University of Canberra, Australia

From the SelectedWorks of Moosung Lee

2016

Ethical issues in (online) social network research in education

Available at: https://works.bepress.com/moosung-lee/77/
Ethical Issues in (Online) Social Network Research in Education

Moosung Lee* · Ewan Wright

Faculty of Education, Science, Technology, and Mathematics University of Canberra
Faculty of Education University of Hong Kong

Bionotes: Moosung Lee is the youngest Centenary Professor, one of the most prestigious professorships, at the University of Canberra. Prior to joining the University of Canberra, he held appointments as Associate Professor and Founding Deputy Director of Education Policy Unit at the University of Hong Kong. He has published numerous articles in high quality academic outlets in the areas of educational leadership and administration, urban education, and comparative education. He was the winner of the AERA Emerging Scholar Award (Division A-Administration, Organization, Leadership) in 2015. Ewan Wright is engaged in doctoral research in the Faculty of Education at The University of Hong Kong, where he holds the Hong Kong PhD Fellowship. His educational background is in Sociology and Political Economy. Prior to joining The University of Hong Kong he worked with various think tanks in the UK and Hong Kong including Centre for Cities, Demos and Civic Exchange. His core research interests are based around inequalities in education, the massification of higher education, and transitions from higher education to employment.

Abstract: In this article we discuss a relatively under-researched area in terms of research ethics, namely, social network research in education, including educational research focusing on social networking websites. Our discussion focuses on several key issues in social network research such as privacy in conjunction with anonymity, voluntary participation with foci on consent forms and second parties. In addition, we shed light on newly emerging issues from social media based research. In so doing, we call for more cautious approaches to social network data collection and analysis in education research.

Keywords: social network research, social media, research ethics, educational research

I. Introduction

It was only about a decade ago when the New York Times branded “social networks” as one of the new ideas of the year (Gertner, 2003). Since then, the concept of social networks has gone beyond the realm of jargon and highly technical discussions. It is now widely permeated into our everyday lives, especially through social networking websites. Academic research mirrors this emergence of social networks particularly over the last decade. Whereas less than 100 social science studies with “social networks” in their abstracts were found by 1975, thirty years later this figure reached 2,500 (Knook & Yang, 2008). Following this, our investigation via the Education Resources Information Center (ERIC) further identified that 1,703 documents, which include “social network” in their abstract, had been published between 2006 and 2015, compared to only 426 over the preceding decade. Of these documents, the majority were related to education settings such as K-12 (36.3%) and higher education/adult education (50.9%).

Notably, almost two-thirds (63.1%) of the research documents related to K-12 settings had been published over the past five years. This is a clear signal of the growing popularity of social network analysis as a research tool or lens for researchers in K-12 education. Despite such an expansion, as Ronald Breiger, the former editor of Social Networks, pointed out, it is rare to find discussions of ethical issues in data collection and analysis of social networks (Breiger, 2005). Indeed, a lack of systematic guidelines for social network research has been raised as a major point of concern (Curtis, 2014). Moreover, an absence of agreed upon guidelines is especially pronounced for educational research targeting K-12 students and teachers (cf. Lee, 2011; Penuel, Sussex, & Korbak, 2006). Within this in mind, we discuss key ethical issues inherently embedded in social network research in education and thereby aim to explore sensible guidelines for research using (online) network data.

II. The Issue of Anonymity

A major barrier to conducting ethical social network research is difficulty in ensuring anonymity. By its
very design, research of social networks may not secure complete participant anonymity in the data collection process (Borgatti & Molina, 2003). Fundamentally, this is because the standard form of interview or survey questions are aimed at getting participants to reveal the names of other people in their social networks (e.g., list the names of your closest friends). In other words, network research in general, and especially network research in organizational settings, have in-built limitations in terms of protecting absolute anonymity (cf. Lee, 2011).

Moreover, such concerns are not limited to the data gathering process. Even in the presentation of research findings, network research does not necessarily secure anonymity. In direct contrast, to “the highly digested outputs” of [conventional statistical] analysis, network diagrams called sociograms are built from raw data that mirror how participants are connected to one another (Borgatti & Molina, 2003, p. 341). While, of course, researchers can use pseudonyms to protect the identity of participants in their network analysis outputs, in certain cases the real names of participants can still be identified. This may be especially problematic when network analysis is conducted in a small organizational setting. In such circumstances, participants may be “re-identified” through some speculative steps and based on only a limited numbers of characteristics—e.g. only one Asian immigrant female teacher in a particular school.

However, there can also be difficulties in securing anonymity in larger organizational settings. To take a real case as an example, a 2008 study aimed to provide a snapshot of the social networks of approximately 1,700 college students at a U.S. university over a four year period by using data from Facebook (Lewis, 2008). Despite certain precautions taken by the researchers, the anonymity of the participants was soon called into question. First, the case university was “re-identified” using information published by the research team including approximate geographical location and student population size. Second, it is arguable that some individual students could be identified through personal information such as gender, nationality, discipline of study, and the structure of their individual social network on Facebook (Zimmer, 2010). While this is a case from a higher education sector, a parallel case could occur in K-12 settings.

Perhaps even more problematically, if such social network data is used as a basis for managerial actions or personnel decisions, then the implications of failing to secure anonymity can become extremely serious (Bogatti & Molina, 2003). To take a hypothetical example, suppose that based on the findings of network analysis, a district office opted to close a particular department or program, which turned out to have inefficient structures and functions in terms of communication channels, resource flows, collaboration opportunities, and so on. Such concern was reflected in a study of the perceptions of educators about the collection and publication of organizational network data. Based on interviews with teachers, principals, and school leaders, three-fifths (60 percent) of participants reported being worried about the sharing of findings with school stakeholders. There was concern that if individuals were found to be relatively isolated from social networks, the results could be used as an accountability tool. In addition, such findings could potentially undermine individual self-confidence, the development of close-knit school communities, and friendships among staff (Penuel, Sussex, & Korbak, 2006).

These examples suggest that in the context of high accountability, network data intended for school-based program evaluation and the like, is likely to generate certain tensions from people involved in such data collection. Put another way, unlike network research for purely academic purposes, network research outputs conducted in organizational settings can directly affect the lives of people, especially when they are used for managerial or personnel purposes. In fact, as Kadushin (2005, p. 146) pointed out, “organizational network analysis with purely academic purposes is relatively rare in that access to organizations usually requires a quid pro quo in that the organization may necessitate a report which could aid them in understanding and possibly reorganizing their organization.” Borgatti and Molina (2003) further noted that mixed academic/consulting network research is perhaps the most dangerous as well as the most common because survey questionnaires typically describe the academic purposes of research but neglect to mention in what ways the research outputs will be presented and utilized by management.
III. Consent Forms as a Safeguard?

Social network researchers should strive to ensure that participants engage in the research on a voluntary basis. One standard procedure for achieving this is to use an informed consent form, like any other research (Lee, 2011). Yet, gaining informed consent does not necessarily safeguard voluntary participation in its own right, especially in organizational settings (Borgatti & Molina, 2005). First, employees may feel under-pressure to provide informed consent due to anxieties about the potential consequences stemming from their non-participation. Second, the anonymity of non-participants may not be secured as their non-participation may be identified from the list of participants. In addition, network research in certain organizational settings for consulting purposes “is not subject to IRB review,” meaning that informed consent forms are not mandatory (Borgatti & Molina, 2005, p. 110).

Despite such implicit organizational pressures, an individual may choose not to participate in organizational network research from the beginning. As Borgatti and Molina (2005) noted, however, non-response cannot be absolutely addressed as omission in network research. For example if a person refuses to fill out a questionnaire, they can be mentioned by other persons in the network research. In such a case, the person who opts not to complete the questionnaire may still appear in the network analysis - i.e. the issue of second parties. In sum, network analysis does not guarantee absolute non-participation. This can become a serious ethical issue particularly when questioning is directed at sensitive topics - e.g. Name your close friends engaging in some form of illicit or undesirable activities such as drug use in your school. The concern here is the fact that network research may stimulate participants to do unethical things (Borgatti & Molina, 2003).

A noteworthy case to demonstrate this issue is the Add Health survey which is one of the largest investigations into the health of adolescents in the U.S. ever undertaken. The survey included asking participants highly sensitive question about drug use and romantic relationships among peers in their social network. For instance, one question asked, Did you ever give INITIALSEX in exchange for drugs or money? (Harris & Urdy, 1998). Given the sensitivity of such data, one can only imagine the fallout if the identities of participants were disclosed. In response, the research team stressed that they took “extraordinary precautions to protect the data from unauthorized use” by for example signing pledges of confidentiality and password protecting all data (Add Health, 2014).

IV. Social Media Websites: A Research Bonanza with Ethical Minefields

Research of social networks using social media websites, especially Facebook and Twitter, has undergone a boom in recent years. Yet, such approaches to research complicate the ethical issues noted above and have created emerging challenges for researchers and ethics committees alike (Buchanan & Hvizdak, 2009).

First, social networking sites can provide an efficient way of recruiting large number of research participants. However, such approaches raise new ethical concerns as data gathered via the recruitment process could be used by social network companies for other purposes. Notably, Curtis (2014) gives the example of an individual responding to an advertisement on a social networking site to participate in a HIV related study. Curtis notes that such sensitive information collected could, in turn, be used for targeted marketing purposes. In addition, interest in the study may be viewable to the other users via functions such as “Like” on Facebook.

Second, as far as consent forms are concerned, more often than not, social media users are not aware of those who observe or crawl through the information uploaded on their social media websites. This can result in tensions in terms of research ethics, especially when dealing with sensitive topics. In a high profile case, researchers came under media spotlight for a study conducted on 689,000 Facebook users over a one-week period in 2012. The study involved researchers manipulating the emotional content of personalized Facebook “News Feeds” to identify how this influenced the mood of users. While the Facebook users were kept anonymous, the study has been widely criticized as consent was not obtained from participants (Hill, 2014). Notably, some of the users were assumed to be adolescents, given that social media sites like Facebook allow access to those as young as 13 years old.

Clearly, in certain cases, such as where social media
users are under 18 years old, conscientious researchers should attempt to reach potential participants in order to obtain consent. However, this is a daunting task given the need to obtain parental consent in such cases. A further problem, as a matter of fact, is that there is no way to ensure 100% whether such participants are really above 18 (or 13) years old given that many social media users use a fake name or avatar on the Internet (Tsikerdekis & Zeadally, 2014).

A further problem, as noted above, is the issue of second parties. Suppose that researchers intend to gather friendship networks through Facebook, for example. Even though they manage to obtain consent forms from their targeted Facebook users, the friendship network data they gathered would include second parties (e.g., other Facebook users linked as friends to the targeted Facebook users) who will never know they participated in the study. Such an event is especially susceptible to occurring when network researchers use any kind of web-based data that are publicly accessible. This is an important but frequently overlooked ethical issue given that research employing social media to crawl through data is rapidly growing, owing to significant benefits in data collection (e.g., big data mining across cities and countries within a relatively short time), which is rarely possible in conventional off-line research.

V. Ethical Validity at Risk

Issues of privacy, voluntary participation and second parties can put network research in a serious dilemma in terms of ethical validity (and therefore measurement validity as well). Ethical validity refers to whether research methods or tools are ethically applied within social contexts. To evaluate ethical validity, researchers are required to reflect on the following questions, as examples: Does the application of the research methods or tools harm particular sub-groups of participants in the research? What potential harm can be done to persons, even including those who opt not to participate in the research, in terms of the application of particular research methods or tools? What “good” is achieved by the application of the research methods or tools? (Caccamise & Newell, 2015, p. 61).

As mentioned earlier, in network research in organizational settings, if findings would be used for managerial or personnel purposes such as promotion, demotion, and layoff, ethical validity could be seriously undermined (cf. Borgatti & Molina, 2003). To avoid this problem, Penuel et al. (2006) provided some practical advice for network researchers in an organizational setting such as schools. Drawing from their data collection experience, they suggested that by removing network survey questions related to privacy, teachers would be more willing to share their professional relationships with colleagues in school for the purpose of school improvement and staff development; “teachers’ concern with privacy did not extend to professional interactions…some teachers saw it as fair to ask about professional interactions” (p. 448). Nonetheless, since the anonymity of network research in small-scale organizational studies (e.g. a case study of a school) cannot be completely secured, potential harm can be done to particular sub-groups of teachers, even when they are asked about professional relationships in the school. Furthermore, teachers who opt not to complete the questionnaire may still appear in the network analysis—i.e., the issue of second parties.

With regard to ethical validity issues in social-media based network research, there seems to be no clear and comprehensive guidelines currently in place. The rapidly growing research fields that use social media sites further complicates the ethical validity issue. As far as the issue of second parties is concerned, social-media based network research is prone to failing the guarantee of non-participation on a larger scale, because it is not likely that most people on social media websites are given a chance to opt not to participate in network research, especially when researchers do data mining from social media websites. In sum, network analysis using social media websites are particularly vulnerable to the guarantee of non-participation.

Finally, the issues of ethical validity are inextricably intertwined with measurement validity. Network research is often vulnerable especially when it comes to gathering data on personal relationships (e.g., friendship) or sensitive relationships (e.g., romantic relationship among adolescents, peer relationship of substance use, sexual relationship). This type of network data collection can result in non-responses, which triggers the issue of measurement validity. The measurement validity of network research is particularly at risk to missing data from non-responses as a network map or sociogram.
could become misleading or incomplete if they do not include key actors (Borgatti & Molina, 2003, p. 339-340). Notably, challenges to network research due to non-responses and related issues about validity are common among network research targeting children. This is largely since in the process of obtaining consent from children, network researchers may face barriers where a) there are difficulties in passing on consent forms to parents and b) parents may feel safer by taking a default position of refusing consent (Kadushin, 2005). In such cases, a systematic absence of particular children can result in biased network analysis. Furthermore, the rapidly growing research field using social media sites complicates the validity issue. For example, there is no guarantee whether Person X is who he/she claims to be on a social media website. Similarly, relationships (or links) on social media websites can potentially mislead researchers in that social media users tend to accept invitations from others on social networking sites such as Facebook or Linked-In, where the degree or intensity of real relationships can be easily masked.

VI. Moving Towards Sensible Guidelines for Network Data

Reflecting all the issues discussed in this paper, it is imperative to develop certain ethical standards and guidelines in order to conduct ethically-sound network research. As a first of its kind, Borgatti and Molina (2005) set out guidelines that are especially relevant for organizational network research (see also Klovdahl’s (2005) work on network research guidelines). Several desirable practices for ethical network research were proposed by Borgatti and Molina, these are as follows (2005, p. 113):

- **Anonymization and aggregation:** Data collected from respondents should be withheld from management until it has been anonymized or aggregated to a group level.
- **Uncoerced participation:** Consent for participation should be obtained by the research team themselves, instead of through management as this could be understood by some as a direct order to participate.
- **Respondent give-back:** While it can be a time-consuming endeavor, provision of individualized feedback to respondents regarding their position in networks is highly beneficial. Without sharing such potentially useful information, the research could be seen as exploitative.
- **True opt-out option:** Organizational members should retain the right to exclude themselves from the study. That is to say that data collected from others about them will be ignored so that they will not appear in any of the analysis.
- **Pre-survey consent:** Researchers should create a “roster-style” survey which contains the names of only those who agreed to participate.

With respect to social media-based network research, as yet, there are no clear guidelines in place. For general safeguards, the recommendations from the Association of Internet Researchers (AOIR, 2012) are useful. As internet research-based protocols, the AOIR’s document provides an ethical platform for researchers’ decision making about questions such as: How are internet-based data being collected, managed, and represented and how are findings presented? At the same time, however, the document is not specifically targeted at network research using social media sites. Indeed, the concern for the lack of specific guidelines of social media-based network research is well reflected in Shapiro and Ossorio’s (2013) recent article in the journal *Science*. They expressed a serious concern for the dearth of systematic and standard ethical guidelines of social media-based network research. Sharing their research experiences, they were concerned that “U.S.-based academic researchers and institutional review boards (IRBs) have received no guidance from the Office for Human Research Protections (OHRP) regarding how to apply human subjects regulating to SNS research” (p. 144). As such, they viewed the IRBs as largely dysfunctional in terms of an internal regulatory body to ensure abidance to ethical issues in social media-based network research.

VII. Unfinished Business

The complexities of ethical issues in social network research have meant that efforts to establish robust ethical standard guidelines remain a largely unfinished business (Lee, 2011; Kadushin, 2005). Moreover, as far as social media based network research is concerned, we still have more questions than answers about ethi-
cal issues, particularly about privacy. As a response, it is crucial that the development and elaboration of ethical standard guidelines is continued apace, especially given the fast growth of network research and the potential impact such research has on participants. Within this in mind, we provided a number of guidelines and suggestions drawing from the existing literature, but we also acknowledge that more practical guidelines from empirical research using network data are needed. In this regard, we think that IRBs’ role should be played out properly and timely, given the continuously-evolving nature of web technologies. At the same time, it is also important for context-specific considerations from individual researchers about what constitutes as harm to human subjects in their studies (AOIR, 2012). As this paper has argued, network researchers have a responsibility to seriously consider ethical issues when conducting their studies. An important and necessary first steps to achieving this would be to take heed of what Kadushin (2005) asked all network researchers to consider i.e. who actually stands to benefit from their network analysis?

References


Received April 3, 2016
Revised April 8, 2016
Accepted April 28, 2016