Question-order effects in social network name generators

James E. Pustejovsky and James P. Spillane
Northwestern University

"Although we sympathize with the characterization of context effects as artifacts, we argue that the processes that result in context effects are interesting substantive phenomena in their own right."

(Tourangeau and Rasinski, 1988, p. 301)

Why study multi-dimensional social networks?

Little attention has been paid to measurement of multi-dimensional social networks and the validity of dimensional comparisons, despite the fact that many research questions require attention to several types of relationships among a given set of people:

- Lazega and Pattison (1999) collect data on co-work, advice, and friendship to study patterns in the relational structure of a law firm spread across multiple offices.
- Cross, Borgatti, and Parker (2001) collect data on five specific motives for advice-seeking behavior in order to study the dimensionality of this specific relationship.
- Bernard, et al. (1990) studies differences between emotional support networks and social support networks.
- Plickert, et al. (2007) studies factors affecting reciprocation rates between people that provide others with emotional support, with minor services, and with major services.
- Ruan (1998) compares the approach of the single network question in the General Social Survey to an exchange-based approach that makes use of eleven name generators, noting that the conclusions are susceptible to question-order effects.

Several different designs can be used to collect network data, among them roster-based methods and recall-based methods. In this paper, we focus on measuring multi-dimensional social networks using recall-based name generators and interpreters.

Five theories of question effects for social network name generators

Following Straits (2000), we draw on studies of question-order effects in behavioral surveys and attitude surveys to identify mechanisms that might be relevant to multiple name generator survey designs. In outlining these effects, we imagine a survey with two name generators.

- ***Fatigue***. Effects would be observed due to fatigue if, in response to the second name generator, a respondent names fewer alters than she otherwise would have, had the first network prompt not been asked. In the extreme, fatigue effects might produce non-response to later name generators. Tourangeau and Rasinski (1988) have proposed that fatigue effects are particularly pronounced in surveys where the overall length depends on the number of items named.
- ***Satisficing***. Satisficing effects occur when a respondent gives an answer that she believes satisfies the request for information, but is not a complete, optimally considered response (Krosnick, 2000). Satisficing is thought to be regulated by task difficulty, respondent ability, and respondent motivation (Krosnick, et al.,1996). In the context of name generator prompts, satisficing would play a role when a respondent decides how many alters to name in response to the second prompt. When confronted with the second name generator prompt, a respondent who does not want to give a full answer could turn to the precedent that she herself set by responding to the first name generator.
- ***Non-redundancy***. In the context of multiple name generator surveys, non-redundancy effects would appear if a respondent omits the names of certain alters in the second name generator because she has already listed the alters in the first name generator (Straits, 2000). The respondent might interpret the second name generator prompt as beginning with the qualification, "Aside from the people you have already named…"
- ***Cognitive priming***. The process of retrieving names from memory for the first generator may start a sub-conscious activation process that brings certain names to the forefront for subsequent name generator questions. If not for the priming effect of the earlier name generator, a respondent might not list certain alters. Evidence from a number of studies has suggested that the social proximity of acquaintances plays a role in how a respondent recalls them from memory when answering a single network question (Brewer, et al., 2005), though it remains to be seen whether the same mechanism might apply across multiple name generators, and name generators that use criterion relationships other than acquaintance.
- ***Question scope redefinition***. Question-order effects may appear because survey respondents use the wording of specific questions, the sequencing of questions, and other facets of the instrument to infer the pragmatic meaning of a question (Schwarz, 1999). Social network name generators are no exception; a respondent must make some assumptions about the sort of names that the question is intended to produce, and will look for contextual clues in order to understand the relationship being described. If a respondent relies on contextual clues from the first name generator to understand the pragmatic meaning of the second generator, the alters that she names may be different from those she would have named in the absence of the first generator.

Instrument design, data collection, and analysis

Our own research project makes use of a multiple name generators and name interpreters to study social capital in elementary schools. We hypothesize that teachers' thought processes differ by subject area, even for the same teachers. Hence, to study advice networks in schools it is problematic to simply ask about teaching in general and necessary instead to study advice networks in particular curricular domains (e.g., mathematics, language arts, etc.). Further, we hypothesize that both advice from their colleagues and advice from people outside of teachers' school buildings play roles in shaping their practice.

In designing a network survey to study social capital in elementary schools, we sought a way to investigate both our substantive questions and one of the methodological issues arising in the design of multiple name-generator surveys. We use a split-ballot experimental design to test whether the order of name-generator prompts in the survey affects the resultant data.

The flow of questions in the name generator section of the survey is tailored in two ways: the respondent's self-described role, and the subject(s) that she teaches. Figure 1 depicts the randomized design of the name generators. Figure 2, panel B shows the appearance of one name generator. Each name generator begins with the same wording: "In the past year, to whom have you gone for advice or information about [SUBJECT PROMPT]?"

Each name generator is followed by a series of name-interpreter questions. For each alter name listed, data is collected on the job or role of the alter, the content of the advice interactions, the frequency of interactions between respondent and alter, and the respondent's rating of the influence of the alter's advice on her practice. Figure 2, panels C and D depict the design of the name interpreter questions.

In the analysis that follows, we use two administrations of our redesigned instrument. One administration involved a sample of 15 public elementary schools and 4 Catholic elementary schools (mostly serving kindergarten through 8th grade) in the same city. The second administration consisted of 10 public middle-schools in a mid-sized city in a different state, all serving grades 6 through 8. Table 1 reports response rates from each sample.

The survey was designed to randomize the order of the math and RWLA prompts only in situations where the respondent either taught both subjects or taught neither. The analysis that follows is limited to those respondents; it therefore excludes teachers and specialists who teach only math, teach only RWLA, or primarily teach other subject classes, but also teach math or RWLA (but not both).

Figure 1. Randomized design of name generators, based on self-described role and subject area(s)

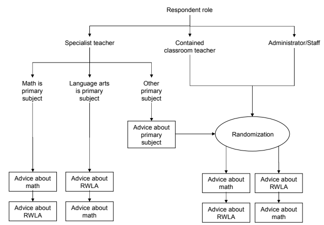
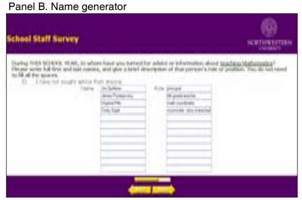


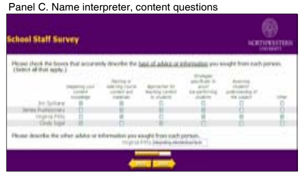
Figure 2. Screen-captures from the sequence of questions in one name generator and name interpreter

Panel A. Instructions

Panel B. Name generator



Panel C. Name interpreter, content questions



Panel D. Name interpreter, frequency and influence questions

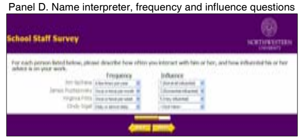


Table 1. Sample sizes and response rates

	19 Elementary Schools	10 Middle Schools
Faculty size	14 to 69	49 to 69
Total staff	544	634
Responses	414 (76%)	548 (87%)
Randomized responses	264	323
Composition	Predominantly contained-classroom primary-grade teachers	Predominantly subject-area teachers, some 6th grade teachers in contained classrooms

Findings

Table 2 presents the average out-degree for each subject-area name generator by treatment. The RWLA name generator experiment reveals significant differences in average out-degree appear in both administrations; respondents in the M/R treatment list about one name fewer than respondents in the R/M treatment. This difference is large in magnitude—on the order of a 40% decrease. The math name generator experiment does not reveal a clear pattern. Note that if the purpose of our research were only to determine whether teachers sought more advice about RWLA or about math, we would reach opposite conclusions if we looked only at the R/M treatment or the M/R treatment.

Table 2. Average out-degree

		Treatment		
Administration	Statistic	R/M	M/R	Difference
Elementary Schools	Respondents	126	138	-12
	RWLA out-degree	3.21	1.91	1.29*
	Math out-degree	1.46	1.70	-0.24
Middle Schools	Respondents	159	164	-5
	RWLA out-degree	2.35	1.47	0.88*
	Math out-degree	1.64	1.32	0.32

* Difference is significant at the 5% level according to a Mann-Whitney test.

Fatigue does not appear to be the primary reason for the observed pattern of differences in out-degree. If fatigue effects were the only cause of the observed differences in mean out-degree, we would expect the distribution of the total number of alters named in both name generators to be similar across treatment groups, because there is no reason for the two randomly-assigned treatment groups to differ in the amount of effort they are willing to exert.

Satisficing could create the observed pattern of average out-degrees by acting on actual differences in average network size between the RWLA advice network and the math network. If the true size of the second network is larger than the reported size of the first network, the satisficing respondent will list only as many names as she did in response to the first name generator, because such a response seems sufficiently complete. We observe that only 33% of respondents in the elementary school M/R treatment group listed more names in the RWLA generator than in the math generator, compared to 68% of R/M respondents. A similar pattern is observed in the middle school sample, though not as large in magnitude.

Table 3. Average number of content-areas checked per alter

		Treatment		
Administration	Name interpreter	R/M	M/R	Difference
Elementary Schools	RWLA	3.0	2.8	0.2
	Math	3.3	2.6	0.8*
Middle Schools	RWLA	2.6	2.5	0.1
	Math	2.6	2.3	0.3*

* Difference is significant at the 5% level according to a Mann-Whitney test.

Non-redundancy effects and cognitive priming effects are not presented, because the randomized design does not allow estimation of the true amount of multiplexity in the two networks. We simply observe that 32% of all alters were named by the same respondent in both generators; in the middle school sample, only 17% of alters were named by the same respondent in both generators.

Question-scope redefinition is observed in two senses. First, content areas are checked more often in the second name generator than in the first. This is true of nearly every specific category. In contrast, the 6th "other" category is checked less frequently in the second interpreter. Second, as table 3 reports, differences between treatment groups in the number of content-areas checked are observed for the math network questions, but not the RWLA network questions. We believe that these differences may be the result of how respondents understand the nature of advice about the different subject areas.

Acknowledgments and contact

Work on this project is supported by the Distributed Leadership Studies (http://www.distributedleadership.org), funded by a research grant from the National Science Foundation (RETA Grant EHR-0412510), with support from Northwestern University's School of Education and Social Policy and Institute for Policy Research. We thank Michael Lach, Gina Grant, Bill Suttles, Ruth Heaton, and Jim Lewis for their assistance with data collection. We also gratefully acknowledge our research team for their continuous support, especially Cindy Sigal, Bijou Hunt, and Virginia Pitts. All opinions and conclusions expressed in this paper are our own and do not reflect the views of any funding agency. Please direct correspondence to James Pustejovsky at Northwestern University, School of Education and Social Policy, 2020 North Campus Drive, Evanston, IL 60201 or j-pustejovsky@northwestern.edu.

Limitations, further research

The effects for which we find evidence are troubling, because they are so closely related to the substantive questions that provoked our research. In particular, satisficing effects and question-scope redefinition effects create biases that would cause us to reach opposite conclusions depending on the order in which the name generators and interpreters were posed.

Our conclusions are limited in several ways, some of which are suggestive of directions for further research:

- Our instrument design does not permit a true test for non-redundancy effects. Without being able to estimate the true level of multiplexity, our conclusions about cognitive priming should also be regarded as tentative.
- We present separate tests for various types of question-order effects. In order to isolate the relative contribution of each effect, an integrated model would be necessary.
- The validity and accuracy of any measurement depends in part on the particular statistic or metric that is applied to the raw data (Costenbader and Valente, 2003). We have focused only on very basic aspects of network data; whether more complex metrics such as closeness or betweenness could be affected by question-order remains to be seen.
- Our results may not generalize to measurement of different criterion relationships. Further methodological research is needed to examine question-order effects using other prevalent criterion relationships.
- One should also be wary of generalizing to other organizational settings. The cognitivist approach to studying name generator accuracy suggests that the recall accuracy depends on the degree to which respondents have a well-developed structure for storing memories of other people (Butts, 2003; Freeman et al., 1987). Biases created by question-order effects may be lessened to the extent that name generators specify criterion relationships for which respondents have good mental models. This question deserves much further investigation.

Design improvements

We conclude with some general suggestions regarding instrument design for multiple name-generator surveys.

- Use complete roster methods whenever possible (Brewer, 2000).
- If possible, randomize the order of name generators.
- Pay attention to the order in which name interpreters are asked. Kogovšek, et al. (2002) have suggested that asking all name interpreter questions about each respondent in turn provides superior validity to asking each name interpreter question in turn about all respondents.
- In situations where roster-based methods are not feasible, we recommend that the instrument be designed to separate the name generator questions from the name interpreter questions:
  - Run all name generators first, and if possible prompt the respondent to keep searching her memory. Techniques such as non-specific prompting, multiple elicitations, or re-interviewing have been recommended to collect robust name generator data (Brewer, 2000).
  - Once a set of alter names has been generated, pose name interpreter questions that asks the respondent to classify the alter into one or more of the criterion relationships of interest. A similar approach has been applied in surveys that collect egocentric network data (see for instance Marin, 2004; Brewer, 2000 also cites a survey by L.M. Jones and C.S. Fischer that apparently uses a similar design).

Literature cited

Bernard, H.R., Johnsen, E.C., Killworth, P.D., McCarty, C., Shelley, G.A., Robinson, S., 1990. Comparing four different methods for measuring personal social networks. Social Networks 12, 179-215.
Brewer, D.D., 2000. Forgetting in the recall-based elicitation of personal and social networks. Social Networks 22, 29-43.
Brewer, D.D., Rinaldi, G., Mogoutov, A., Valente, T.W., 2005. A Quantitative Review of Associative Patterns in the Recall of Persons. Journal of Social Structure 6.
Butts, C.T., 2003. Network inference, error, and informant (in)accuracy: a Bayesian approach. Social Networks 25, 103-140.
Costenbader, E., Valente, T.W., 2003. The stability of centrality measures when networks are sampled. Social Networks 25, 283–307.
Cross, R., Borgatti, S.P., Parker, A., 2001. Beyond answers: dimensions of the advice network. Social Networks 23, 215-235.
Freeman, L.C., Romney, A.K., Freeman, S.C., 1987. Cognitive structure and informant accuracy. American Anthropologist, New Series 89 (2), 310-325.
Kogovšek, T., Ferligoj, A., Coenders, G., Saris, W., 2002. Estimating the reliability and validity of personal support measures: full information ML estimation with planned incomplete data. Social Networks 24, 1-20.
Krosnick, J.A., 2000. The threat of satisficing in surveys: the shortcuts respondents take in answering questions. Survey Methods Centre Newsletter 20, 4-8.
Krosnick, J.A., Narayan, S.S., Smith, W.R., 1996. Satisficing in surveys: initial evidence. In Braverman, M.T., Slater, J.K. (Eds.), Advances in survey research. San Francisco, Jossey-Bass, pp. 29-44.
Lazega, E., Pattison, P., 1999. Multiplexity, generalized exchange and cooperation in organizations: a case study. Social Networks 21, 67-90.
Marin, A., 2004. Are respondents more likely to list alters with certain characteristics? Implications for name generator data. Social Networks 26, 289–307.
Plickert, G., Côté, R.R., Wellman, B., 2007. It's not who you know, it's how you know them: who exchanges what with whom? Social Networks 29, 405-429.
Ruan, D., 1998. The content of the General Social Survey discussion networks: an exploration of General Social Survey discussion name generator in a Chinese context. Social Networks 20, 247-264.
Schwarz, N., 1999. Self-reports: how the questions shape the answers. American Psychologist 54 (2), 93-105.
Straits, B.C., 2000. Ego's important discussants or significant people: an experiment in varying the wording of personal network name generators. Social Networks 22, 123-140.
Tourangeau, R., Rasinski, K.A., 1988. Cognitive processes underlying context effects in attitude measurement. Psychological Bulletin 103 (3), 299-314.