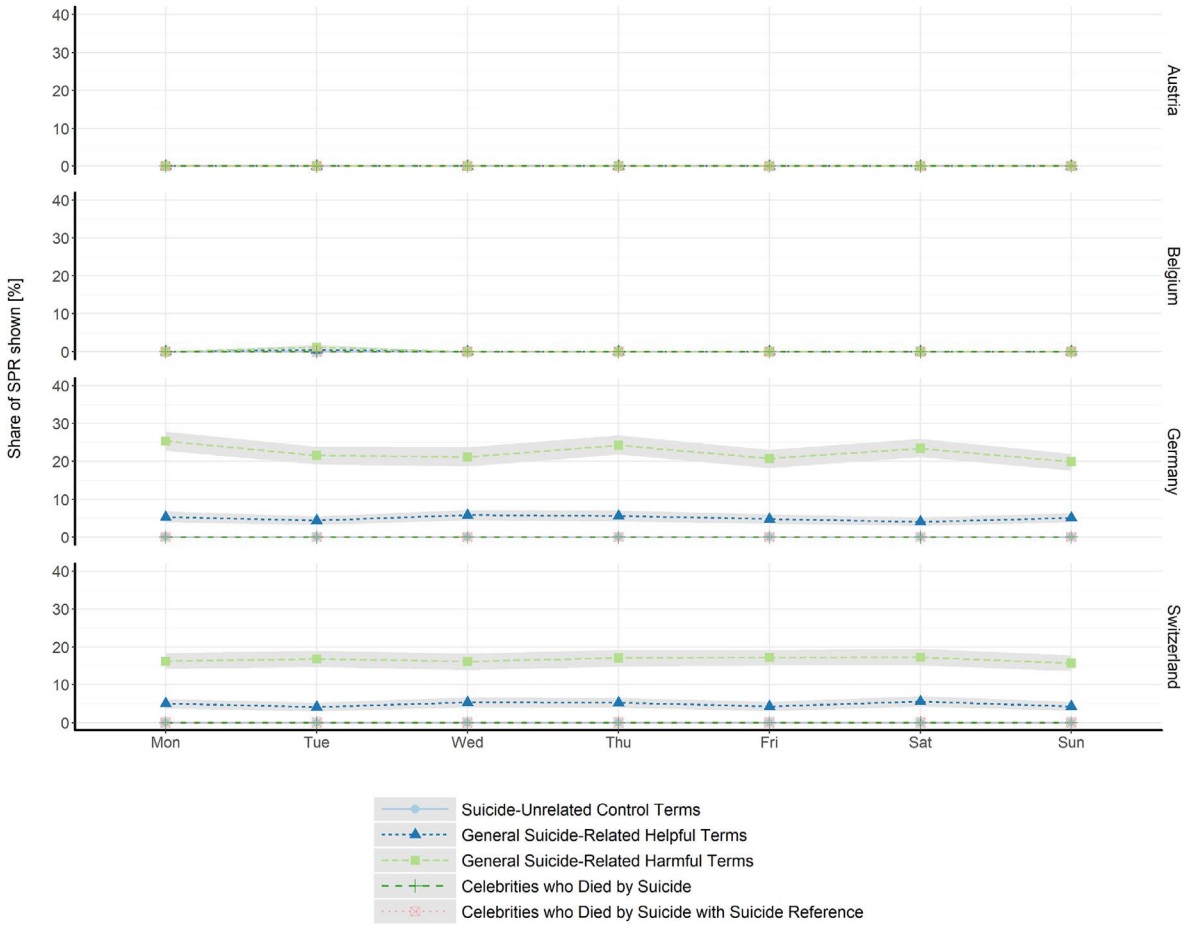

Fig. 1. Display frequencies of Google's suicide-prevention result (SPR) per day in Austria, Belgium, Germany, and Switzerland after using five different categories of search terms, either helpful or harmful, but also with either a general or celebrity reference to suicides, or a suicide-unrelated control group. The gray area represents the percentile-bootstrapped 95% confidence bands of SPR display frequencies based on 1000 replications.

reveal possible malfunctions of the methodological paradigm: (1) the virtual agents randomly took screenshots of the search to validate the findings; (2) at random, agents navigated to an IP-locating service (i.e., ipinfo.io) in order to check the VPN location. After each run, the Firefox environment was terminated, and the procedure started all over again.

In total, this setup yielded $N = 137{,}937$ searches. These searches were almost equally distributed across Belgium ($n = 39{,}254$ searches), Germany ($n = 38{,}942$ searches), and Switzerland ($n = 40{,}848$ searches); in Austria, where we only used one server, approximately half the number of searches were performed; namely, $n = 18{,}893$. All collected data and analyses have been published on the Harvard Dataverse under https://doi.org/10.7910/DVN/VX6QVB.

4. Results

4.1. Preliminary analysis: SPR display rate for general suicide-related searches

As Fig. 1 depicts, the SPR was displayed at a moderately high frequency in Germany and Switzerland but was never displayed in Austria or Belgium. Regarding the latter, there is a small peak on April 16th (see Fig. 1) due to a methodological issue: one Belgian server unexpectedly reverted to its physical server location in Frankfurt, Germany. Although unplanned, this short switch to Germany, a country with higher display frequencies (see Fig. 1), provides further evidence for the idea that the SPR is not equally displayed to all users, and depends on the country where the users are located when they search for their suicide-related information. Addressed simultaneously, this analysis provides evidence for a digital divide.

We predicted that general suicide-related harmful search terms would elicit higher display frequencies compared to general suicide-related helpful search terms. In Germany, harmful terms, $M = 22.3$, 95% *CI* [21.4, 23.2], elicited significantly higher display frequencies compared to helpful terms, $M = 5.0$, 95% *CI* [4.5, 5.5], as indicated by the non-overlapping confidence intervals that are based on bootstrapping. Similarly, in Switzerland, harmful terms, $M = 16.6$, 95% *CI* [15.8, 17.5], elicited significantly higher display frequencies compared to helpful terms, $M = 4.8$, 95% *CI* [4.3, 5.3]. Interestingly, although the display frequencies were almost identical in Germany and Switzerland for helpful terms, display frequencies significantly differed in both countries for harmful terms. Consistent with our expectations, the control terms never led to the display of the SPR.

4.2. Main analysis: SPR display rate for celebrity-suicide searches

As the primary contribution of the present study, we tested whether Google presented the SPR when virtual agents searched for celebrities who had died by suicide without (e.g., “robin williams”) or with a suicide cue (e.g., “robin williams selbstmord”). For a total of $n = 55{,}307$ searches across all four countries related to these two celebrity-related search term categories, strikingly, the SPR was never displayed. This outcome was extremely surprising given that at least the celebrity searches with a suicide cue were predicted to lead to the display of the SPR in countries where the SPR generally shows up (i.e., Germany and Switzerland).

For a more thorough understanding of this finding, we compared

the display frequency rates for the search term “selbstmord” (Engl.: “suicide,” a term from the “general suicide-related harmful” category), with the rates for the combined search term (i.e., celebrity + suicide cue; e.g., “robin williams selbstmord,” Engl.: “robin williams suicide”). Virtual agents who searched for the term suicide alone (“selbstmord”) saw Google's SPR in about half of the searches (Germany: 53.32%, n = 353/662; Switzerland: 49.64%, n = 342/689). Despite this finding, the SPR disappeared entirely when the name of a celebrity was added: the display frequency was zero when users searched for combined search terms (“robin williams selbstmord”). We return to this striking finding in the discussion section.

5. Discussion

Internet search engines are an important gatekeeper for suicide-protective information. Google acknowledges this fact and presents potentially helpful referrals to professional crisis intervention to some of its users. The present study tested whether the search engine presents this preventive information, the SPR, when users also search for celebrities who have died by suicide. References to crisis interventions and suicide-prevention resources are especially relevant for vulnerable users in this context and may help them overcome their suicidal crises. In the four countries with German-speaking populations that this study examined, Google *never* displayed its SPR when users searched for the names of celebrities who had died by suicide. Of interest, the SPR was shown in approximately half of all searches when users only provided the term “selbstmord” (Engl.: “suicide”). Strikingly, the additional inclusion of the name of a celebrity who had died by suicide made the suicide-protective information completely disappear. When users searched for the same term *and* the name of a celebrity who had died by suicide (e.g., “avicii suicide”), the SPR was never displayed. Given that celebrity suicide is an especially important area where detrimental imitation effects are possible, the absence of the SPR in these searches must be deemed especially problematic from the perspective of suicide prevention.

This finding gives us insights into Google's algorithmic decision-making within the area of suicide prevention. Apparently, the previously suggested “search-term” hypothesis (Scherr et al., 2019a, b) finds support here. The best yet simplest explanation is that Google uses profanity filters; that is, keyword lists that define whether or not the SPR will be displayed. It is possible that more complex (machine-learning) algorithms are working in the background as well, but the results of the present study (see the finding related to different display rates of “selbstmord” versus “aviici selbstmord”) are consistent with a simple keyword account. Apparently, only the exact use of certain keywords triggered the display of the SPR—an easy-to-change fact that could help in disseminating helpful and potentially life-saving suicide-prevention information to vulnerable users in times of acute crisis. We provide a full list of key terms (which are based on previous research) in the online supplement. This compilation may help search-engine providers to adapt their algorithmic decision-making accordingly.

Moreover, the present analyses also replicated previous findings on a digital divide in access to mental health information. There were large differences in the display rates of the SPR between the four German-speaking countries. While the SPR was presented at moderately high frequencies in Germany and Switzerland, it never appeared in Austria or Belgium. Thus, it did matter where the vulnerable individuals were when they searched for suicide-related information. Search engines may thus contribute to (global) social inequalities.

The search engine Google deserves praise for starting to implement suicide-prevention information. Nevertheless, the display rate of the SPR found in the present study for Germany still exists in stark contrast to the substantially higher rates in English-speaking countries (Haim et al., 2017; Scherr et al., 2019a, b). The between-country differences in SPR display rates among four countries shed further light on the inequalities in access to preventive (mental) health information. Some countries seem to profit more from opportunities provided by new digital media technologies. Interestingly, country or population size does not seem to matter given that Switzerland is similar to Austria yet has higher SPR display rates.

6. Limitations

The study has limitations. First, the searches made by our virtual agents were not conducted in the direct aftermath of the deaths of the celebrities. We held the observation period constant to ensure that the context was comparable. It is possible that the SPR display frequencies are different shortly after a celebrity suicide. Although we do not expect this given the high temporal stability of the search results (Haim et al., 2017), a future study may wish to test variations directly, subsequent to celebrity deaths by suicide. Second, we restricted the analysis to four countries with German-speaking majority and minority populations. Display rates may certainly be different in countries with other languages (see Scherr et al., 2019a, b). Third, in total, we used seven celebrities who were determined to be known in all four countries. Display frequencies may be different for other celebrities. Given that none of the seven celebrities elicited the display of the SPR, we do not expect different display rates for different celebrities; however, this is an empirical question and thus also up to a future study. Fourth, we used seven consecutive days for the observation period. As already noted, search results are relatively stable over time (Haim et al., 2017). Nevertheless, display frequencies may vary (slightly) over longer time periods.

7. Conclusions

Digital media technologies not only shape our daily lives, but they also play a paramount role for (global) suicide prevention. Offering suicide-prevention information in an info box (SPR) on top of suicide-related search results is a promising way to inform the public about where to find help in an acute suicidal crisis. Currently, Google's underlying programming could be optimized. Search-engine providers clearly have a social responsibility in this regard. By increasing the display frequencies of suicide-protective information, search engines may better contribute to saving the lives of vulnerable individuals.

Author contributions

Florian Arendt: Conceptualization, Methodology, Formal Analysis, Writing-Original draft preparation. **Mario Haim**: Conceptualization, Methodology, Formal Analysis, Writing-Original draft preparation. **Sebastian Scherr**: Conceptualization, Methodology, Formal Analysis, Writing-Original draft preparation.

References

Arendt, F., Scherr, S., 2017a. Optimizing online suicide prevention: a search engine-based tailored approach. Health Commun. 32 (11), 1403–1408. https://doi.org/10.1080/10410236.2016.1224451.

Arendt, F., Scherr, S., 2017b. The impact of a highly publicized celebrity suicide on suicide-related online information seeking. Crisis J. Crisis Interv. Suicide Prev. 38, 207–209. https://doi.org/10.1027/0227-5910/a000455.

Arendt, F., Scherr, S., Pasek, J., Jamieson, P., Romer, D., 2019a. Investigating harmful and helpful effects of watching season 2 of 13 Reasons Why: results of a two-wave U.S. panel survey. Soc. Sci. Med. 232, 489–498. https://doi.org/10.1016/j.socscimed.2019.04.007.

Arendt, F., Scherr, S., Romer, D., 2019b. Effects of Exposure to Self-Harm on Social Media: Evidence from a Two-Wave Panel Study Among Young Adults. New Media & Society. Advance Online Publicationhttps://doi.org/10.1177/1461444819850106.

Biddle, L., Donovan, J., Hawton, K., Kapur, N., Gunnell, D., 2008. Suicide and the

internet. BMJ 336 (7648), 800–802. https://doi.org/10.1136/bmj.39525.442674.AD.
Chen, Y.-Y., Tsai, P.-C., Chen, P.-H., Fan, C.-C., Hung, G.C.-L., Cheng, A.T., 2010. Effect of media reporting of the suicide of a singer in Taiwan: the case of Ivy Li. Soc. Psychiatr. Psychiatr. Epidemiol. 45 (3), 363–369. https://doi.org/10.1007/s00127-009-0075-8.
Fahey, R., Matsubayashi, T., Ueda, M., 2018. Tracking the Werther effect on social media: emotional responses to prominent suicide deaths on twitter and subsequent increases in suicide. Soc. Sci. Med. 219, 19–29. https://doi.org/10.1016/j.socscimed.2018.10.004.
Fink, D.S., Santaella-Tenorio, J., Keyes, K.M., 2018. Increase in suicides the months after the death of Robin Williams in the US. PloS One 13 (2), e0191405. https://doi.org/10.1371/journal.pone.0191405.
Haim, M. (in press): Agent-based testing: an automated approach toward artificial reactions to human behavior. Journal. Stud..
Haim, M., 2019. ScrapeBot. A Selenium-based tool for agent-based testing. Source code Available at: https://github.com/MarHai/ScrapeBot/.
Haim, M., Arendt, F., Scherr, S., 2017. Abyss or shelter? On the relevance of web search engines' search results when people google for suicide. Health Commun. 32, 253–258. https://doi.org/10.1080/10410236.2015.1113484.
Kennedy, M., O'Leary, A., Beck, V., Pollard, K., Simpson, P., 2004. Increases in calls to the CDC national STD and AIDS hotline following AIDS-related episodes in a soap opera. J. Commun. 54, 287–301. https://doi.org/10.1111/j.1460-2466.2004.tb02629.x.
Lester, D., Rogers, J., 2012. Crisis Intervention and Counseling by Telephone and the Internet. Charles C Thomas, Springfield, IL.
Mann, J.J., Apter, A., Bertolote, J., Beautrais, A., Currier, D., Haas, A., ... Marusic, A., 2005. Suicide prevention strategies: a systematic review. Jama 294 (16), 2064–2074. https://doi.org/10.1001/jama.294.16.2064.
Niederkrotenthaler, T., Fu, K.-W., Yip, P.S.F., Fong, D.Y.T., Stack, S., Cheng, Q., Pirkis, J.E., 2012. Changes in suicide rates following media reports on celebrity suicide: a meta-analysis. J. Epidemiol. Community Health 66 (11), 1037–1042. https://doi.org/10.1136/jech-2011-200707.
Niederkrotenthaler, T., Voracek, M., Herberth, A., Till, B., Strauss, M., Etzersdorfer, E., ... Sonneck, G., 2010. Papageno v Werther effect. BMJ Br. Med. J. (Clin. Res. Ed.) 341. https://doi.org/10.1136/bmj.c5841.
Scherr, S., Haim, M., Arendt, F., 2019a. Equal access to online information? Google's suicide-prevention disparities may amplify a global digital divide. New Media Soc. 21 (3), 562–582. https://doi.org/10.1177/1461444818801010.
Scherr, S., Arendt, F., Frissen, T., Oramas, J., 2019b. Detecting intentional self-harm on Instagram: development, testing, and validation of an automatic image-recognition algorithm to discover cutting-related posts. Soc. Sci. Comput. Rev. 21, 562–582. https://doi.org/10.1177/0894439319836389.
Scherr, S., Reinemann, C., 2016. First do no harm: cross-sectional and longitudinal evidence for the impact of individual suicidality on the use of online health forums and support groups. Comput. Hum. Behav. 61, 80–88. https://doi.org/10.1016/j.chb.2016.03.009.
Shoemaker, P., Vos, T., 2009. Gatekeeping Theory. Routledge, New York, NY.
Stack, S., 1987. Celebrities and suicide: a taxonomy and analysis, 1948-1983. Am. Socio. Rev. 52, 401–412.
Stack, S., 2005. Suicide and the media: a quantitative review. Suicide Life-Threatening Behav. 35, 121–133. https://doi.org/10.1521/suli.35.2.121.62877.
Sonneck, G., Kapusta, N., Tomandl, G., Voracek, M., 2007. *Krisenintervention und Suizidverhütung* [Crisis intervention and suicide prevention]. UTB, Vienna, Austria.
Till, B., Tran, U.S., Voracek, M., Niederkrotenthaler, T., 2017. Beneficial and harmful effects of educative suicide prevention websites: randomised controlled trial exploring Papageno v. Werther effects. Br. J. Psychiatr. 211 (2), 109–115. https://doi.org/10.1192/bjp.bp.115.177394.
Võ, M.L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M.J., Jacobs, A.M., 2009. The berlin affective word list reloaded (BAWL-R). Behav. Res. Methods 41, 534–538. https://doi.org/10.3758/BRM.41.2.534.
Wasserman, D., 2016. Suicide: an Unnecessary Death. University Press, Oxford, England.
World Health Organization, 2019. Suicide data. Retrieved from. https://www.who.int/mental_health/prevention/suicide/suicideprevent/en/.
Zeiger, R., 2010. Helping you find emergency information when you need it. November 11, Retrieved from. https://googleblog.blogspot.com/2010/11/helping-you-find-emergency-information.html.