Project Title: The Digital Connections Study: Social Networking Technology and its Impact on Health-Risk Behaviors among Homeless Youth
Principal Investigator: Anamika Barman-Adhikari, PhD
Co-Investigator: Kimberly Bender, PhD
DU IRB Protocol #: 978668-1

I.   Background:

Social networking sites (e.g., Facebook) present an important data source for understanding the social context of youths' health behaviors, including information about individuals' attitudes and social norms regarding substance use and sexual risk behaviors. For example, a recent study of adolescent users found that 41% posted about substance use, 24% posted about sex, and 14% posted about violent behavior.[1] Importantly, studies have found that adolescents who are exposed to friends' risky online displays of behavior are more likely to engage in risky behaviors themselves.[2]

Despite their lack of resources and high transience, social media use among homeless youth is surprisingly pervasive. Previous studies by the PI and Co-I find almost 80% of homeless youth report using social media weekly (only 9% do not have a social media profile), and almost 90% of youth who used social media prefer to use Facebook.[3,4] In fact, social media may represent a particularly valuable outlet for homeless youth to maintain social connections, as they typically lack stable phone and address contact information. To date, only three studies, conducted by the PI and her collaborators, have assessed social media use and health-risk behaviors among homeless youth.[6,7,8] These studies suggest that social media use is associated with both positive and negative outcomes; these outcomes generally depend on whom youth communicate with (i.e. family, prosocial friends, street friends, caseworkers, or strangers) and what they talk about (i.e. safe or riskier topics). But, all three existing studies use survey data, which is limited by its reliance on self-reports and retrospective recall. Observational research, in which researchers download participants' actual social media interactions, overcomes these limitations.

Social media websites (such as Facebook) allow researchers unprecedented access to data generated in a real-world environment through an Advanced Programming Interface (API).[9] To our knowledge, no published work exists that has collected observational social media data on homeless youth. Supplementing this observational data with in-person assessments of youths' social networks and reports of their behavior outside of Facebook will provide an innovative view into these youths' social worlds, leading to actionable insights regarding how to use social media to positively influence homeless youths' health. Preliminary data collected through focus groups by the PI (in December, 2015) with 30 homeless youth at the partner agency found youth are interested in social media as a way to receive health prevention programming.[5] A first, important step in developing such

preventative programming, is to understand clearly homeless youths' social media use and how social media interactions influence risk behaviors.

The current proposal thus aims to:

> 1. Explore and describe participants' both past year (retrospective) and post recruitment (prospective) (6-month timeframe) social media interactions by creating a computer program that will pull archival (retrospective) data from participants' Facebook timeline.
> 2. Investigate the relationship between observed social media interactions (both whom participants connect to and what they talk about) and subsequent self-reported attitudes, norms, and behaviors among homeless youth.

II. Design

Study Design Overview

The study will utilize a combination of *observational and traditional survey methods*, to investigate the aims proposed above. Recruitment will take place at the Urban Peak drop-in center and shelter; Urban Peak is a non-profit organization serving homeless youth located in Denver. A purposive sampling design will be used to recruit participants {N=200; eligibility criteria: ages 18 through 26, owner of a Facebook profile (for at least a year), ability to speak and write in English}. Recruiters will be present at the agency, over one month, for the duration of service provision hours to approach and screen youth, and invite participation.

For youth who qualify for the study and consent to participation, there will be two sources of data. First, with participant permission, observational retrospective data from participants' [past-year] and post-recruitment (6-months) ongoing Facebook interactions will be downloaded. Second, youth will participate in a self-administered self-report survey regarding their characteristics, beliefs, social relationships and behaviors at baseline, 3-month follow up, and 6-month follow up.

We are aware that collecting observational data from social network technology platforms is still a relatively new and growing field of inquiry. While still a nascent field of study, several empirical studies have utilized social media data (i.e. observational data) to study young adults' interactions and how such interactions impact their engagement in a wide array of health behaviors (such as alcohol use, hookah use, sexual risk behaviors etc.) [10, 11] These studies have carefully considered and addressed ethical considerations and guidelines associated with working with such new forms of data. [12, 13, 14] We have followed ethical guidelines established and evaluated by these studies in designing our own protocol.

Some of the key guidelines provided by researchers [12, 13, 14] in ensuring participant confidentiality and privacy in social media based studies are as follows:

1. **Participants should have clear knowledge about what information is going to be extracted from their data, and who will be able to access** it. We will accomplish this by giving participants information on data privacy issues and how their data is going to be used during the informed consent. In general, studies have found that participants in Facebook-based studies seem to feel comfortable sharing those extensive records with researchers. [15,16]

2. **Data are de-identified after collection and no attempts are made to re-identify them.** Just like traditional research in which identifying information is initially collected, but de-identified later, researchers recommend similar ways in which to de-identify the data and only retain unique participant identification codes (PID's) to analyze the data later. In the context of our study, this means that even though we collect identifying information when we first download data from social media, they will be assigned Personal Identification Numbers (PID's) and therefore not have any identifying information when the data is stored and analyzed.

3. **Participants can retract, at any time, information that researchers have collected based on their consent.** In order to facilitate this, our team will provide contact information during the consent process, including contact information online that can be responded to quickly in case a participant changes their mind about consenting to their data being analyzed. Additionally, participants (or Facebook users) themselves have the ability to revoke or limit access to data once it has been collected through their privacy settings.

Additionally, **federal regulations** require IRBs reviewing data mining on Facebook to consider

(1) whether the study will minimize risks to the extent that it can; (2) whether the risks that the study poses to participants exist in proportion to the benefits; (3) whether the proposed subject selection process is equitable; (4) whether informed consent will be sought, and if so, whether it will be sought in an appropriate manner; and (5) whether data collection and storage will take place in such a way that subjects will be protected, and in particular, whether the privacy and confidentiality of subjects will be adequately protected.[14]

In designing our protocol, we have tried to adhere rigorously to the above standards.

Study Sample and Recruitment Strategy

Young people will be recruited through research assistants approaching youth in agency common areas, introducing the study and inviting interested youth into a private office to screen for inclusion. Researchers will be present on site several days each week and available to answer questions about the study during those times.

To qualify for the study, youth must be between the ages of 18-26 and be homeless or unstably housed according to the federal definition established by the United States Department of Housing and Urban Development. This is operationalized as either: living in a place not meant for human habitation (streets, car), emergency shelter, transitional

housing and hotels or living in temporary living situations (with friends, family) in which they cannot stay for 30 days or more. Participants will also be required to confirm that they have a Facebook account (at least for a year) and that they can read and understand English (see appended eligibility screener for screening questions).

Procedures

When youth are approached for study recruitment, research assistants will broadly explain eligibility criteria (18-26, homeless or unstably housed), the study purpose (to understand how youth use social media and how the topics they discuss via Facebook relate to their experiences and behaviors offline), method (that this will involve us getting access to their social media timelines, i.e. what they have posted or commented on in the past one year, through a web-based app that automatically downloads this data once we get their permission) and self-report participation (approximately 1 hour self-administered survey completed at baseline, 3 month follow up, and 6 month follow up).

Youth who are interested in the study will be screened for eligibility. Eligibility criteria will be assessed by a trained research assistant who will ask participants about: where they slept last night, how long they can stay at that location, their age and whether they have owned a Facebook profile for a year (see appended screening questions). For youth who meet eligibility criteria, the research assistant will seek informed consent for participation. For those who do not meet eligibility, the research assistant will thank them for their time and discontinue the interaction.

To consent participants, the research assistant will review all aspects of the study in detail with the participant aloud and answer any questions. The consent form (see appended consent form) clarifies the overall objectives of the study as well as details the two aspects of data collection (observational data from Facebook and self-report data via survey). A hard copy of the consent form will be given to each participant to keep. If youth consent to the study, they will indicate so by clicking "I agree to participate" electronically (on the survey docked on **Qualtrics** accessed by an IPad) to proceed. If youth click "I do not agree to participate", the research assistant will answer any questions and end the interaction if youth do not wish to continue.

To allow access to observational data via Facebook, the second screen after clicking "I agree to participate" will ask participants to enter their email and Facebook profile information. For those who do not consent to collection of observational data via Facebook, they will be excluded from the study. Participants who consent will then be directed to a link (via the project IPad through a web-browser). By clicking this link, the participants will be directed to a web-based interface that links with youth's Facebook profiles with consent procedures that will allow us to collect data from their Facebook timelines. The study participant will be given information about what information is collected from their profiles before they consent to this part of the study (i.e. their posts, their likes, their updates and their demographic information).

A link to the information sheet/privacy policy (see appended) will be posted on the website where they access the application. The research assistant will review this privacy policy

aloud with each participant. This is in accordance with the Facebook policy (https://www.facebook.com/legal/terms).

Once participants have agreed to the terms of use, we will be able to start retrieving data from their *timelines (i.e. past-year retrospective data from their FB timeline as well as post-recruitment Facebook interactions for a period of six months)*. Such *third party applications*, which collect information from Facebook profiles, are commonplace and most youth have several such apps installed in their Facebook accounts. The app does not appear in any way on participants Facebook profiles.

Researchers will not "friend" any of the participants. If at any time, participants want to revoke their participation in the study, they will be given contact information for researchers and once we receive that request, we will remove their data from our database.

Data collected will be limited to posts, comments, status updates, links, shares etc. that are ***only generated by the participant*** in accordance with privacy policies set by the social networking website (i.e. Facebook) (https://www.facebook.com/about/privacy/) and guidelines established by respected researchers.[12, 13]

We are also aware that because of the aims and study design we will inadvertently be gathering information from Facebook friends (i.e. secondary subjects or SS) of study participants. Only communication that passes between the study participant and the SS will be available to download. No other data of the Facebook friend not tied to the study participant can be downloaded. We will only access data that the study participant has access to. Although the names of the participants or the SS will not be stripped or replaced on the initial download, their names will be replaced with number codes when the data are combined with the self-report survey data in order to protect the privacy of the both participants as well as SS's. Thus, when the data is being analyzed, there will be no names associated with the data obtained from Facebook. The main purpose for not replacing the names initially is that we want to understand whether youth in our study communicate with each other (i.e. if participants within the study communicate with each other on Facebook). This is important to assess as it has important intervention implications. Learning how homeless or unstably housed youth connect with other homeless youth over social media can inform whether social network technology can be used to deliver peer-based interventions for homeless youth who are often not easily reachable through traditional place-based intervention platforms. We will only be mapping how study participants are linked to one another and take out identifying information for all primary and secondary subjects once data is combined (see data safety and monitoring plan below for more details). Our protocol on secondary subjects was designed to align with guidelines provided by Kosinki and colleagues.[12]

Participants will be informed that their Facebook profile content would be viewed but that no content would be posted on the participant's profile. Data initially collected will be archival in nature (i.e. data will be pulled from youth's Facebook timeline from the day they are recruited going back one year).

In addition to giving consent for observational data collection via Facebook, participants will also provide consent for completing a self-administered survey on a research tablet. Youth will be asked to provide their Facebook ID and email at the start of the survey because that will enable us to link the observation data that from their Facebook timelines to the self-administered survey data.

Youth who participate will complete the self-administered survey, which will take approximately 45 minutes to 1 hour. They will receive $30 in gift cards to a local food vendor as incentive for completing the self-administered baseline survey. Survey data will be gathered using the Qualtrics software using the University of Denver license. Data will be gathered on secure, password protected iPads and data will be uploaded to a secure server at DU within 24 hours after data collection, and then deleted from the iPad.

Follow up self-administered surveys (at 3 months and 6 months) will be collected on-line by providing a link via Facebook private message to each participant. These interviews will assess the same self-reported attitudes, norms, and engagement in health risk behaviors (i.e., substance use and sexual risk) using standardized instruments as assessed during the baseline self-report survey. Youth will be compensated $15 via on-line gift cards to local food vendors for each follow-up survey completed.

Measures

Self-report surveys: The self-report assessments (see appended survey instrument) administered at baseline, 3-month follow up, and 6-month follow up) will address several domains, including:
1. Demographics (age, race/ethnicity, gender, sexual orientation, education, etc.).
2. HIV and STI testing and treatment history (date of most recent test; what STIs tested for; diagnoses and treatment; perceived HIV susceptibility)
3. Health care access (location of and reason for last healthcare visit, unmet need for physical and mental healthcare, prescription medication access, etc.)
4. Homelessness history (where slept last night, experiences of literal homelessness, length of lifetime homelessness, reasons for homelessness)
5. Sexual risk taking (unprotected anal and vaginal sex, exchanging sex for money, drugs or a place to stay, meeting sex partners online, sex under the influence of drugs/alcohol, recent multiple and concurrent sexual partners)
6. Recent drug use (past 30 days' alcohol, marijuana, meth, heroin, injection drug use, prescription drug misuse etc.)
7. Technology access and use (whether youth have access to internet, frequency and comfort of technology use etc.)
8. Mental health (depression, anxiety)
9. Victimization (experiences of interpersonal victimization such as assault and robbery in past 3 months)
10. Social network questions (A name generator will be utilized to collect the social network data. Participants will be asked about five people that they communicate with most on Facebook (first and last name and Facebook usernames of these participants). Collecting both first and last name as well as Facebook usernames of their Facebook

friends will enable us to link them through our observational data later. We are especially interested in understanding whether who they report (or perceive) they communicate with most on Facebook and around what topics is concordant with what we find out from our observational data. After they nominate these five people, they will be asked follow-up questions about the nominated people (i.e., Relationship (e.g. demographics, duration of tie, type of tie, HIV and substance use risk behaviors, methods of connecting, social norms).

Observational data downloaded via Facebook: This will include: the most common topics discussed, the format in which these are discussed (status update language, shared links), and how postings about health risk behaviors are socially reinforced (comments, "shares", "likes") by others.

III. Potential Risks:

Psychological risks

Self-report survey: Some of the questions asked on the survey may be upsetting for young people, as they ask about sensitive topics such as experiences of victimization and current risk behaviors such as sexual activity and substance use.

To address these potential risks, we will take the following steps: informed consent procedures will spell out the time, topics and task demands of the research. Further, informed consent procedures will emphasize that participation is voluntary and that participants have the right to skip questions or to terminate involvement with no penalty. Graduate-level research assistants will be trained to be sensitive to fatigue and distress in participants and will offer breaks or will terminate sessions if a participant is highly distressed or uncomfortable. In rare cases where youth become emotional during assessment, they will be immediately referred to Urban Peak staff who will be onsite and are specifically trained to help youth deal with these issues. In the event that the young person wants to talk to an agency staff person during or after the interview, he/she will have access to the services that the agency staff offer such as counseling, case management, and referral services. All youth will be provided with a list of resources (see appended resource inventory) along with the consent form for the study.

In the event that a young person reveals information that they are homicidal or suicidal, Urban Peak staff will be available for immediate risk assessment and facilitate reporting and connection with appropriate mental health services. All research team members will be trained in mandated reporting issues, as well as protocols specific to this research project. Reports of abuse or neglect will be handled according to guidelines in state law. If the graduate-level research assistant is concerned about participant safety, they will immediately contact Urban Peak staffs on site as well as the study PI by cell phone.

In regards to the 3-month and 6-month self-report surveys conducted online, a screen will appear that thanks participants for their time and provides them with referral contact information (crisis line and Urban Peak) should they want support or services related to any of the topics they shared in the survey.

Facebook/observational data: In regards to the observational data (i.e. data retrieved from participants Facebook profiles) data will be retrospective looking back from several points in time. At baseline, data will be recovered from past year. Moving forward after study recruitment, data will be recovered at 3 months and at 6 months post-recruitment with each time point recovering data from the previous 3 months. These data will be downloaded into a database and any identifying information will be stripped off before merging these data with baseline data. While participants or secondary subjects may express risky thoughts, feelings or behaviors, such as suicidal ideation (SI) or homicidal ideation (HI) (or other illegal activities) on their timelines, this data will not be observed in real time on an individual basis. Rather, observational data will be archival (i.e., collected after the fact), will be collected in aggregate across participants, and will be de-identified upon retrieval. Thus, because this study does not attempt to personally monitor any of these youth's Facebook profiles on an ongoing basis, there will be no risk for the researchers discovering SI or HI in real time for any individual youth. Any high-risk postings are monitored instead by Facebook's internal suicide risk protocols that have been appended below. IRB guidelines require that for a suicide or suicide attempt to be reportable, this would have to be at least possibly related to the study intervention procedures.[17] Our study procedures (especially as it relates to collection of observational data) do not involve intervention, and we therefore believe that our research will not exacerbate suicide risks among our participants.

As mentioned above, Facebook has set forth Community Standards for which they take action for specific online behaviors. Among these they have specific policies regarding "violence and threats" as well as "self-harm". The information from the Facebook Community Standards website is pasted below and can be accessed at https://www.facebook.com/communitystandards. Additionally, the information related to illegal activities and self-harm is set forth in the privacy policy https://www.facebook.com/about/privacy/.

Based on the actions taken by Facebook regarding self-harm, homicide, as well as potential illegal activities, it seems unlikely that data from participants or secondary subjects will contain these concerning behaviors. Additionally, Facebook provides access to suicide prevention programs for any user that may be suicidal. Thus, we feel that Facebook addresses this issue before the data would even come to us (see below community standards for more details).

From the Facebook Community Standards Page:
"Violence and Threats
Safety is Facebook's top priority. We remove content and may escalate to law enforcement when we perceive a genuine risk of physical harm, or a direct threat to public safety. You may not credibly threaten others, or organize acts of real-world violence. Organizations with a record of terrorist or violent criminal activity are not allowed to maintain a presence on our site. We also prohibit promoting; planning or celebrating any of your actions if they have, or could, result in financial harm to others, including theft and vandalism.

Self-Harm
Facebook takes threats of self-harm very seriously. We remove any promotion or encouragement of self-mutilation, eating disorders or hard drug abuse. We also work with suicide prevention agencies around the world to provide assistance for people in distress.

From the Data Use Policy Page/Privacy Policies:
Responding to legal requests and preventing harm
We may access, preserve and share your information in response to a legal request (like a search warrant, court order or subpoena) if we have a good faith belief that the law requires us to do so. This may include responding to legal requests from jurisdictions outside of the United States where we have a good faith belief that the response is required by law in that jurisdiction, affects users in that jurisdiction, and is consistent with internationally recognized standards. We may also access, preserve and share information when we have a good faith belief it is necessary to: detect, prevent and address fraud and other illegal activity; to protect ourselves, you and others, including as part of investigations; and to prevent death or imminent bodily harm. Information we receive about you, including financial transaction data related to purchases made with Facebook Credits, may be accessed, processed and retained for an extended period of time when it is the subject of a legal request or obligation, governmental investigation, or investigations concerning possible violations of our terms or policies, or otherwise to prevent harm. We also may retain information from accounts disabled for violations of our terms for at least a year to prevent repeat abuse or other violations of our terms."

IV. Benefits to participants or future benefits

It is unlikely that participants will receive any direct benefits from their participation in the study. The participants will benefit indirectly since the results of the research will be used to design interventions aimed at preventing high-risk behaviors (HIV, substance use, mental health crisis) that can be executed through a social networking technology based platform for this otherwise hard-to-reach population.

V. Confidentiality:

Data safety and monitoring plan

All possible measures to ensure the anonymity and confidentiality in this project will be taken. This will involve taking multiple steps and be done separately for self-administered survey data and observational data.

Once we have a finished collecting self-report data from participants, identifying information will be separated from the data collected. Storing identifiers will be as follows: each subject is assigned a **personal identifier (PID),** this is recorded on a separate database on a separate computer with their name and contact information (i.e. emails and Facebook username); no data will be attached

to this identifying information. The database with the identifying information will be destroyed after the study is concluded. The electronic data (stripped of identifying information) will be stored in an encrypted database and PI and study personnel will have access to the full database.

In regards to the observational data retrieved through Facebook, we will follow similar steps as outlined above. The web-based interface created to collect youth's Facebook data will have names, Facebook usernames and unique Facebook URLs. These data will be stored on a password-protected server at the University of Denver Graduate School of Computer Science. Once Facebook data is retrieved from the web-based application, it will be downloaded into a SQL enabled database. The initial data will have identifying information for participants. However, once linkages to PID numbers are created identifying information will be removed. These data (collected by the app) will be stripped of identifiers before being matched by PID to the survey data, so that no identifying information is ever included in the data bases which contain private information about sexual behaviors or HIV, HCV and STI status. The de-identified data will be kept indefinitely.

Only members of the research team who 1) must access data to complete research-related tasks; and 2) have signed a confidentiality agreement will have access to data. The PI and Co-PIs will have full access to data and will oversee access by the research team. Except in cases of mandated reporting, information about individual participants will not be released. Aggregate findings across all participants will be provided in a report to Urban Peak. This report will not identify individual participant responses.

Due to the sensitive nature of the data collected, we will also apply for a **Certificate of Confidentiality from NIH** to protect research participants against compulsory legal demands for identifying information or identifying characteristics of a research participant.

References

1. Moreno, M. A., & Kolb, J. (2012). Social networking sites and adolescent health. Pediatric Clinics of North America, 59(3), 601-612.

2. Huang, G. C., Unger, J. B., Soto, D., Fujimoto, K., Pentz, M. A., Jordan-Marsh, M., & Valente, T. W. (2014). Peer influences: The impact of online and offline friendship networks on adolescent smoking and alcohol use. Journal of Adolescent Health, 54(5), 508-514.

3. Barman-Adhikari, A., Yashioka-Maxwell, A., Bender, K., Hall, R. & Rice, E. (in press). Social Networking Sites and Sexual Health behaviors among Homeless Youth. Journal of Health Communication.

4. Pollio, D., Batey, S., Bender, K., Ferguson, K., & Thompson, S.J. (2013). Technology use among homeless young adults in two U.S. cities. Social Work, 58, 173-175.

5. Barman-Adhikari, A. (2015). What do youth want out of social media based interventions? Unpublished Data. DU IRB Project #: 836982-1

6.Barman-Adhikari, A., & Rice, E. (2011). Sexual health information seeking online among runaway and homeless youth. Journal of the Society for Social Work and Research, 2(2), 88-96.

7. Rice, E., Monro, W., Barman-Adhikari, A., & Young, S. D. (2010). Internet use, social networking, and HIV/AIDS risk for homeless adolescents. Journal of Adolescent Health, 47(6), 610-613.

8. Young, S. D., & Rice, E. (2011). Online social networking technologies, HIV knowledge, and sexual risk and testing behaviors among homeless youth. AIDS and Behavior, 15(2), 253-260.

9. Facebook (2014). Facebook Login and Login Review. https://developers.facebook.com/docs/facebooklogin/ v2.2.

10. Moreno, M. A., Arseniev-Koehler, A., Litt, D., & Christakis, D. (2016). Evaluating college students' displayed alcohol references on Facebook and Twitter. Journal of Adolescent Health, 58(5), 527-532.

11. Brockman, L. N., Christakis, D. A., & Moreno, M. A. (2014). Friending adolescents on social networking websites: a feasible research tool. Journal of interaction science, 2(1), 1-9.

12. Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. American Psychologist, 70(6), 543.

13. Moreno, M. A., Goniu, N., Moreno, P. S., & Diekema, D. (2013). Ethics of social media research: common concerns and practical considerations. Cyberpsychology, Behavior, and Social Networking, 16(9), 708-713.

14. Solberg, L. B. (2010). Data mining on Facebook: A free space for researchers or an IRB nightmare? Journal of Law, Technology and Policy, 2010(2).

15. Dwyer, C., Hiltz, S., & Passerini, K. (2007). Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. AMCIS 2007 proceedings, 339.

16. Stillwell, D. J., & Kosinski, M. (2015). myPersonality Project Website.

17. Behavioral research (2016). Reportable events. Retrieved October 21st, 2016 from https://irb.research.chop.edu/behavioral-research