Studying Closed Communities On-line: Digital Methods and Ethical Considerations Beyond Informed Consent and Anonymity

Ylva Hård af Segerstad, Christopher Kullenberg, Dick Kasperowski & Christine Howes

Introduction

In 2014 researchers from Facebook and academia conducted a massive-scale experiment on emotional contagion through social networks (Kramer, Guillory, & Hancock, 2014). In the experiment the news feeds of nearly 700,000 Facebook users were manipulated to see whether positive or negative news affected their emotions. The study sparked an intense discussion of ethical guidelines and informed consent in the international research community. Many argued that it breached ethical guidelines by lacking informed consent (cf. Vitak, Shilton, & Ashktorab, 2016). However, another dimension of this episode concerns the limitations of what researchers are allowed to do as compared to other professionals. The core feature of the Facebook platform is to manipulate content using algorithms to optimize it for marketing purposes.[1] Thus, Facebook is allowed to do for *commercial* purposes what researchers are not allowed to do for *scientific* purposes.

As social media is tremendously rich in data, we argue that the question of ethics must be posed in close proximity to the methods and techniques used. Since social media platforms are constructed with the aim of collecting as much user data as possible, we suggest that foreseeable ethical difficulties can be managed by *reducing* the amount of data collected. The perspective of this chapter is that ethical assessments are realized in the application of methods, and that ethical considerations need to be integrated into the hands-on work of collecting, storing, and analyzing data.

This chapter explores how ethical principles can be used to inform and modify digital methods in order to conduct responsible internet research. Rather than abandoning the possibilities that rich data and digital methods offer in instances where informed consent is problematic, we instead discuss how digital tools can be used to achieve anonymization when studying online communities.

We use the case study of a vulnerable, closed Facebook group to illustrate when and how informed consent and anonymity can be achieved and how digitally generated research data urges us to rethink these notions and their relation to data quality. We make use of Jensen’s distinction between *found data* – online material that can be found, such as images, texts, and other digital traces left by interacting users; and *made data* – data created by an active research intervention, such as surveys, interviews, and participatory field work (Jensen, 2012). This distinction is important for understanding the ethical implications inherent in various types of data.

The chapter is structured as follows: We begin with a presentation of our data, followed by a discussion of the relevant ethical principles. We, then, consider the ethical considerations taken in the methodological process of studying the closed community. The chapter ends with a discussion of informed consent and anonymization in relation to digital methods.

---

[1] https://www.facebook.com/business/news/update-to-facebook-news-feed (accessed April 21, 2016)

The Case: A Closed Facebook Group

Our case study is a Facebook discussion group for a community of bereaved parents coping with the loss of a child. This group originates from a physical Swedish peer grief support association and is maintained and moderated by administrators from the association. The community has more than 1400 members,[2] who produce between 10-40 posts and 100-300 comments per day. The closed status of the group means that only members who have been approved by the administrators are able to access it. Members of this community are peers: They share the experience of having lost a child. In most contemporary Western societies, the death of a child is a near-taboo subject which is so uncomfortable that it is often avoided in everyday encounters (Brotherson & Soderquist, 2002). Moreover, implicit norms require bereaved parents not to grieve for their dead child too openly, too intensely, or for too long. Bereaved parents' grief is thus stigmatized, limiting their opportunities to cope with it.

Social media offers new opportunities for getting in touch with peers in a community where they may talk about their dead children and share experiences and feelings. In a study exploring bereaved parents' use of social media as a resource for coping with the loss of a child, we have conducted surveys and interviews with members of the community. Results show that important functionalities of the group are its closed nature, the shared experiences, and the immediate and constant access to the community through digital and mobile technologies. Members are often in a vulnerable state and perceive the community as a safe haven in which they can share things that cannot be shared outside the group, without being "judged" by the norms of the majority culture. The closed status of the group is regarded as a prerequisite for the community to function (Hård af Segerstad & Kasperowski, 2015a).

When users sign up for a Facebook account, they comply with the conditions stated in Facebook's user agreement, i.e. that all content produced by users is owned by the platform provider and may be sold to third parties.[3] What this entails and the extent of Facebook's practice of monitoring and harvesting their interactions, is not always evident to the user (Zimmer & Proferes, 2014). The everyday interactions of the bereaved parents in the community are informed by high expectations of privacy (Zimmer & Proferes, 2014), which are, in practice, contradicted by Facebook's terms and conditions. Thus, research on this group highlights several ethical principles for responsible internet research.

Ethical Principles in Relation to Studying On-line Communities

In our case-study, made data (as in Hård af Segerstad & Kasperowski, 2015a) pose fewer ethical challenges than found data. These different types of data emphasize different ethical considerations and have different implications for the many ethical guidelines for internet research, which have been proposed and discussed on a more or less regular basis since the early 2000s (e.g. Ess & AoIR ethics working committee, 2002; Markham & Buchanan, 2012; EU, 2015). These guidelines are usually formulated in general terms, with the recognition that *in practice* research ethics must be developed in context and cannot be applied universally (Markham & Buchanan, 2012). As Vitak et al. (2016) argue, "discussions of research ethics should be rooted in practice, if we are to have a constructive debate around how to update ethical principles for online research" (p. 10). Principles for internet research emphasize that "the greater the vulnerability" for the subject of research "the greater the obligation of the researcher to protect" (Markham & Buchanan, 2012; p. 4).

---

[2] As of April 2016.

[3] https://www.facebook.com/terms (accessed April 26, 2016)

Moreover, since “digital information […] involves individuals” the balance between “rights of subjects” and “social benefits” must be monitored carefully, with ethical considerations continuously addressed in a “deliberative process” during research (pp. 4-5).

In our study of bereaved parents on Facebook, we have undertaken a deliberative process with administrators and moderators of the community, discussing subjects’ rights and potential social benefits at length. This has resulted in informed consent from the administrators, and expectations regarding how the research will benefit the studied community. This highlights the requirement to manage expectations – an aspect of ethical concern which is not commonly addressed. In our study of bereaved parents, the expectation is that the research will produce knowledge that can be used to mitigate or rectify societal issues or inform political decision-making. We have also been engaged in ethical deliberation with colleagues in the international research community (Hård af Segerstad, 2013; Hård af Segerstad & Kasperowski, 2015b). Such ongoing dialog is essential for keeping up with methodological and technological developments in rapidly changing digital environments.

However, the vulnerability of a community in which individuals communicate and share highly sensitive content obliges the researcher to move beyond informed consent given by only administrators and moderators. The biomedical heritage of ethical guidelines is based on harm-reduction principles, where full informed consent is a gold standard, i.e. all participants are informed about the purpose of the study, participation is voluntary, and withdrawal is possible at any time. Following McKee and Porter (2009), we agree that informed consent from all subjects is desirable in cases where vulnerable individuals are under study.

It can be argued that members’ expectations of a high degree of privacy, vital for the community to function, set a context for research in which informed consent is required (cf. Vitak et al., 2016). When it comes to data from interviews and surveys this is attainable, however, the issues are more complex when studying the interactional data and content generated by some 1400 people. These data are produced by human subjects – not a faceless crowd, but vulnerable individuals. As Vitak et al. (2016) argue, the principle of “respect for persons” goes beyond the scope of informed consent and anonymity in the data collection process. Even if this principle is fulfilled, simply asking for consent may lead to unforeseen methodological and ethical consequences.

Vanderhoven, Schellens, Vlacke, and Raes (2014) argued that the need for written informed consent from the participants in their study of teenagers’ Facebook profiles was waived, as “obtaining informed consents would have jeopardized the reliability of the study” because the “[t]eenagers could have changed their Facebook-profile […] before observations took place.” Although the researchers took care to make sure that the dataset remained anonymous – by encrypting the real names, merely substituting real names cannot guarantee anonymity (cf. Zimmer, 2010).

Similar issues arise with our data. Additionally, we face ethical problems during the period of data collection because the community is a vital resource for group members’ coping strategies. Once the users are aware that they are involved as research subjects, they may adapt their behavior and lose trust in the supporting function of the group.

Access to data has traditionally depended on an interpersonal exchange between researchers and individuals giving their consent, however, such access now tends to go through large, private companies such as Google, Twitter, and Facebook. The members of the closed community on Facebook could be conceptualized as having already given consent in obtaining their Facebook accounts, through *institutional and policy rules of consent* (Beauchamp, 2011). However, for our purposes this is insufficient; we require *autonomous authorization* by every member of the community (Beauchamp, 2011; following McKee &

Porter, 2009 a.o.), i.e. each user intentionally agreeing to something which they have adequate understanding of and being free of any internal or external pressures to do so.

Consequences of trying to obtain autonomous authorization (which may include members leaving the group when in need of support) must also be related to the demanding criteria of "full disclosure and complete understanding" as a gold standard of informed consent. These criteria are difficult to meet and it is therefore reasonable to explore if partial informed consent combined with carefully crafted digital methods for anonymization might be a viable route for responsible internet research.

Anonymization is always a trade-off between the comprehensiveness of anonymization and the integrity of the research quality. As discussed in Ohm (2010) one must consider the difficulties of fully protecting anonymized datasets from the possibility of re-identification of individuals. Aggressive suppression, i.e. deleting all direct identifiers (name, social security number, etc.) and indirect identifiers (other information which, when combined, may allow identification of an individual, e.g. year of birth, occupation, postcode, etc.), can render data almost useless for research since the independent variables are more or less removed. To generalize rather than to suppress identification might better balance research possibilities with anonymization. A third alternative is aggregation, which involves using statistics to summarize results, "coarsening" data by decreasing the granularity of personal information (e.g. only providing a person's county rather than their address) or detaching the utterance from the speaker. Yet another possibility is the use of "safe rooms" (cf. Watteler & Kinder-Kurlanda, 2015) in which archives are secured (physically or by using secure software environments). This way, only authorized researchers are allowed to access sensitive data, which in turn makes review of source data possible.

A Step-by-step Discussion on Research Ethics in Practice

In this section we discuss how ethical issues must be addressed at all stages of the research process and how they can be re-thought from an integrative perspective. In this way the often technical hands-on work can be connected back to the ethical issues and overall scope of the research process.

Step 1: Access to Data

A closed group on Facebook is only accessible to approved members, and can thus be regarded as a private, or semi-private, space. As researchers, however, we have to take into account how community members perceive of their privacy (cf. Zimmer & Proferes, 2014).

Closed communities devoted to highly sensitive and private topics, such as child loss, have previously been largely inaccessible to researchers. We have access to this community because one of the researchers is a member. Previous approaches by other researchers have been declined, due to the sensitive subject matter and the bereaved parents' fear of being judged by the norms of the majority culture. In Western societies, for example, a common conception is that the best way to cope with grief is to find closure and 'let go' of the deceased (cf. Klass, Silverman & Nickman. 1996). Many of the respondents in our case study report that they cannot grieve openly and fear that being subject to research would pathologize them (Hård af Segerstad & Kasperowski, 2015a). In our surveys and interviews, members welcomed our research, with some articulating hope that it would contribute to a better understanding of their stigmatized situation outside the closed community. They expressed trust in a researcher who is a peer, even providing administrator status. The association maintaining the group also officially endorses the research (minutes of the annual meeting of the grief support organization, 1 March 2015).

The main ethical problem here is the difficulty of grasping what access to data means in digital environments. As a trusted administrator of the Facebook group, the researcher has access to more information than most users are aware of, complicating the notion of informed consent while relying on a fragile trust-based agreement between community and researcher.

Step 2: Collecting Found Data

There are two main approaches to collecting data from social media platforms, with different ethical consequences. The first involves accessing the material from the perspective of the user interface, referred to as “browsing”, “lurking” or “webnography”. In this approach, most meta-data is removed by the graphical user interface, and only a limited amount of data can be saved and stored, for example, by creating screen shots, copy-pasting text or saving images. However, this limitation does not preclude the ethical considerations needed to access the personal profiles of users without consent.

The second approach uses another entry point for accessing data. Originally constructed as a way for software developers to create applications and services on top of social media platforms such as Facebook, Twitter, and Wikipedia, the *Application Programming Interface* (API) can be used to extract a large amount of machine-readable data (Batrinca & Treleaven, 2015). However, because this technology was made for programmers rather than researchers, the built-in design choices are based on capacity and cost-related problems rather than ethical standards. Thus, even though APIs are usually restricted to collecting publically available data (cf. Bessi et al., 2015), the very act of collecting data can be problematic, as researchers can build a database of people who subscribe to specific political content, discuss medical information, or belong to vulnerable groups. In this respect, Facebook is a particularly critical case, as they require people to provide “real names and information”.[4] Although this policy is not always adhered to (Hern, 2015), the users cannot hide their identity without breaching the terms of service, and the publically available data therefore tend to be accurate.

A crucial problem here is that the default setting for most off-the-shelf software is to collect as much data as possible. This is tempting from a methodological perspective but can lead to ethical dilemmas. Unlike conventional survey methods, data from online environments have a real-life connection, which is difficult to disconnect from the study at hand. For example, as shown in Figure 1, the raw data from Facebook are structured around a unique fifteen-digit ID-number making every individual user traceable both across the Facebook platform and all related web sites and third-party services connected to Facebook. Moreover, the data have a very high resolution, beyond the knowledge of most users. Each post, comment and the like is precisely time-stamped and every image is preserved. In this way, the data are “more real” when retrieved from the API than when viewed by the users themselves.

```
{
    "from": {
        "name": "Jane Doe",
        "id": "234678943234046"
    },
    "created_time": "2013-04-21T19:32:45+0000",
    "message": "This is a Facebook comment.",
    "id": "196544587141019_754236",
    "user_likes": false,
    "can_remove": false,
    "like_count": 0
}
```

Figure 1: Data and meta-data from a Facebook comment, retrieved using the Facebook Graph API and stored as a json object.

Using ready-made software or developing your own programs makes it possible to collect hundreds of thousands of interactions with little effort. In the latter case it is possible to collect only the data needed for the research question as a way of minimizing the dangers

---

[4] https://www.facebook.com/terms (accessed April 25, 2016).

of surplus data and to obfuscate the data in such ways as to render them less identifiable, by "suppressing identities" or "generalizing identities" (Ohm, 2010).

The *principle of data-minimization* suggests that one should only retrieve the data needed to pursue a specific research question. The benefit of this approach is that the ethical question of sensitive data becomes directly related to the aims and purpose of the study, making it easier to determine the relation between the collection of data and any possible consequences. For example, if the research question asks for the social structure of an online community, the study would need to collect data about user interactions (e.g. who talks to whom). As such studies can potentially reveal sensitive information about specific people, a judgement on whether to collect data or not can be made before operationalizing the research questions further.

```
{
    "from": {
        "name": "dd3785ad4561af3e97a773d3526469ce6f6028388",
        "id": "0d6a50eb4f7b89d338f4a6b58364d26dd0db73e20684a8ebfd837d5e717aa43d"
    },
    "created_time": "2013-04-21",
    "message": "This is a Facebook comment.",
    "id": "92222e379b55e71d7427424c3097266b6b73eae19fc7530babaccf7f4ee634cb",
    "user_likes": false,
    "can_remove": false,
    "like_count": 0
}
```

Figure 2: Anonymous Facebook comment. The real name and the identification numbers are replaced with cryptographic hashsums and the exact time stamp is removed (cf. Jones et al. 2013).

While complete anonymization of found data is close to impossible, minimizing the data that can be connected back to a real individual is a viable rule of thumb. There are several strategies for achieving partial "anonymization" of data using encryption technologies that replace identifying information with cryptographic signatures, thus preserving data consistency (e.g. Vanderhoven et al., 2014). In Figure 2, only the content of the Facebook comment is left intact and the timestamp is restricted to the date. The identifying data are replaced with cryptographic hashsums, which preserves data integrity – avoiding duplicate data and preserving the possibility of retrieving what different users say. Even though it is still theoretically possible to reconstruct the source data using only the textual content of a post, this would require direct access to the Facebook database or collection of the same data again.

Step 3: Analyzing Data

When analyzing the collected material, the data minimization strategy outlined above ideally avoids unintentional re-identification of individual users (Vitak et al., 2016; Zimmer, 2010). This is easier to achieve when studying aggregated phenomena, such as frequency of posts, time of day when posts are written, or quantitative patterns of interactions (likes, comments, natural language processing of text corpora), where only the precise units of analysis are extracted from the collected data. But in cases where individual users are the subject of in-depth analysis, the anonymization process is crucial, unless informed consent from each individual has been obtained.

The contradiction between anonymity and analytical clarity (c.f. Ohm, 2010) must be handled with special attention to patterns which can be used to re-identify individuals. For example, there are usually external variables that structure the analysis of natural language in Facebook posts and comments. The analysis might explore at what time of day certain words are used, if there are gender or age differences, or patterns in the turn taking of conversations. Even if the identities of the users have been removed or obfuscated, there is a risk that a determined third party will be able to reconstruct the data and re-identify users. Patterns in metadata can be very revealing in context. For example, knowing what time of the day a person has a lunch break or wakes up might be enough to connect a pattern back to an individual user. This risk increases with small online communities, where participants share intimate experiences and know each other personally.

Step 4: Dissemination/Publication of Data

Once the analysis has been performed, the results and sometimes the source data will be inspected, disseminated, and made publically available. Current science policy is promoting the idea of maximal open data (EC, 2016; NSF, 2015). This has ethical implications for sensitive data, which need to be carefully thought through in each case. These emerging ethical problems should be considered at an earlier stage of the research process. Data extracted from digital environments are, to varying degrees, searchable and traceable once they are made public. Open platforms such as Twitter provide accurate means of searching through online content, whereas Facebook has a more restricted interface, with full-text search only possible within confined parameters. However, the extraction of data from the APIs makes it possible to create independent search functionalities that in turn enable further re-identification of data.

This means that publication of direct quotes must be avoided. In cases where quotes are used, individual permissions must be obtained making sure that research subjects are aware that their anonymity may be breached. Such permission is, however, impossible to attain if the data have already been anonymized in the collection and analysis stages. Another strategy is to use “composite narratives” (cf. Davidson & Letherby, 2014) where direct quotes are rewritten to avoid searchability. However, this strategy involves the modification of data, which necessarily reduces data quality.

Disseminating results to academic peers, especially when attempting to replicate a study or inspect the source data for verification (e.g. by anonymous reviewers), means that all prior steps in the research process are also put to the test in terms of research ethics. The structure of source data is always more sensitive than aggregate and published results. As there might be unforeseen methods of data analysis that lead to accidental re-identification of individuals, it is important to also consider any potential information that can be extracted from the data.

Further ethical issues arise in our case when research subjects explicitly ask us to use their real name and that of their deceased child in publications and disseminations of results. Researchers in death studies have argued that in some cases, anonymity may be disrespectful to both the bereaved and the deceased (cf. Walter, 1996).

The questions of anonymity and informed consent thus also need to be considered in light of what this entails for the research subjects both in the short and long term in relation to their evolving grief process. This affects issues of informed consent of individual members in the community and the expectations of both individuals and the grief association.

Discussion

In this chapter we have discussed the notions of informed consent and anonymity in relation to studying a closed and vulnerable group on social media. We have attempted to tie together methodological considerations and the demand for high-quality data with possible strategies for minimizing harm through the principle of data minimization, where the collection of data is constrained by the specific research questions. We urge researchers to avoid relying on ready-made software solutions, which are often constructed for purposes other than research and thus lack “built-in” ethical considerations. We also argue that general ethical principles must be re-thought and adapted to very specific contexts that change quickly as technology develops.

When working with found data informed consent is difficult to attain and may disturb the community studied as well as affecting data reliability, as individuals alter their behavior due to the presence of researchers. Indeed, found data were never designed to be researched in the first place. When posting information on social media users do not generally consider the possibility of being systematically observed, studied, and disseminated in research outputs. Anonymizing and reducing the data to the bare minimum required to answer a specific research question is viable. However, such approaches need to be carefully monitored throughout the research process and take into account the consequences of dissemination and data sharing that may be required in peer review.

The alternative is to strictly adhere to informed consent in all instances of research. As a consequence, found data will be off-limits for academic and scholarly research, leaving such data (including all phenomena made up by such data) to marketing bureaus, intelligence agencies, and other entities not bound by research ethics. Such a limitation on research could be counter-productive in the long term, as it risks casting a shadow on online behaviors and phenomena that need to be analyzed with proper scientific methods. The challenges imposed upon us by the rapid digitalization of society cannot be met without the increased knowledge produced by responsible research.

References

Batrinca, B., & Treleaven, P. C. (2015). Social media analytics: A survey of techniques, tools and platforms. *AI & Society*, 30(1), 89–116.

Bessi, A., Mauro C., Davidescu, G. A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Science vs conspiracy: Collective narratives in the age of misinformation. *PLoS ONE*, 10(2), e0118093.

Beauchamp, T. L. (2011). Informed consent: Its history, meaning, and present challenges. *Cambridge Quarterly of Healthcare Ethics, 20*(04): 515–523.

Brotherson, S. E., & Soderquist, J. (2002). Coping with a child’s death. *Journal of Family Psychotherapy, 13*(1-2), 53–86.

Davidson, D., and Letherby, G. (2014). Griefwork online: Perinatal loss, lifecourse disruption and online support. *Human Fertility*, *17*(3): 214–217.

EC (2016). *Open Innovation, Open Science, Open to the World – a vision for Europe*. Retrieved from the European Commission website http://bookshop.europa.eu/en/open-innovation-open-science-open-to-the-world-pbKI0416263/

Ess, C., & AoIR Ethics Working Committee (2002). *Ethical Decisionmaking and Internet Research. Approved by AoIR, November 27, 2002.* Retrieved from http://www.aoir.org/reports/ethics.pdf.

EU (2015). *Opinion 4/2015: Towards a new digital ethics – Data, dignity and technology European Data Protection Supervisor*. Retrieved from https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2015/15-09-11_Data_Ethics_EN.pdf

Hern, A. (2015, November 2). “Facebook relaxes ‘real name’ policy in face of protest”. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2015/nov/02/facebook-real-name-policy-protest

Hård af Segerstad, Y., & Kasperowski, D. (2015a). A community for grieving: Affordances of social media for support of bereaved parents. *New Review of Hypermedia and Multimedia*, *21*(1-2), 25-41.

Hård af Segerstad, Y. (2013). Trust and reliance: Ethical and methodological challenges of studying sensitive topics online. An insider's dilemma. *Paper in the Panel "(Research) Ethics at the Edge: Case-Studies and New Directions" at Internet Research 14.0: Resistance and Appropriation, the fourteenth International Conference of the Association of Internet Researchers*. Denver, CO: AoIR

Hård af Segerstad, Y. & Kasperowdski, D. (2015b). Opportunities and challenges of studying closed communities online: digital methods and ethical considerations. *Paper in the Panel "Internet Research Ethics: New Contexts, New Challenges – New (Re)solutions?". Internet Research 16.0: Digital Imaginaries, The Sixteenth International Conference of the Association of Internet Researchers.* Phoenix, AZ: AoIR.

Jensen, K. B. (2012). Lost, found, and made. In *The Handbook of Global Media Research*. Wiley-Blackwell (pp.433-450). Retrieved from http://dx.doi.org/10.1002/9781118255278.ch25.

Jones, J. J., Bond, R. M.., Fariss, C. J., Settle, J. E., Kramer A. D., Marlow, C., & Fowler J. H. (2013). Yahtzee: An anonymized group level matching procedure. *PLoS ONE*, 8(2), e55760.

Klass, D., Silverman, P. R. & Nickman, S. L. (eds.) (1996). *Continuing Bonds- New Understandings of Grief*. Philadelphia PA: Taylor & Francis

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *PNAS: Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8788-8790. Doi:10.1073/pnas.1320040111

Markham, A., & Buchanan, E. (2012). *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)*. Retrieved from www.aoir.org/reports/ethics2.pdf

McKee, H. A., & Porter, J. E. (2009). *The Ethics of Internet Research: a Rhetorical, Case-Based Process*. New York, NY: Peter Lang.

NSF (2015). *Open Government Plan (3.5).* Retrieved from The National Science Foundation website www.nsf.gov/pubs/2015/nsf15094/nsf15094.pdf

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anoymization. *UCLA Law Review, 57,* 1701–1777.

Vanderhoven, E., Schellens, E., Valcke, M., & Raes, A.(2014). How safe do teenagers behave on Facebook? An observational study. *PLoS ONE, 9*(8).

Vitak, J., Shilton, K., & Ashktorab, Z. (2016). Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW'16)* (pp. 941-953). New York, NY: ACM.. Doi=http://dx.doi.org/10.1145/2818048.2820078

Walter, T. (1996). A new model of grief: Bereavement and biography. *Mortality*, *1*(1), 7-25.

Watteler, O., & Kinder-Kurlanda, K. E. (2015). Anonymisierung und sicherer Umgang mit Forschungsdaten in der empirischen Sozialforschung. *Datenschutz Und Datensicherheit - DuD*, *39*(8), 515–519.

Zimmer, M. (2010). “But the data is already public”: On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313–325.

Zimmer, M., & Proferes, N. J. (2014). A topology of twitter research: Disciplines, methods, and ethics. *Aslib Journal of Information Management*, *66*(3), 250–261.