Social Networks 66 (2021) 114–124

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Network Canvas: Key decisions in the design of an interviewer-assisted network data collection software suite

M. Birkett [a,*], J. Melville [a], P. Janulis [a], G. Phillips II [a], N. Contractor [b], B. Hogan [c]

[a] Northwestern University Feinberg School of Medicine, Department of Medical Social Sciences, Chicago, IL, United States
[b] Northwestern University Kellogg School of Management, Department of Management and Organizations, Evanston, IL, United States
[c] University of Oxford, Oxford Internet Institute, Oxford, UK

ARTICLE INFO

Keywords:
Networks
Data collection
Software
Social networks

ABSTRACT

Self-reported social network analysis studies are often complex and burdensome, both during the interview process itself, and when conducting data management following the interview. Through funding obtained from the National Institute on Drug Abuse (NIDA/NIH), our team developed the Network Canvas suite of software – a set of complementary tools that are designed to simplify the collection and storage of complex social network data, with an emphasis on usability and accessibility across platforms and devices, and guided by the practical needs of researchers. The suite consists of three applications: Architect: an application for researchers to design and export interview protocols; Interviewer: a touch-optimized application for loading and administering interview protocols to study participants; and Server: an application for researchers to manage the interview deployment process and export their data for analysis. Together, they enable researchers with minimal technological expertise to access a complete research workflow, by building their own network interview protocols, deploying these protocols widely within a variety of contexts, and immediately attaining the resulting data from a secure central location. In this paper, we outline the critical decisions taken in developing this suite of tools for the network research community. We also describe the work which guides our decision-making, including prior experiences and key discovery events. We focus on key design choices, taken for theoretical, philosophical, and pragmatic reasons, and outline their strengths and limitations.

Introduction

Problem statement

The collection of network data is a high priority for researchers across a variety of domains, but it also poses enormous methodological and practical challenges (Valente and Pitts, 2017; Marsden, 2011; Perry et al., 2018; Adams, 2019). For example, numerous studies have demonstrated the sensitivity of network data to research design decisions, such as the choice and phrasing of specific name generators (Adams, 2019) or the ordering of items (Hlebec and Ferligoj, 2002). The quality and reliability of network data has also been shown to be highly impacted by the subjective aspects of the interview experience (Tourangeau et al., 2000), such as a participant's perception of completion burden (Silber et al., 2019; Mc Carty et al., 2007), their sense of study confidentiality or the sensitivity of the data to be captured (Perry et al., 2018). Equally important, but somewhat less documented in the research literature, are the upstream challenges experienced by network researchers who wish to quickly design, deploy, and analyze the data from their studies. For example, there are few survey softwares which are designed to capture social network data and consequently, many researchers rely on in-person paper and pencil interview techniques which can also be prohibitively slow and resource-intensive (Perry et al., 2018; Mc Carty et al., 2007; Herz and Petermann, 2017). Access is even further limited as even the simplest network dataset often requires extensive manipulation to get to the point of analysis (Adams, 2019) –which can be a major barrier for researchers uninitiated in working with graph structures. These challenges are only amplified when researchers wish to capture "complex" network data, such as that which includes multiple types of actors and relationships across multiple points in time (Adams, 2019). Because of these challenges, this work is often limited to researchers with substantial training in network science,

* Corresponding author at: Department of Medical Social Sciences, CONNECT Complex Systems and Health Disparities Research Program, Northwestern University Feinberg School of Medicine, 625 N Michigan Ave, Suite 1400, Chicago, IL, 60611, United States.
E-mail address: birkett@northwestern.edu (M. Birkett).

https://doi.org/10.1016/j.socnet.2021.02.003

Available online 24 February 2021

extensive resources, or the expertise to create bespoke software solutions.

Objective

In this paper we detail how the creation of a suite of tools for collecting social network data (termed Network Canvas, 2021) emerged from and speak to a variety of the methodological and organizational challenges that exist in the implementation of social network studies. We describe how the design of this suite of applications responded to many of these challenges both through a philosophical approach to interviewing focusing on co-construction of social network data and through various means of engagement with the relevant academic communities. We articulate how this design path has facilitated an integrated series of applications for network data, but in doing so reveals a variety of constraints, some of which are still considered out of scope for our software, but not for social network data collection. In articulating this approach, we seek to be reflexive about our design in relation to the knowledge produced from such techniques. In designing this software, it was apparent that we could not meet every need, although we aspired to create a tool with broad applicability in social network studies. This reflexivity leads us to consider what sorts of work are or are not possible using such a system, and thus what sort of claims can or cannot be made from this approach.

The software suite resulting from our design process and community engagement is called Network Canvas and is comprised of three applications:

Architect: an application for researchers to visually design interview protocols, without the need for coding or scripting.

Interviewer: an application which allows researchers to administer interview protocols to study participants in a controlled environment.

Server: an application for researchers to manage the interview deployment process, gain real-time insight into the progress of their data collection, and export data for analysis.

These three applications work together to provide an integrated workflow from interview design up to the point of analysis. Researchers start in the Architect Application where they build an interview protocol. Then, researchers move between the Server and Interviewer Applications to deploy and administer that interview protocol, as well as collect and manage participant data. Finally, data are exported from Server and are available in a number of formats for further analysis. Fig. 1 provides a representation of this workflow.

Within this paper we provide further details on the challenges which have limited the capture of network data, and describe our efforts to design software that might increase the ease and efficiency for both researchers and participants while also discussing the limitations and constraints of our approach.

This Network Canvas workflow was designed to meet a variety of scenarios as articulated in social network research. However, through our engagement with the network community and our reflection on specific design choices, we understand that our tool is not ideal for all circumstances. The design of our tool both enables and constrains particular research applications. We hope that by critically explicating our influences and our design process, we will not only provide a greater understanding of our tool and the boundaries of it, but we will also inspire reflexivity within the social network community on how priorities shape the software we use to conduct our work – and how these impact which questions are able to be asked and by whom, and the underlying quality and the structure of the resulting data.

The paper proceeds as follows. We first articulate our prior experiences and our engagement with relevant stakeholders. We then discuss the challenges for participants in the deployment of social network studies and the ways we sought to address these challenges. They mainly concern the in-situ experience and impacts on data quality. We then turn to researcher-facing challenges that mainly concern accessibility and efficiency. We then reflect on how our particular approaches compare and contrast with existing software in the field. We highlight how different software is oriented to meet different needs of researchers. We conclude by suggesting next steps for software development for social network studies broadly.

Experiences and discovery activities which inform our design

To contextualize both our understanding of the challenges of network data collection and our approach to the design of a solution, we first articulate two studies by the development team that directly inspired the current incarnation of our software. Network Canvas emerged through a collaboration to design a network data capture tool that would meet the personal network data capture needs of these two specific and very different projects.

Study 1

At Northwestern University, a longitudinal cohort study – RADAR – required personal network data collection from over 1,000 participants who would provide data every 6 months for up to 5 years. The scale and longitudinal nature of this project required an efficient method to capture social, sexual, and drug-use connections between egos and alters, and alters and alters. While researchers on this project had previously successfully employed analog participant-aided sociograms (PAS), for example using whiteboards, magnetic 'poetry' cards and dry erase markers to capture network details (Kuhns et al., 2015), this process introduced a number of downstream challenges for researchers. For example, ties were captured via photographs of the whiteboard, which required researchers to enter the data at a later point. And the capture of multiplex ties required multiple photographs to be entered for each study participant, further slowing an already tedious translation. Therefore, researchers at Northwestern were interested in using digital tools which retained the strengths of PAS, but streamlined the workflow from data collection to analysis.

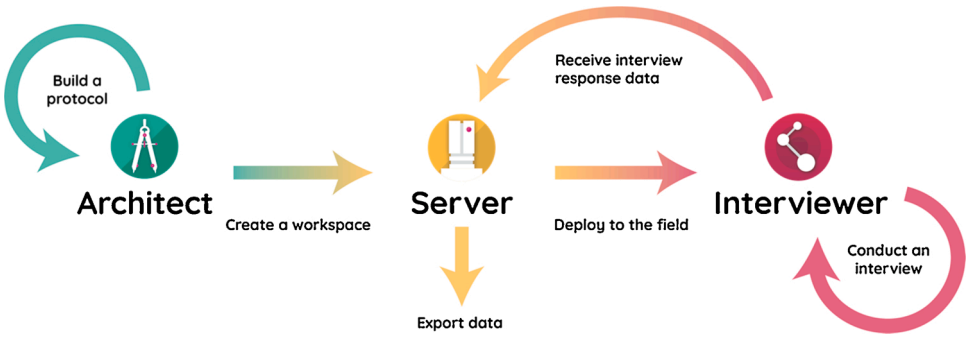

Fig. 1. The three applications of Network Canvas providing an integrated workflow.

Study 2

At University of Oxford, researchers were developing a digital framework to extend paper-based approaches to social network data collection. This was being developed to qualitatively explore the role that the composition and structure of personal networks play in the outcomes of people with long-term disabilities. The study researchers wanted to minimize the complexity of the interface, while providing large intelligible features that might be easily interacted with on large touchscreens by individuals with significant accessibility needs or cognitive impairments. Further, this study focused on in-depth discussions of the multiplex relations between family members, peer support, carers, and other relevant members of a person's personal community. Thus, there was a considerable emphasis on how to make the digital presentation of social network data not just intelligible, but sufficiently flexible to foster conversations involving multiple actors relevant to the respondent. This might be contrasted with pen-and-paper diagrams that are generally fixed once the nodes and edges are laid down. Early prototypes built a flexible perceptual 'map' of the individual's personal network, which would be used as an elicitation device within the qualitative interview.

The initial collaboration across these two study teams led to the production of a software framework (further detailed in Hogan et al., 2016) which drew from both analog and digital approaches. While this approach was a solution to each team's immediate study needs, it was not scalable to the broader social network community. In particular, limited development resources meant the prioritization of features related to the participant-facing front-end of the application with compromises made at the back-end. For example, early versions of Network Canvas were self-contained applications that embedded both the specific interview protocol and the general technology in the same container. One could not load a different protocol or otherwise alter the look and feel of the interview instrument without entirely rebuilding the software. To a limited extent, it was possible to go and hand-edit files in the application, but this required prior experience with extensive features of HTML5 and graph database technologies. This limited both our ability to modify the survey instrument, and other's ability to utilize the tool. Furthermore, the software utilized data structures that were amenable to rapid storage and editing within the program, but that required greater wrangling prior to analysis, further limiting the tool to only those with extensive expertise in parsing of network data structures. Therefore, our team pursued a new collaboration to extend Network Canvas so that it might be scalable for the broader social network community, able to be customized for a variety of research settings while still allowing for the secure and simple management of social network data.

It was important to ensure that the design of Network Canvas reflected the needs of the broader social network and public health community. Therefore, key discovery activities were built into the software development plan to aid in the prioritization of a variety of features. These activities were not intended to be systematic investigations, but instead to be used by the team to understand the activities and characteristics of current and potential social network researchers in order to inform the tool's development. Activities included the formation of a Scientific Advisory Board and ongoing relationships with two University test sites who received funding to work collaboratively with our team. Activities also included an anonymous survey conducted of the broad social network and public health research community. Utilizing initial information gathered in that survey, three targeted focus groups were then conducted. And finally, numerous software demonstrations and feedback sessions were held between 2017 and 2018 - three at Network Conferences, six conducted during visits to specific network-focused research labs, and one to the Centers for Disease Control and Prevention. In addition to this work, we also maintained an active presence at social network and public health conferences, reviewed alternate and existing software for network data collection, and learned through pro bono support, where possible, to early adopters of Network Canvas within their own research studies and teaching endeavors.

While all of these experiences have played a role in informing our understanding of data collection challenges and our approach to research-facing design, as we specifically reference insights obtained from the anonymous survey of the social network community, we next provide additional detail about that data collection.

Survey of the social network community

The survey of the social network community was meant to provide our team a broad understanding of the challenges facing researchers who either utilize or wish to utilize social network data. Survey invitations were distributed via six social network listservs or social or behavioral health listservs which frequently discussed network data collection (e.g., INSNA's SOCNET, the Duke Network Analysis Center, the American Evaluation Association, the College on Problems of Drug Dependence, the Society for Community Research and Action, and the American Public Health Association HIV/AIDS Section), via listservs for the 19 NIH-funded Centers for AIDS Research sites, and via 22 personal invitations to leaders in the network community. The survey was administered online via Qualtrics. Questions were asked about network training experiences, access to and experience with collecting network data, familiarity with specific network analysis tools or packages, researcher access to hardware that might be used to capture network data, and the key hurdles experienced in the collection of personal network data. Finally, the respondent's willingness to be contacted for a follow-up interview was assessed.

Surveys were completed by 181 individuals, with 156 of those individuals indicating that their anonymized responses could be reported upon in future research studies. Of those 156, approximately 68 % identified themselves as faculty in academic institutions and 16 % as current students. Approximately 62 % indicated having collected personal network data before. Key findings are discussed under Researcher-facing Design Considerations.

Next, we draw upon the existing literature and the information gathered by our team to describe our understanding of the challenges which have limited the capture of network data and how these have factored into our software design. We have split the discussion across two groups, first describing design considerations for participants, then for researchers.

Participant-facing design considerations

Numerous studies have shown that network data quality suffers when participants are confused, bored, or lack trust in the interview process. For example, as indicated by Perry, Pescosolido, Borgatti (Perry et al., 2018), all survey research is impacted by social desirability bias, satisficing – or providing low quality responses when a participant perceives the survey as too cognitively demanding, and interviewer effects. However, as network capture is frequently perceived by participants as cognitively demanding, as well as highly invasive, the potential for poor data quality increases (Herz and Petermann, 2017; Vehovar et al., 2008). Furthermore, the reliability of data is highly impacted by research design decisions. Network data can be influenced by the order of name generators (Yousefi-Nooraie et al., 2017), the method of eliciting alter-to-alter ties (Eddens and Fagan, 2018), or even the visual stimulus used to capture this data (Hollstein et al., 2020). Accordingly, the inherent complexity of this process demands tools that are highly usable and minimize cognitive burden on participants in effort to maximize data validity.

Our approach to participant-facing design

In an effort to improve the experience of participants, our work draws from a long history of visual approaches to personal network data

capture, and visual methods in social science more broadly. As comprehensively reviewed by Freeman, visualization has been central to the capture and study of social relationships since at least Moreno's earliest work on sociograms (Freeman, 2000). Through visually displaying the underlying network, these approaches allow participants to efficiently access and interact with the most fundamental form of their personal network. Concretely, this means that participants can react to the entirety of nominated alters as a visual "set," which allows them to generate intuitions about the network object, and better negotiate complexities such as alter-alter linkages and network boundaries. Furthermore, allowing participants a view of their network as a coherent object has been shown to democratize the research in a way that leads to greater engagement, participation, and personal investment in the interview (Baum et al., 2006).

An early and notable example of the application of visual methods to personal network capture is Antonucci's bullseye diagrams for studying social convoys across the life course (Antonucci, 1986). Pahl and Spencer (Pahl and Spencer, 2004) expanded Antonucci's work by enabling people to arrange their friendships in concentric circles with Post-Its. One of the challenges of this approach, however, was how to manage the capture of indirect alter-alter ties, and how to do this for a large number of alters. For example, a single relational network of 20 alters might lead to 190 additional queries. The PAS approach (Carrasco et al., 2008; Hogan et al., 2007) was developed as a response, to take full advantage of the intuitiveness of visual approaches, while capturing indirect alter-alter ties in a less burdensome fashion. Similar to Pahl and Spencer (Pahl and Spencer, 2004), researchers in Wellman's NetLab used Post-Its, but they also captured indirect ties by drawing lines and encircling nodes. This approach was shown to be usable and intelligible to research participants (Hogan et al., 2007), as well as to allow participants to deeply engage with their data (Dubois and Ford, 2015).

While in some respects analog methods like the PAS are quite intuitive and simple, they can also be quite resource intensive. For example, once the ink dries, the initial flexibility of the medium disappears. It becomes difficult to modify the PAS without starting again, making the capture of multiple relational ties difficult. Indeed, our initial investigations confirm that participants take advantage of the malleability of digital sociograms, most notably by rearranging their graphs in ways which appear to make them more intelligible to themselves, according to some coherent principles related to latent social structures (Hogan et al., 2017). Furthermore, analog data capture approaches require that researchers invest significant time and resources in coding and entering the data into a database, often including techniques such as double-entry and review of audio/video transcripts to aid with ambiguous coding (Kuhns et al., 2015).

The challenges with coding and preserving visual analog data motivated researchers to explore computer-assisted approaches to network data collection. A number of these alternative approaches have been developed over the last 20 years, with most of them allowing for the immediate creation of structured data and greatly improving efficiency. A few notable examples are Mc Carty's *EgoNet* (Egonet, 2021), *EgoWeb 2.0* (Kennedy, 2021), *Anamia* (ANAMIA Egocenter, 2013), *GENSI* (Stark and Krosnick, 2017), and *VennMaker* (Vennmaker, 2021), many of which allow for visualization and real-time representations of personal networks (Table 1) (Mc Carty and Govindaramanujam, 2005). Network Canvas combines some of the key strengths of both the traditional PAS and these initial software – borrowing visual elements of the PAS to both represent and input relational data while also allowing the efficient capture of structured data. But in an effort to improve research participant experience, we have also incorporated best practices in interface design – such as the use of physicality (Norman, 2007), HCI principles (Macaulay et al., 2009) such as Heer & Shneiderman's rules for interactivity (Heer and Shneiderman, 2012), and design simplicity (McGrenere et al., 2002) – in order to produce a software which is easily understood and navigated and builds the trust and autonomy of our users. We believe this focus on interface design allows our tool to more fully replicate the interactivity and engagement of a physical PAS within research software.

For example, as seen in Fig. 2, although Network Canvas utilizes a screen, it intentionally draws upon tactility, micro-interactions, and the physical metaphors of drawing, dragging, and binning to provide participants with multiple means to indicate features either of their social network or of their relationships to network members. Relational ties are created by tapping on one node and then another. Network members are added and represented in the interview by nodes which can be dragged and rearranged by the participant. Information about relationships can be generated by dragging nodes into bins for categorical items or for ordinal items performing a kind of "card sort" by dragging nodes onto a contiguous color gradient scale. Outside of social networks, researchers have long utilized visual and tactile approaches to facilitate conversations about abstract and complicated phenomena – like identity and social experiences. Where language alone might be limiting, the use of visual metaphor allows better communication about how complicated ideas or systems fit together. For example, physical approaches have been shown to simplify engaging with complex ideas by making them tangible, while also increasing engagement and reducing tedium. Gauntlett (2014) demonstrated success using Lego blocks and props with participants so that they could create scenes that represented their identity, while bringing in figurines as network members. Similarly, Guggenheim et al. (2013) used a literal sandbox with toys with participants to explore strategies for imagining disaster scenarios and means for emergency provisions. By imbuing our tool's interface with physical metaphors, we hope to allow anyone encountering our software for the first time to quickly understand how to use it, thereby allowing the elicitation of a social network to be a simple and engaging process with minimal interviewer-oversight necessary.

Another user-centered design principle which we have prioritized is design simplicity (Macaulay et al., 2009; McGrenere et al., 2002), or avoiding the unnecessary inclusion or foregrounding of any technical features which are not relevant to data capture and which could potentially cause confusion for research participants. For example, in order to maximize the intuitiveness of the software for research participants, we have de-emphasized the development of some features often requested by network researchers, as these might confuse a participant with no familiarity with social network analysis. Features like automatic resizing of nodes and automatic layouts based on connectivity – which alter the look or position of a node based on something outside the control of the participant – have been intentionally omitted as they may lessen the ability of the participant to understand what is displayed as well as undermine the participant's sense of control of the interview.[1] Thus, Network Canvas interfaces have been designed to minimize clutter and guide the user through the interview protocol.

Our focus on providing the participant with an excellent research experience also led to Network Canvas being interviewer-assisted – or that the interview must occur in the presence of an interviewer and on an interviewer-controlled machine. While this fundamentally limits the utility of our tool for researchers who are unable to conduct in-person data capture, the higher priority was to maintain the interview experience and trust of research participants. For example, a design which allowed data capture on cell phones would have limited our ability to ensure high quality visual and tactile elements. The small screen size would impact readability as well as the number of nodes that could be displayed on a screen. Further, touch engagement is notoriously finicky to optimize across a wide-range of mobile devices, risking interface inconsistencies, participant frustration, and mitigating data quality. Finally, after the interview, researchers would have little control over

[1] Although node size cannot be altered to indicate differences, we have designed other interfaces where relative importance, strength, or frequency can be easily reported. For instance, the ordinal bin interface allows for indicating valued differences between nodes.

Table 1
Network Data Capture Software Comparison Across Key Areas.

Software Name	Technical Foundation	Mode of Administration	Flexibility	Co-creation	Workflow	Openness
Anamia (ANAMIA Egocenter, 2013)	Adobe Flash	Self	Designed for a specific use case. Not able to be customized easily.	Focused entirely around interactive visual construction of sociogram.	Does not provide assistance for interview creation or data management.	Uses proprietary technology
EgoWeb 2.0 (Kennedy, 2021)	JavaScript	Interviewer or self	Has been used for a wide variety of data collection projects globally. Allows user to configure types of networks collected.	Sociograms are able to be rendered with data collected during interview.	Runs as a traditional webserver, which may need technical expertise; promotes easy data export in familiar formats.	GNU General Public License v2.0 (GNU General Public License, 2007)
Enso[a] (Enso, 2021)	JavaScript	Self	Flexibility in interview prompts, but not in node types.	Can include interactive construction of sociogram	Primary development focus is client-server architecture, not providing overarching workflow.	GNU General Public License v2.0 (Kennedy, 2021)
GENSI (Stark and Krosnick, 2017)	HTML5 and JavaScript	Self	Has been used across large-scale online studies. Works best with larger displays and networks smaller than eight.	Can include interactive construction of sociogram	Requires knowledge of javascript and server administration	Unknown (Enso, 2021)
Network Canvas	CSS3, HTML5, and JavaScript	Interviewer-assisted, participant-led	Allows multiple node and relational types, cognitive social structures, roster-input.	Can include interactive construction of sociogram. Participant intelligibility prioritized.	Integrated; no special scripting required	GNU General Public License v3.0 (codaco, 2021)
Trellis (Trellis, 2020)		Interviewer or self	Designed for use with multiple devices; relies on straightforward design and standard controls	Less focus on interaction with network, alter-alter ties, and sociogram	Robust project management is a design priority, meaning a straightforward and linear workflow	MIT (Trellis, 2020)
VennMaker (Vennmaker, 2021)	Javascript	Interviewer	Default settings encourage use of nodes as people, but flexibility exists	Focused around interviewer-led construction of interactive sociogram.	Server-side implementation (MyNetworkMap) to support back-end data management	MIT (formerly closed source) (Vennmaker, 2021)

[a] Formerly known as OpenEddi.

Fig. 2. Participant building their sociogram within the Network Canvas Interviewer App.

the transmission of potentially sensitive data over unsecured networks. For many researchers, the in-person and interviewer-assisted requirements may be an insurmountable limitation. Many researchers rely solely on remote survey administration because of their population of interest, such as highly stigmatized populations where the increased anonymity of remote capture may ease recruitment. Others may rely on remote data capture because it tends to require less overhead. The 2020 pandemic has all but halted human subjects research, further emphasizing the importance of the future expansion of our tool for remote data capture. However, our initial design of our tool to allow interviews to be conducted on a research participant's device – across a number of Internet browsers, screen sizes, resolutions, and networks – would have insurmountably sacrificed our ability to ensure useable interfaces and secure data.

Validating the participant-facing design

Initial validations of our participant-facing digital interfaces have shown research subjects to find Network Canvas useable and allow for the production of high quality data (Hogan et al., 2016, 2019). For example, research participants completing Network Canvas produce a comparable number of personal contacts and sexual partners as those captured via PAS and other validated measures of sexual risk behavior, but that data collection took substantially less time to complete (Hogan et al., 2016). Data generated by Network Canvas has also been shown to be reliable over time, with no significant differences in the number of edges generated and only slight differences in the number of alters reported (Hogan et al., 2019). Furthermore, contrary to what might be expected based on other longitudinal network research, we found that the overall number of contacts increased between two longitudinal waves, suggesting that the tool did not (in that context) suffer from motivated underreporting (Hogan et al., 2019). And finally, participants rated their interactions with the tool overall, as well as their interactions with specific interfaces, as highly useable (Hogan et al., 2016). Quotes from participants about their experience with Network Canvas or any suggestions included "No, I really enjoyed my time."; "It was pretty fun using the technology!"; "Completely comfortable"; and "It all was pretty cool to see visually". Because the feedback was positive and no issues were raised, these validations did not have an impact on participant-facing digital interface design.

Researcher-facing design considerations

Researchers also face challenges in the design, deployment, and analysis of social network studies. While, in the last several years extensive discussions of the capture of ego-centered data have appeared (Perry et al., 2018; Adams, 2019; Crossley et al., 2015; Mc Carty et al.,

2019) which have included at least some discussion of these challenges, overall far less attention has been paid to documenting these upstream challenges in the research literature. While our understanding of the field draws from this existing literature, by necessity our knowledge of the challenges inherent in network data capture have been informed by both our development team's personal experiences with data capture and our team's interactions with the broader social network community. Within the next sections we will highlight the major challenges identified through these various sources and how our software has been shaped in response.

Across our work, the most salient and consistent finding was a great need to simplify the process of designing a network interview protocol and managing the data obtained from such interviews. Both from our own experiences and from what has been reported over numerous discovery activities within the social network community – researchers struggle to not just capture social network data, but to make it analysis-ready as quickly as possible. For example, in our survey of the social network community we found that the vast majority of researchers struggled with developing network instruments, deploying network instruments in the field, and getting their network data into an analysis-ready form. As an illustration of this, only 4.8 % *Strongly Agreed* that "Managing and accessing personal network data is simple" while 37.5 % *Strongly Disagreed*. One of the primary drivers of this difficulty appears to be the technical expertise required to work with or work around the limitations of existing tools. For example, 59 individuals responded to the open-ended item '*In your opinion, what are the key current hurdles in the collection of ego-centered data collection?*' and numerous mentioned difficulties due to technical complexity and lack of access to simple tools. For example:

"Accessing simple data collection tools"
"Having access to free/low-cost and modifiable electronic data collection tools"
"Lack of knowledge and lack of software to help make it easy"
"Efficient programs that can collect data in a clean user-friendly way but also the output is clean and easy to use."
"The biggest challenge is the output portion."
"Preparing analysis-ready data sets"

We found these challenges to be amplified in social and behavioral health researchers. For example, across numerous site visits we found that these researchers frequently relied upon data managers or analysts who had little expertise in working with network data structures or datasets which require extensive manipulation. Furthermore, despite the efficiencies introduced by utilizing a precursor to Network Canvas within our own RADAR longitudinal cohort study, our team was still required to create and maintain an extensive data management plan before data was clean, accessible, and analysis ready.

Based on this knowledge of the community and our own experience – we prioritized the creation of a free software tool which allowed researchers to easily design their own survey and allowed the immediate receipt of structured data – while removing many barriers of technological literacy. This was especially important for individuals whose primary training may not have been within social networks, but who wanted to begin to incorporate network data capture into their studies.

Beyond designing a tool which simplified the design of survey instruments and receiving data, another challenge we encountered was the diversity of topics and approaches used in the network community. Therefore, it was important for our team to design a tool flexible enough to accommodate the majority of researchers. For example, researchers who capture complex quantitative data as well as qualitative researchers who utilize networks more as an object to be reflected upon (Ryan et al., 2014; Edwards, 2010). Furthermore, we wanted to meet the needs of both researchers who utilized name generators to elicit personal network data as well as those who utilized "rosters" or who conducted recognition-based studies where participants chose network members from predetermined lists.

Another challenge we identified was that many existing academic software – for pragmatic reasons – have been built for specific studies rather than as customizable general tools for a wide social network audience. For example, *Anamia* (ANAMIA Egocenter, 2013) was designed specifically for a study of eating disorders. While understandable, designing for a single study limits flexibility as software may overly prioritize the needs of certain types of researchers over the needs of a more general research audience. Furthermore, these tools are often built by small teams who are unable to continually update and maintain the technology. And, particularly if they aren't developed openly and are locked behind paywalls or within a researcher's archives, often these tools will not receive continual maintenance or be inaccessible except to those with extensive technical knowledge.

Our approach to researcher-facing design

Designing for efficiency, ease, and access

Due to the above challenges, many of Network Canvas's design decisions have been shaped by our prioritization of increasing the access, the ease, and the efficiency of network data capture, for researchers - as well as participants. Making Network Canvas usable for a wide audience thus involved designing a means to flexibly build studies in Network Canvas that was as intelligible to researchers as the Network Canvas interview experience was to participants, as well as building a way to consolidate and securely manage data coming out of Network Canvas that lowered the barriers normally faced with wrangling high dimensional social network data. Next, we detail a few specific design decisions and how they work to increase the access, ease, and efficiency of network data capture for social network researchers.

Our first priority was simplifying the workflow for researchers, by building the Software Suite around three applications which were interoperable and streamlined the workflow. Survey protocols are designed in Architect and then deployed and managed in Server. Interviews are administered in the Interviewer App and data is centrally received by Server where it can then be output for analysis and modelling.

While study participants only interact with the Interviewer App, researchers utilize all three applications. We believe this integrated workflow increases efficiency, as it allows researchers to immediately create a protocol, administer it, and export the resulting social network data on the very same day - resolving the need for manual data entry. Furthermore, this workflow allows that once an interview protocol is designed, a researcher can immediately deploy it to a number of linked field devices – an especially important feature in RADAR and other large NIH-funded studies. Interviews can then be run on the field device without the need of Internet access. Then, after the interviews are run, each of these field devices can then immediately securely transfer encrypted interview data from the device to the Server Application. Server not only is a host for study data, but it also provides dashboard of summary descriptives on completed interviews and allows the researcher to export a single data file with all observations collated.

A second priority was designing our tools to require low technological literacy. For example, like the Interviewer App, the Architect App also utilizes easily intelligible graphical user interfaces with easily comprehensible elements. Interview protocols are built by assembling protocol stages onto a physical timeline that reflects the order of interview stages as they are displayed in the Interviewer App. The Server App also contains visual interfaces, including a dashboard to immediately monitor key metrics of surveys conducted. Finally, the Server App allows for the easy export of data into familiar file formats like .csv and .graphml. We intentionally avoided software design which would require scripting or complicated server configuration.

In another attempt to increase access to network data capture, we prioritized creating a flexible suite of tools that had as few constraints on the type of interview that could be designed, as possible. This meant

attempting to consciously resist embedding any prescriptive notions of how to best capture social network data within our software design. Instead, we built the tool by focusing on the most fundamental building blocks of networks – nodes and edges – while leaving the specific nature of the operationalization, framing, and capture of these entities within the interview entirely under the researcher's control. This allows our software to readily accommodate research where nodes may represent anything from places, to concepts, to social contexts, to people.

To better illustrate how this decision impacts the software design in practice, consider how the process of creating an interview within the Network Canvas software compares to other survey tools. Of course, Network Canvas has many overlapping features with standard survey software (e.g., Qualtrics and SurveyMonkey), such as built-in methods for capturing ordinal, categorical, numerical, and textual data. However, since Network Canvas is built to be "network-centric" (as opposed to the "variable-centric" approach common in other survey software), a researcher building an interview protocol in Network Canvas will first designate the node and edge types that will comprise the network, and from there will choose the appropriate interface screens to capture and describe these network entities. This means that where standard survey software would begin with the network researcher creating a series of items, along with an elaborate "loop and merge" logic to artificially simulate a network data model, Network Canvas structures the entire interview creation process around the network entities themselves, with all other data collection happening relative to these entities.

Once these entity types are defined, the process of creating an interview in Network Canvas is further structured around the core tasks common to most network interviews: (1) name generation, (2) name interpretation, and (3) creation of alter-alter ties. Each of these is accomplished through the selection and configuration of a series of interfaces. Once again, flexibility was prioritized. While each specific interface was optimized for a particular type of task (e.g., name generation; edge creation; ordinal binning; narrative interviewing), many other aspects of how an interface behaves are configurable. In the case of name generation, researchers are first able to choose between using roster data, allowing free nomination, or using both systems simultaneously. If using roster data, dedicated interfaces exist in Network Canvas that are designed specifically for working with rosters of different sizes, which implement filtering, sorting, and searching as needed, in a configurable fashion. For researchers not utilizing rosters, there is once again complete flexibility in the way that nodes are created: both an ultra-quick nomination interface is available (when only a name or label is required), or a fully functional form system can be used, which supports collecting any number of additional attributes at the point of elicitation.

While these interfaces themselves are not particularly novel or innovative, their power is in their ease of configurability so that researchers are able to easily customize study protocols that meet their exact needs and style of measurement. For example, within Fig. 3 we demonstrate two potential approaches for obtaining valued alter-alter ties – (A) where ties between alters are first indicated as being present or not, with a follow-up tie interpretation interface for each tie indicated within the sociogram, to (B) where ties are elicited on the sociogram in a stepped manner. The choice of the best approach to utilize must be directed by a study team's careful consideration of the context of the study, such as the data expected to be generated and the participant population that will be engaged. For example, follow-up tie elicitation within (A) may become tedious for studies which anticipate a large number of alter-alter ties, while ties elicited on the sociogram (B) might be more difficult for study populations with low dexterity. The choice of approach also has implications for the structure of the data on the backend, with value of the tie as being saved as an attribute on a single edge within (A) representing closeness of connection, or with the value of tie closeness inferred by the presence or absence of three distinct edge types within (B), which are stored as either present or absent. Said more simply, the former approach produces a column of valued data, perhaps ranging from 1 to 3; the latter produces perhaps three columns of binary data, all containing either a 1 or a 0.

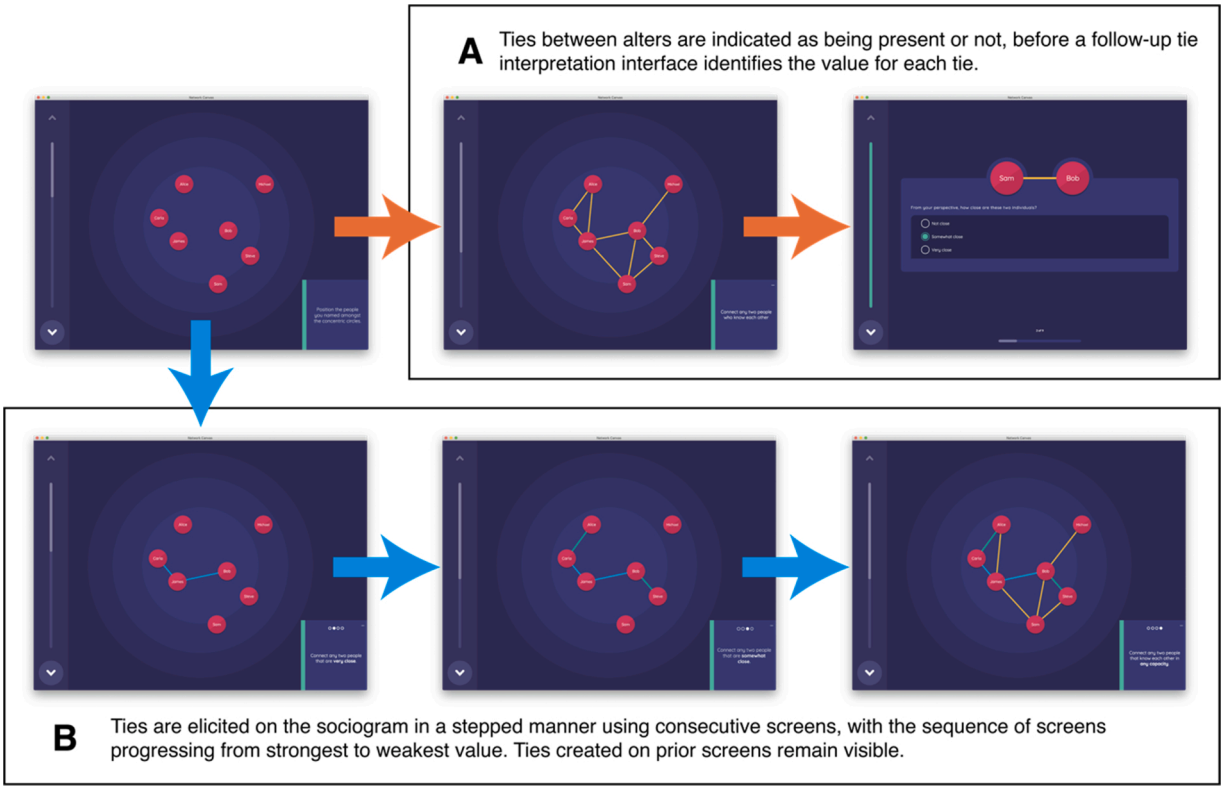

Fig. 3. Two example workflows within Network Canvas for capturing valued alter-alter ties.

Where network ontologies differed so fundamentally that a single interface would not suffice, we have endeavored to create dedicated interfaces for certain tasks. An example of this is the "narrative interface," which is designed specifically to facilitate discussion, storytelling, or reflection on the sociogram-as-object by the participant. It provides the facility for free annotation of the sociogram (including support for pen or stylus input), and allows the researcher to create collections of node and edge attribute data called "presets" that visually re-code the sociogram around a specific theme. Another example of this configurability is that we specifically designed the software to accommodate any visual stimulus for arranging one's network and eliciting alter-to-alter ties, as researchers have their own different approaches, each with unique strengths and weaknesses (Hollstein et al., 2020).

To further maximize the reach of the software to the broadest community of researchers, we have prioritized ensuring that our software is compatible on a wide variety of devices and operating systems. While we recommend that the Interviewer App is run on touch-enabled field devices – such as iPads, Surface laptops, or desktops with touch-enabled displays – in order to take advantage of survey touch capabilities, this is not a requirement.

The final point which demonstrates our concern with efficiency, ease, and access – we have constructed Network Canvas protocol files to be easily shareable. We see that this decision may enable a mechanism that the social network community has often lacked: an efficient technical means to facilitate replication studies. Once constructed, Network Canvas protocol files (.netcanvas files) are easily shared. Any external data, images, text, and code are all bundled as a single. netcanvas file that can be hosted on a server, sent through email, or otherwise transferred between colleagues. This means that researchers can easily modify or directly implement existing research protocols.

Designing a free and sustainable software

One additional key challenge facing not just our tool – but all academic software – is sustainability (Howison and Herbsleb, 2011). Because of this, we have prioritized conducting open and collaborative development as we strongly believe that only through the buy-in of the research community will the project be sustainable.

Key to the "open" element of this strategy is that our development happens in full public view using the GitHub social coding platform, and is entirely open source. GitHub allows full access not only to our source-code and precompiled binaries, but also to our issue tracking system, our project management boards, and our developer documentation. This gives any interested party a clearly articulated path to submitting bug reports or feature requests, interrogating our decision-making, or simply learning more about the project. The open-source nature of our code enables external researchers to review and audit any implementations that might make a difference to their work, such as key security and privacy functionality.

Crucially, our use of GitHub also enables a key element of our collaborative model, since researchers can directly contribute additional code that we can integrate into the main codebase. This collaborative development is vital to the long-term sustainability of the software, as it will help ensure our work remains relevant. It is conceivable that once funded development of Network Canvas concludes, the research community itself will be able to maintain the technical aspects of the project, in the spirit of existing open-source community run projects such as Linux, Apache, and Firefox. This would represent a pronounced democratization of the development of research tooling, when compared with prior closed source "expert led" approaches.

The second element of the collaborative ethos we undertook was to engage with varied research communities, and directly incorporate feedback and functionality requests into our development roadmap. This strategy had three key components:

(1) *The establishment of "test sites" within key research communities*. These test sites were chosen because of their expertise in network data collection and health, as well as their willingness to "road test" alpha and beta versions of the software by attempting to implement planned research and highlight inflexibilities and limitations with our approach. For example, one collaboration revealed the importance of designing several styles of roster interface, in particular one which accommodated exceptionally large rosters of several thousand nodes. To meet this need, we built an interface that allows researchers to import large rosters into their surveys. On this interface, researchers can customize the display and search properties of the imported dataset within which participants can search, view, and select nodes. In another example, collaborators advocated for the development of an entity resolution module, a feature that matches and consolidates nodes named across multiple ego networks using machine learning to reduce the burden of reviewing potential matches, which led us to pursue funding to develop and integrate that feature within the Server App (Janulis, 2017).
(2) *Discovery activities including workshops, surveys, focus groups, and the convening of a Scientific Advisory Board*. We used these mechanisms to canvas opinion on development priorities, specific functionality relevant to specialized research communities, and general project direction.
(3) *The production of high-quality promotional materials and documentation*. In conjunction with our workshops, extensive documentation and promotional materials are designed to raise the profile of the project, as well as to enable students and researchers outside of academia to learn to use the software at no cost, thereby growing the community.

The final manifestation of this decision can be found in our licensing model. In order to enshrine the principles of openness and collaboration, and to provide a guarantee to collaborators and the community that their effort will not be privatized and sold back to them, the Network Canvas Software Suite is, and will always be, free. Our definition of "free software" is taken from the Free Software Foundation's General Public License (GPL) (GNU General Public License, 2007), which is the license through which we distribute. This is a popular open source license that provides anyone who wants to use our software with certain guaranteed freedoms, including (non-exhaustively):

- Freedom to use the software commercially, without restriction;
- Freedom to modify the software however they see fit; and
- Freedom to distribute the software.

These freedoms are protected by some requirements built into the GPL license:

- Any modifications must be made available under the GPL license and
- You must include a copy of the original source code with any versions you distribute, making it possible for others to modify.

By locking our code to this licensing model, we hope to create a sense of transparency, fairness, and trust. We do not believe it would be palatable to ask for feedback or contributions from others within this community to improve the software, if there were a risk that these improvements could be hidden behind fees, black-box code, or prohibitive licensing terms

Limitations of our design approach

Every software design decision both enables and constrains particular research applications, and researchers must think carefully when selecting tools based on their goals. Next, we outline several of the limitations of our work in order to assist readers in determining the utility of our tool for their circumstances.

Our first limitation is that, to date, we have only limited evidence of

the validity of our software – with all evidence coming from two studies (Hogan et al., 2016, 2019). Both of these studies provide initial demonstrations of the high quality of data provided by our participants, through either empirical examinations of construct validity or data reliability over time (Hogan et al., 2016, 2019). However, we have not explicitly evaluated our approach against other digital tools, nor examined how choice of specific Network Canvas interfaces or their customization impacts the underlying data (e.g., participant-aided sociogram versus dyad census), as others have done with related digital tools (Eddens and Fagan, 2018). Indeed, as demonstrated by Hollstein et al. (2020), even small changes in the look of a visual data collection tool can lead to substantial differences in the resulting data. For example, in one small study we found meaningful differences between the number of alter-alter ties collected via Network Canvas compared to a dyadic census (Janulis et al., 2019). Furthermore, while we have found initial evidence of the usability and efficiency of our interface design for research participants (Hogan et al., 2016), thus far we lack any empirical evidence about the usability and efficiency of our tool for researchers. To further refine our design and provide researchers with greater guidance of best practices in the use of our tool, our group plans to expand our evaluation into all these aspects of our tool.

A second limitation to our work is that we intentionally designed our interview to be interviewer-assisted, or that interviews are run from the Network Canvas Interviewer App which is installed on an interviewer-controlled field device. As described earlier in this manuscript, this decision was made to mitigate a number of design, data quality, and security issues which would have resulted from interviews being run on a participant's personal device, and that we believed would have insurmountably sacrificed our ability to ensure useable interfaces and secure data transfer. A consequence of this decision though is that our tool does not currently support remote surveys. Indeed, from our discovery events with the social network and public health communities, many researchers rely on remote interviews that occur on cell phones, or through email links to web surveys. And for those individuals, other software may be a better fit. That said, our team is actively exploring other software models, such as the addition of a separate component to our base software which would allow the administration of remote interviews. A remote option which fully or partially removes the interviewer from the interview may increase interview convenience and improve the ability for data to be captured anonymously – both of which might improve the data quality on sensitive measures or the response rates of hidden, hard-to-reach, or marginalized populations. However, a remote option may perhaps be most ideal for researchers less concerned with the transmission of sensitive identifiable data, and those who do not believe participant buy-in or usability will be lessened by the removal of the interviewer from the interview.

Of course, having interviews be interviewer-assisted introduces other potential limitations – such as the potential to influence respondent disclosure. To mitigate this we suggest that researchers, particularly when capturing sensitive data, think critically about opportunities to build trust with research participants. For example, interviewer training on the building of rapport may be fundamentally important. Researchers may also wish to structure interviews so that sensitive data are captured during periods when the interview device is handed off to the respondent so as to allow greater privacy.

While we have attempted to build a tool that is usable across broad array of populations, including vulnerable and disabled populations, Network Canvas may not be a strong fit for all participants. For example, the visual elements and touch that is emphasized within our design may be difficult to navigate for older populations, those with visual impairments, and those who may have less technological literacy. Through our future work we will continue to examine the usability and acceptability of our tool in new populations – for example, we recently began a project examining the acceptability of our tool in a population of Nigerian men at risk for HIV. However, we also encourage researchers to think carefully about how our tool may fit with the accessibility needs of their participants.

While we sought to ensure that our tool was useable for multiple research communities, we have found that the flexibility of Network Canvas can a double-edged sword. Our experiences with early users have found that those who have relied on bespoke methods are often reluctant to deviate from their known style of interview. And this has led some users to initially build overly tedious interview protocols which do not take full advantage of the efficiencies of a digital interface. For example, many initial protocols will often replicate long screens of per alter Likert-style items. While there are instances when researchers should closely replicate prior interview protocols (i.e., for data comparability), we suspect that the flexibility of the tool causes a high learning curve for researchers. It is not enough to build a tool, as we have learned that we also must teach users how to think like Network Canvas before the community begins to take full advantage of the design features of Network Canvas. For example, many researchers are used to carefully determining item wording when designing a measure, but building a Network Canvas protocol also requires that researchers consider the overall flow of the protocol and the look and design of specific interfaces. These, along with consideration of the needs of the participant population, will help ensure that researchers utilize Network Canvas in ways which maximize ease and efficiency for research participants.

While at first our training materials shied away from providing too many examples, as we wanted users to experiment and be creative within their research protocols – we have actually found that concrete demonstrations of the software's capacities are vital in providing researchers foundational understanding that will lead to them taking full advantage of the software. As our team becomes more adept at training users, we are keen to learn alongside them as they discover creative applications of Network Canvas within their own research domains.

We have also found that having a large and interdisciplinary development team, and one which values input from the community, is both a strength and a limitation. Our team spans fields such as psychology, sociology, epidemiology, public health and computer science. This diversity has provided our group with a broader understanding of the needs of the field, and has likely enabled us to build a more generalizable tool than already existed. However, collaborative work is slow. While we believe that we end with an overall better product, this work takes time and open and consistent communication about individual and team priorities, so that we are able to converge upon a common vision. It also introduces complexities in licensing and governance which we have attempted to mitigate by being as open, transparent, and democratic in our governance, as possible. Finally, seeking and incorporating the user feedback takes a great deal of time. For example, while our group built in numerous opportunities for interaction with test sites and users, we have found these activities to be most meaningful when a researcher's personal data needs are in alignment with our specific feedback needs. For example, some of the most insightful feedback has come from students and early career researchers who are just launching their research programs. All of this work takes time. Our software entered Beta in 2019 and remained there until the end of 2020. An unfortunate but necessary frustration for researchers who may have been waiting for our work.

Finally, producing a robust and flexible software for the community takes a great deal of resources. We have been lucky to have been supported by the NIH and we hope to continue in this model for many years. While soft money is never certain, we hope that by developing the software openly, we will attract a community that will help sustain Network Canvas over the years.

We do not assert that our tool will be best for all circumstances. As pointed out above, the Interviewer App has been designed for in-person interviewer-assisted interviews. Researchers who are constrained to remote data capture may find our tool less useful. Furthermore, researchers of patient populations with low technical savvy or those with visual or physical impairments may have greater difficulty navigating interview screens. And perhaps most importantly, researchers who have uncomplicated network data needs may be better served by lower

technology approaches or by embedding their network questions into already familiar research tools. For example, for studies where the majority of the data collection is at the individual- vs the network-level, embedding a single name generator within a REDCap survey might suffice for some research studies.

Conclusion

As outlined in this paper, our team has attempted to build a sustainable software suite which is easy, flexible and efficient for researchers and empowering and intelligible for participants. As outlined throughout this paper, we believe that our software tool will relieve several bottlenecks which have substantially impacted the ability to capture data, as well as hindered progress within the field of social networks.

For the next few years our team is focused on refining and evaluating the core software, training users, and extending the Software Suite. Of note, a high priority for us is the development of new interfaces for geospatial and temporal data, the construction of an online library for the sharing of Network Canvas protocols, and exploring ways of expanding node elicitation to cover different measurement intervals with even lower burden. Finally, we are interested in expanding our Software Suite so that it is not just a tool for researchers, but a tool for public health professionals to conduct real-time disease investigation, including the ability to meaningfully analyze survey data collected with Network Canvas from a central portal.

By attempting to make it easier and less resource intensive to capture data, we hope researchers from diverse domains reconsider the importance and value of capturing network data directly from participants versus utilizing publicly available (e.g., GSS, Add Health) or trace network data (e.g., Twitter), as has been said by numerous others – these are not interchangeable and reflect different aspects of human social behavior (Perry et al., 2018; Adams, 2019; Mc Carty et al., 2019). Furthermore, by making it easy to share and deploy specific network research protocols – we hope our tool furthers reproducibility in the field of social networks and enables researchers to have a more nuanced understanding of network data capture methodology and protocol design, and the reliability of methods across various populations and settings.

As data collection tools both shape and constrain research in fundamental ways, we hope by explicating the underlying design principles and process by which we came to those principles, we clarify the strengths and limitations of the Network Canvas Software Suite to best inform researchers in their efforts to capture structural data.

Ackowledgements

This study was supported by the National Institutes of Health: R01DA042711, MPIs: Birkett & Phillips; K08DA037825, PI: Birkett; R21LM012578, PI: Janulis.

References

Adams, J., 2019. Gathering Social Network Data, 1 ed. SAGE Publications, Thousand Oaks.
ANAMIA Egocenter, 2013. Ecole des Hautes Etudes en Sciences Sociales (EHESS) [computer program]. Centre Edgar-Morin (CEM-IIAC), Paris.
Antonucci, T., 1986. Hierarchical mapping technique. Generations: J. Am. Soc. Aging 10, 10–12.
Baum, F., MacDougall, C., Smith, D., 2006. Participatory action research. J. Epidemiol. Community Health 60 (10), 854–857.
Carrasco, J.A., Hogan, B., Wellman, B., Miller, E.J., 2008. Collecting social network data to study social activity-travel behavior: an egocentric approach. Environ. Plann. B Plann. Des. 35 (6), 961–980.
codaco. Network Canvas. https://github.com/codaco/Network-Canvas.
Crossley, N., Bellotti, E., Edwards, G., Everett, M., Koskinen, J., Tranmer, M., 2015. Social Network Analysis for Ego-nets.
Dubois, E., Ford, H., 2015. Qualitative political communication| trace interviews: an actor-centered approach. Int. J. Commun. 9, 2067–2091.
Eddens, K., Fagan, J.M., 2018. Comparing nascent approaches for gathering alter-tie data for egocentric studies. Soc. Netw. 55, 130–141.
Edwards, G., 2010. Mixed-method Approaches to Social Network Analysis. National Center for Research Methods.
Mc Carty C. Egonet. https://github.com/egonet/egonet.
Eddens K. Enso. http://iuni.iu.edu/projects/enso.
Freeman, L.C., 2000. Visualizing social networks. J. Soc. Struct. 1 (1).
Gauntlett, D., 2014. The LEGO system as a tool for thinking, creativity, and changing the world. In: Wolf, M.J.P. (Ed.), LEGO Studies: Examining the Building Blocks of a Transmedial Phenomenon. Routledge, New York.
GNU General Public License, 2007. GNU Operating System. https://www.gnu.org/licenses/gpl-3.0.en.html.
Guggenheim, M., Kraeftner, B., Kroell, J., 2013. 'I don't know whether I need a further level of disaster': shifting media of sociology in the sandbox. Distinktion Scand. J. Soc. Theory 14 (3), 284–304.
Heer, J., Shneiderman, B., 2012. Interactive dynamics for visual analytics. Commun. ACM 10 (2), 45–54.
Herz, A., Petermann, S., 2017. Beyond interviewer effects in the standardized measurement of ego-centric networks. Soc. Netw. 50, 70–82.
Hlebec, V., Ferligoj, A., 2002. Reliability of social network measurement instruments. Soc. Networks 14 (3), 288–306.
Hogan, B., Carrasco, J.A., Wellman, B., 2007. Visualizing personal networks: working with participant-aided sociograms. Field methods 19 (2), 116–144.
Hogan, B., Melville, J.R., Philips 2nd, G.L., et al., 2016. Evaluating the paper-to-Screen translation of participant-aided sociograms with high-risk participants. Proc. SIGCHI Conf. Hum. Factor. Comput. Syst. 2016, 5360–5371.
Hogan, B., Simonetto, P., Archambault, D., et al., 2017. Drawing sociograms is not quite like drawing graphs: evaluating the use of touchscreen participant-aided sociograms. In: European Conference of Social Networks (EUSN). Mainz, Germany.
Hogan, B., Janulis, P., Phillips 2nd, G.L., et al., 2019. Assessing the stability of egocentric networks over time using the digital participant-aided sociogram tool Network Canvas. Netw. Sci. 1–19.
Hollstein, B., Topfer, T., Pfeffer, J., 2020. Collecting egocentric network data with visual tools: a comparative study. Netw. Sci. 8 (2), 223–250. https://doi.org/10.1017/nws.2020.4.
Howison, J., Herbsleb, J.D., 2011. Scientific software production: incentives and collaboration. In: Paper Presented at: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work. Hangzhou, China.
Janulis, P., 2017. Implementing and Evaluating a Machine Learning Tool for Entity Resolution in Drug Use and Sexual Contact Networks of YMSM. NLM/National Institute of Health.
Janulis, P., Philips 2nd, G.L., Neray, B., Hogan, B., Melville, J.R., Birkett, M., 2019. An experimental comparison of alter-to alter tie elicitation: participant-aided sociogram versus traditional dyadic census. In: International Sunbelt Social Network Conference. Montreal, Canada.
Kennedy D. EgoWeb 2.0. https://github.com/qualintitative/egoweb.
Kuhns, L.M., Birkett, M., Muth, S.Q., et al., 2015. Methods for collection of participant-aided sociograms for the study of social, sexual and substance-using networks among young men who have sex with men. Connect (Tor) 35 (1).
Macaulay, C., Sloan, D., Jiang, X., et al., 2009. Scientific software development projects that integrate usability and user- centered design methods can have an impact that goes beyond a limited set of users. IEEE Softw. 26 (1), 92–102.
Marsden, P., 2011. Survey Methods of Network Data. Sage Publications, London.
Mc Carty, C., Govindaramanujam, S., 2005. A modified elicitation of personal networks using dynamic visualization. Connections 26 (2), 61-19.
Mc Carty, C., Killworth, P.D., Rennell, J., 2007. Impact of methods for reducing respondent burden on personal network structural measures. Soc. Netw. 29 (2), 300–315.
Mc Carty, C., Lubbers, M., Vacca, R., Molina, J.L., 2019. Conducting Personal Network Research: A. Practical Guide. The Guilford Press.
McGrenere, J., Baecker, R., Booth, K., 2002. An evaluation of a multiple interface design solution for bloated software. Paper Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
Network Canvas (2021) https://www.networkcanvas.com/.
Norman, D., 2007. The next UI breakthrough, part 2: physicality. Interactions 14 (4), 46–47.
Pahl, R., Spencer, L., 2004. Personal communities: not simply families of 'fate' or 'choice'. Curr. Sociol. 52 (2), 199–221.
Perry, B.L., Pescosolido, B.A., Borgatti, S.P., 2018. Egocentric Network Analysis: Foundations, Methods, and Models. Cambridge University Press, Cambridge, UK ; New York, NY.
Ryan, L., Mulholland, J., Agoston, A., 2014. Talking ties: reflecting on network visualisation and qualitative interviewing. Sociol. Res. Online 19 (2), 1–12.
Silber, H., Schroder, J., Struminskaya, B., Stocke, V., Bosnjak, M., 2019. Does panel conditioning affect data quality in ego-centered social network questions? Soc. Netw. 56, 45–54.
Stark T. GENSI. http://www.tobiasstark.nl/GENSI/GENSI.htm.
Stark, T.H., Krosnick, J.A., 2017. GENSI: a new graphical tool to collect ego-centered network data. Soc. Netw. 48, 36–45.
Tourangeau, R., Rips, L.J., Rasinski, K.A., 2000. The Psychology of Survey Response. Cambridge University Press, Cambridge, U.K. ; New York.
Trellis (2020) https://trellis.yale.edu/.
Valente, T.W., Pitts, S.R., 2017. An appraisal of social network theory and analysis as applied to public health: challenges and opportunities. Annu. Rev. Public Health 38, 103–118.

Vehovar, V., Manfreda, K.L., Koren, G., Hlebec, V., 2008. Measuring ego-centered social networks on the web: Questionnaire design issues. Soc. Netw. 30 (3), 213–222.
Vennmaker. https://github.com/VennMaker/vennmaker-source.
Yousefi-Nooraie, R., Marin, A., Hanneman, R., Pullenayegum, E., Lohfeld, L., Dobbins, M., 2017. The relationship between the position of name generator questions and responsiveness in multiple name generator surveys. Sociol. Methods Res. 48 (2), 243–262.