

2014

# Identifying and Addressing Sources of Measurement Error in the Norbeck Social Support Questionnaire

Eileen Gigliotti  
*CUNY College of Staten Island*

William Ellery Samuels  
*CUNY College of Staten Island*

June Como  
*CUNY College of Staten Island*

Tofiga Gamidova  
*CUNY College of Staten Island*

Brian Gordon  
*CUNY College of Staten Island*

*See next page for additional authors*

**[How does access to this work benefit you? Let us know!](#)**

Follow this and additional works at: [http://academicworks.cuny.edu/si\\_pubs](http://academicworks.cuny.edu/si_pubs)

 Part of the [Nursing Commons](#), and the [Social and Behavioral Sciences Commons](#)

---

## Recommended Citation

Gigliotti, Eileen; Samuels, William Ellery; Como, June; Gamidova, Tofiga; Gordon, Brian; and Acerios, Peter, "Identifying and Addressing Sources of Measurement Error in the Norbeck Social Support Questionnaire" (2014). *CUNY Academic Works*.  
[http://academicworks.cuny.edu/si\\_pubs/65](http://academicworks.cuny.edu/si_pubs/65)

This Article is brought to you for free and open access by the College of Staten Island at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu).

---

**Authors**

Eileen Gigliotti, William Ellery Samuels, June Como, Tofiga Gamidova, Brian Gordon, and Peter Acerios

# Identifying and Addressing Sources of Measurement Error in the Norbeck Social Support Questionnaire

Eileen Gigliotti, PhD, RN  
William Ellery Samuels, PhD  
June Como, EdD, RN  
Tofiga Gamidova, RN, MS NP-BC  
Brian Gordon, RN, BS  
Peter Acerios, RN, BS

*College of Staten Island, The City University of New York*

**Background and Purpose:** The Norbeck Social Support Questionnaire (NSSQ) provides a rich portrait of one's perceived social support including network size/composition and quality of support. Analyses of quantitative research reports and critiques of the NSSQ revealed non-negligible measurement error. We document evidence of measurement error, report potential sources of this error, and present forthcoming psychometric testing. **Methods:** Quantitative evidence of measurement error from the NSSQ literature provided the basis for initial hypotheses concerning sources of error in network nomination and support ratings. We then conducted cognitive interviews to investigate these hypotheses. **Results:** Cognitive interviews revealed evidence of respondents' miscomprehension and response option bias. **Conclusions:** The current nomination process coupled with the lack of a "not applicable" response option and embedded examples in tangible Aid items reduces the accuracy of NSSQ subscores.

**Keywords:** Norbeck Social Support Questionnaire; measurement error; cognitive interviews; social support measurement

More than 30 years ago, Norbeck, Lindsey, and Carrieri (1981, 1983) began the process of developing the Norbeck Social Support Questionnaire (NSSQ), a widely used measure of one's perceived availability of affect, affirmation, and aid. However, as the term implies, *developing* a valid and reliable measure is an ongoing process. Since that time, they and others have been engaged in furthering its development through psychometric testing (Byers & Mullis, 1987; Gigliotti, 2002, 2006; Gigliotti & Samuels, 2011; Norbeck, 1985, 1995; Norbeck et al., 1983) as well as through analysis and critique (Gigliotti & Samuels, 2011; House, Kahn, McLeod, & Williams, 1985; Norbeck, 1995).

Benefitting from statistical advances, results of a confirmatory factor analysis (Gigliotti, 2002) showed that the NSSQ does indeed conform to its intended three-factor structure. However, these same results also found that the NSSQ contains meaningful

measurement error that can decrease power of a statistical test and result in failure to accept the alternative hypothesis when it is true (Type II errors). Our review of quantitative NSSQ data allowed us to formulate initial hypotheses concerning the primary sources of this error. We then conducted a series of cognitive interviews to qualitatively investigate these hypotheses and to identify further sources of measurement error.

The purpose of this article is threefold: First, we document the evidence (Gigliotti & Samuels, 2011; House et al., 1985; Norbeck, 1995) of measurement error in the NSSQ subscales and present our initial hypotheses. Second, we report methods and results of our cognitive interviews. Third, we present our revised hypotheses, proposed revisions to the NSSQ, and forthcoming psychometric testing.

## BACKGROUND

### Conceptual Framework

The NSSQ is based on Kahn's (1979) conceptualization of three social support functions: Affect, Affirmation, and Aid. Affect refers to positive regard; Affirmation is others' endorsement of one's behaviors, perceptions, and views; and Aid is the giving of material or symbolic help. The NSSQ measures each functional dimension by a two-item subscale (see Table 1). These three dimensions constitute the quality of support one perceives receiving from their network. The NSSQ also incorporates Kahn's "convoy" model, the evolving network one travels through life with, by measuring the size of one's network and the types of relationships (spouse, family, friends, etc.) that comprise it.

The NSSQ (Norbeck et al., 1981, 1983) stands out among measures of social support in two important ways. First, the NSSQ differentially catalogues both the quality and quantity of support one perceives. Like analyses of one's social network, the NSSQ quantifies the network's size (and composition). Unlike network analyses, the NSSQ also measures the quality of support by asking respondents to rate the level of Affect, Affirmation, and Aid

**TABLE 1. Norbeck Social Support Questionnaire Social Support Items**

Designation	Item
Affect1	How much does this person make you feel liked or loved?
Affect2	How much does this person make you feel respected or admired?
Affirm1	How much can you confide in this person?
Affirm2	How much does this person agree with your actions or thoughts?
Aid1	If you needed to borrow \$10, a ride to the doctor, or some other immediate help, how much could this person usually help?
Aid2	If you were confined to bed for several weeks, how much could this person help you?

From "Development of an instrument to measure social support," by J. S. Norbeck, A. M. Lindsey, and V. L. Carrieri, 1981, *Nursing Research*, 30, p. 265. Copyright 1981 by Lippincott Williams and Wilkins. Reprinted with permission.

one perceives receiving. As will be discussed further in the following text, however, the quality and quantity of support may be conflated in ways that can only be fixed through changes to the instrument.

Second, unlike global measures of support, the NSSQ combines measures of network size/composition with support ratings to give a detailed accounting of perceived availability of Affect, Affirmation, and Aid from each network relationship. Moving from general to specific hypotheses, NSSQ data allows one to investigate effects of (a) the entire network's total functional support (all three subscales combined), (b) the entire network's scores from each subscale, (c) total functional support from each relationship (e.g., spouse, family), and (d) each subscale's score from each relationship (e.g., spouse's affect). Thus, the NSSQ provides a uniquely rich portrait of one's perceived social support.

### Administration and Scoring

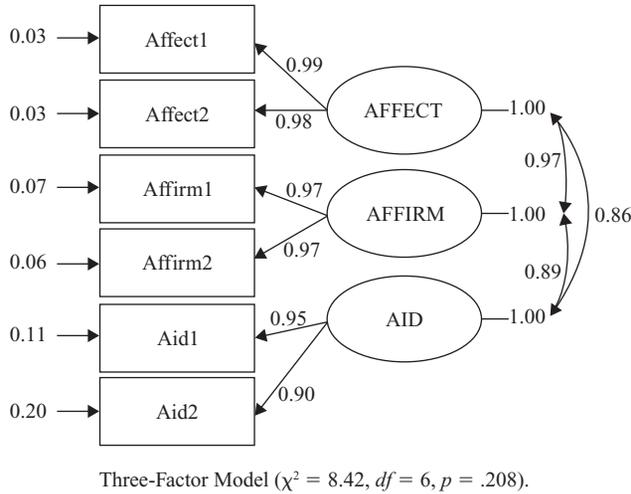
At present, the NSSQ is self-administered in paper and pencil format, in a two-stage process. In the first stage, respondents see 24 numbered spaces on the left side of the first page and are asked to "list each significant person in your life on the right." Consider all the persons who provide personal support for you or who are important to you." After listing up to 24 network members and indicating their relationship to the respondent, respondents are instructed to turn a half page and then begin the second stage. This stage entails rating each network member (0 = *none at all* to 4 = *a great deal*) on six items, two items per half page, regarding perceived levels of Affect, Affirmation, and Aid (see Table 1). Two items on the fourth half page inquire about duration of each relationship and frequency of contact. Finally, a full fifth page asks respondents about the loss of any relationships. Affect, Affirmation, and Aid subscale scores are then formed by summing respondents' ratings (0–4) of network members' perceived support.

## MEASUREMENT ERROR

Results of Gigliotti's (2002) confirmatory factor analysis are shown in Figure 1. On the far left of Figure 1 are the standardized residuals indicating the variance in those items not accounted for by the tested theoretical model, Norbeck's original three-factor model. These residuals reflect unexplained variance (error). With large residuals of .11 and .20, the Aid subscale is affected most by this error; the Affirmation subscale (.06–.07) is affected less, and the Affect subscale (.03) is least affected. However, despite the magnitude, all such error reduces the accuracy of subscores. As explained in the following text, evidence suggests that both network nomination (network size/composition) and subsequent support ratings were responsible for much of this measurement error.

### Network Nomination

**Network Size.** The NSSQ is uniquely designed to measure both the level of support and network size separately. These factors should be largely independent: A smaller, stronger network could provide similar total support to a larger, weaker network. The NSSQ measures size by asking respondents to nominate network members; it measures support by summing respondents' ratings (0–4) of up to 24 network members' support. A valid



**Figure 1.** Confirmatory Factor Analysis Results. From “A confirmation of the factor structure of the Norbeck Social Support Questionnaire (NSSQ),” by E. Gigliotti, 2002, *Nursing Research*, 51(5), p. 281. Copyright 2002 by Lippincott Williams and Wilkins. Reprinted with permission.

measure of support therefore depends on two things: First, the criteria by which network members are nominated is consistent among all respondents. Second, all nominated network members are accurately rated.

It appears that the NSSQ’s support measures do not currently create scores that are independent of network size. Support scores depend largely on the number of network members a respondent nominates and the reported (Gigliotti & Samuels, 2011) high, positive correlations between network size and support subscores reflect this: Affect,  $r_s = .94-.95$ ; Affirmation,  $.90-.92$ ; and Aid,  $.81-.82$ . Indeed, shortly after the NSSQ’s creation, House and colleagues (1985) cautioned that variations in network size were a primary source of error in NSSQ support scores.

However, although support scores are strongly affected by reported network size, allowing network size to vary is theoretically sound. Kahn’s (1979) model holds that network size relates to theoretically relevant factors such as extraversion–introversion (Pollet, Roberts, & Dunbar, 2011), and African Americans generally have smaller, more family-centered networks than do European Americans (Ajrouch, Antonucci, & Janevic, 2001). Moreover, family size is a major determinant of network size (Kahn & Antonucci, 1980), so members of cultures that advocate larger families will tend to have larger networks. Thus, it is important to retain the NSSQ’s measurement of network size and composition; it is also important to differentiate it from social support as much as they represent separate domains.

Nevertheless, if natural network size variations were the only source of extraneous variance for support scores, then averaging support scores (support score divided by network size) should remove the effect of network size, allowing the two domains to be measured separately. However, results of the first systematic investigation of the efficacy of averaged NSSQ subscores to decrease the role of network size as a source of measurement error in support scores—and thereby increase the power of a statistical test in three separate

samples (Gigliotti & Samuels, 2011) showed that only averaged Affect (mean increase in power = .47) and Affirmation (mean increase in power = .49) subscores resulted in consistent and substantial gains in statistical power compared to unaveraged subscores. In contrast, averaged Aid subscores improved more modestly in two samples (mean increase in power = .17), and power actually decreased by .19 in a third sample.

**Network Composition.** As previously noted, the NSSQ is based on Kahn's (1979) model of the social support network. Similar to other social network models (Dunbar & Spoor, 1995), Kahn's model depicts the network as a layered structure with each successive outer layer decreasing in relationship intensity (Antonucci & Akiyama, 1987; Kahn & Antonucci, 1980). Such models hold that respondents typically begin by nominating core (e.g., family and close friends) network members and then possibly extending their listed networks outward to include peripheral network members such as neighbors and work associates (Kahn & Antonucci, 1980). Importantly, all social network models hold that only core supporters are expected to provide all three types of support and especially tangible aid.

Given the tenets of Kahn's (1979) model and taking into account natural variations in network size, NSSQ respondents who nominate smaller networks likely restrict their nominations to core supporters. In contrast, respondents who nominate larger networks are more likely to include increasingly more peripheral network members who are not expected to provide a great deal of all three types of support: It is likely that not all peripheral network members like the respondent (affect), serve as confidants (affirmation) *and* provide tangible help (aid). Tangible help is the least likely form of support that a peripheral network member provides.

Critically though, because of the two-stage nomination/rating process previously described, NSSQ respondents are not aware that they are nominating network members whose support they will subsequently rate. It is only when they turn the first half page, after nomination is completed, that the first two support items are revealed. Thus, some respondents may make superfluous network nominations. In fact, Norbeck and her colleagues' (1983) test-retest results showed that respondents nominate fewer network members ( $t = 2.26, p = .03$ ) on the retest. It appears that respondents become more selective in their nominations once they know the purpose of network nomination and have knowledge of the support context in question.

It also important that respondents must rate all network members on all three types of support; even later-nominated members who may not provide all types (especially tangible aid). In other words, at present, network nominees cannot be temporarily "dropped" from the network if a support question does not pertain to them. Hence, lack of a not applicable response option impacts support scores.

## Support Scores

As House and colleagues (1985) pointed out, support scores are formed by summing respondents' ratings (0–4) of all network members. Because network size is allowed to vary (1–24), respondents can inflate their support scores by nominating artificially large networks. Averaged scores (support score divided by network size) should allow researchers to separate measures of network quality from network quantity. However, averaging presents a problem because without a not applicable response option, all nominees must be counted in the denominators (network size) even if they are rated 0 (*none at all*). This is because an investigator does not know if a respondent meant that the nominee should

be able to give them that type of support but does not, or that they would not be expected to provide that type of support. In such cases, the numerators (support scores) cannot keep pace with the inaccurately large denominators (network size) and score restriction occurs.

Score restriction was Norbeck's (1995) concern regarding averaged scores, and in the case of averaged Aid subscores, her concern was warranted (Gigliotti & Samuels, 2011). That is, although all *averaged* subscores decreased as network size increased (mean correlations: Affect =  $-.07$ ; Affirmation =  $-.02$ ; Aid =  $-.19$ ), only averaged Aid scores showed statistically significant negative correlations with network size. This is reflective of Kahn and Antonucci's (1980) findings that only core supporters are likely to provide tangible aid. Averaged Affect and Affirmation subscores are likely also affected but to a lesser degree than averaged Aid subscores.

### Initial Hypotheses

Based on these observations, we hypothesized that a portion of the measurement error in all NSSQ support subscores, and especially in Aid subscores, results from two inter-related sources: erroneous network nominations and subsequent erroneous support ratings. Specifically, we hypothesized that some respondents are unduly influenced by the presence of 24 network nomination spaces and, without knowing that they will subsequently rate these nominees on support items, nominate more network members than they would have nominated had they known the purpose of network nomination. It is likely that erroneous nominees tend to be peripheral network members who would not be expected to provide a great deal of all three types of support. However, their spurious inclusion not only artificially increases both network size and summed support scores, it also makes averaging unable to fully remove network size from mean support scores.

Furthermore, we hypothesized that, once nominated as a "significant person in your life" and in the absence of a not applicable response option, some respondents are reluctant to give low Affect and Affirmation ratings because of a social-desirability-like response option bias. It is likely that some respondents find it more acceptable to give low Aid ratings (Kahn & Antonucci, 1980). We reasoned that this strongly contributes to the fact that all three averaged subscores decreased as network size increased, but only averaged Aid subscores showed a statistically significant negative correlation with network size.

## METHOD

### Revisions to the Norbeck Social Support Questionnaire

To qualitatively test our hypotheses, we revised the NSSQ format (into what we call NSSQ-R.1) making the nomination/rating process more transparent. The NSSQ-R.1 asks respondents to nominate and rate one network member on all items before nominating and rating the next network member. Thus, respondents know the purpose of network nomination and the types of support in question before nominating the second network member. We then conducted cognitive interviews using Willis's (2005) protocol, asking respondents to "think aloud" as they completed either the NSSQ or the NSSQ-R.1. Our purposes were to (a) compare network nomination decisions made by participants completing either the original NSSQ or the NSSQ-R.1; (b) identify participants' difficulties, during subsequent support ratings, as a result of these network nomination decisions; and (c) evaluate the utility of adding a not applicable response option to the 0–4 support ratings currently used.

Willis (2005) notes that cognitive testing is an iterative process because new problems are sometimes identified during testing. Therefore, after the first round of cognitive interviews, and based on results of data analysis, we made two revisions to the NSSQ-R.1, creating the NSSQ-R.2. One revision was anticipated: adding the not applicable response option. The second revision addressed a previously unnoticed problem with the wording of the Aid items. We then conducted a second round of cognitive interviews/data analysis to identify any further problems participants had with this newly revised instrument, the NSSQ-R.2.

## **Participants**

With the institutional review board approval, we interviewed 16 community-dwelling adults in Phase 1. Seven participants completed the original NSSQ, and nine completed the NSSQ-R.1. In Phase 2, we interviewed six community-dwelling adults. All six participants in Phase 2 completed only the NSSQ-R.2.

## **Data Collection**

Participants were interviewed at a private site of their choice by one of the authors. Investigators first explained the study and obtained written consent to participate and to be audiotaped. Following Willis's (2005) protocol, participants were orientated to the cognitive interview process by being orally walked through an unrelated sample item by the investigator; they were then asked to think aloud as they completed either the NSSQ, NSSQ-R.1, or NSSQ-R.2. After completing their assigned version of the NSSQ, participants were asked the relevant structured probe questions reproduced in Table 2. Finally, participants were shown the alternate format(s) and asked which they preferred and why.

Although the data collection protocol was designed as a think aloud followed by structured probe, all participants benefitted from further coaching during the interview, for example, by being asked, "How did you arrive at that answer?" or "You seemed to hesitate, why?" This prompted participants to verbalize their decisions—the primary goal of a cognitive interview. In addition, most participants requested further clarification concerning how many network members to nominate and the written instructions were provided verbally: "You do not have to fill in all the network nomination spaces."

## **Analytic Approach**

Following Willis's (2005) recommended analytic approach, investigators reviewed their own audiotapes and notes that were handwritten during the interview and then identified/coded problems using a formal coding schema, as described in the following text. We expanded on this approach by reviewing and coding all interviews, not only our own, and then holding several group debriefing sessions to share impressions. We made the decision to not transcribe audiotapes because our intent was to evaluate respondents' comprehension and decision-making processes and found out effective. Hence, we deemed that any additional benefits transcription may have offered were not outweighed by the resources needed. We also noted Willis's caution that, in projects with purposes such as ours where the intent is to gain an impression of the participants' encountered difficulties, transcription can result in an overabundance of data and the hazard of not being able to "see the forest for the trees" (p. 161).

**TABLE 2. Structured Probe Questions**

NSSQ and NSSQ-R.1	Original NSSQ Only	NSSQ-R.1 Only	NSSQ-R.2 Only
What did the phrase “significant people in your life or people who provide personal support to you or who are important to you” mean to you?			What types of immediate help came to mind when you read that question?
How did you decide who to list in your support network?			What types of help over a period of several weeks came to mind when you read that question?
	Did the presence of 24 spaces influence you to list people?	What made you decide to stop adding people to your support network?	Why do you think those particular examples of help came to mind?
	What did you think when you turned the first page and saw that you had to rate each network member? Did you feel “duped”?		Did you find the “n/a” option to be helpful?
Did you find that some of the questions did not pertain to some of the people in your network?			
Can you give me some examples?			
If yes, how did you handle that? What score did you give those people and why?			

(Continued)

**TABLE 2. Structured Probe Questions (Continued)**

NSSQ and NSSQ-R.1	Original NSSQ Only	NSSQ-R.1 Only	NSSQ-R.2 Only
Would a not applicable category have helped?	Do you wish you had listed fewer or more people?	After rating the first person, did you decide not to list a person or some people because all the questions did not pertain to them?	
How easy or difficult was it to distinguish between those people who provided a great deal and those who provided a moderate amount of support?			

*Note.* NSSQ = Norbeck Social Support Questionnaire; NSSQ-R = revised Norbeck Social Support Questionnaire.

We used Willis, Schechter, and Whitaker's (as cited in Willis, 2005) coding schema: (a) comprehension: misunderstanding directions or questions; (b) recall: being unable to retrieve information; (c) bias/sensitivity: answering in a way the respondent thinks is socially acceptable; (d) response category: having a mismatch between response categories and the question; and (e) logical processing: showing difficulty in adhering to the cognitive process chain. In Phase 1, we sought to identify difficulties with both network nomination and support ratings. In Phase 2 respondents were only required to identify difficulties with support ratings.

## RESULTS

### Phase One

#### *Network Nomination Difficulties (NSSQ vs. NSSQ-R.1)*

*Brief Summary.* All seven participants who completed the original NSSQ identified comprehension difficulties during network nomination. In contrast, none of the nine participants completing the NSSQ-R.1 noted difficulties with comprehension. Both groups experienced difficulties with bias/sensitivity.

*Comprehension.* Participants completing the original NSSQ reported little difficulty understanding the network nomination directions and readily nominated network

members. All seven participants indicated quickly creating and adhering to a more or less restrictive decision rule for nominating network members (e.g., “people I am close to,” “the core and then the friends”). Similar to Kahn and Antonucci’s (1980) findings, all participants also noted nominating network members in order of relationship closeness or frequency of contact. Importantly however, none of those completing the original NSSQ reported understanding that the network members they nominated would then be rated on support items. One, who had listed 14 network members, commented, “I would have listed five people if I knew I had to answer all these questions!” Two reported feeling “duped” when they turned the page; one of these two noted that she would not have completed the ratings.

In contrast, all participants completing the NSSQ-R.1 based their nomination decisions on the type of support being rated rather than strictly on closeness of relationship or frequency of contact. Their statements reflect an understanding of what is being asked of them: “Once you read the questions, you knew what type of social support we [participants] were asking for”; “. . . others are acquaintances not support”; and “You are talking about real, real support here [pointing to the items].” Another participant noted, “It [the term *significant person*] didn’t mean a lot until I got to here [points to items] and, when I started seeing what the questions were, I sort of got the construct of what you were asking for.” All said they understood the extent of their involvement once they rated the first network member.

*Bias/Sensitivity.* All participants who completed the original NSSQ indicated they were influenced by the presence of 24 spaces, and all asked, “How many do other people usually list?” Although no one nominated 24 network members ( $M = 10.80$ ;  $SD = 2.94$ ), typical remarks included, “24 spaces! I don’t have that many [network members] . . . I would feel very inadequate.” Five of these seven participants noted that they nominated too many network members: “Having 24 spaces pushed me into my peripheral support and not sticking to my core.” One participant who nominated 14 members said, “If there were not 24 spaces I would have omitted 2, no maybe 3, maybe 4.” Another stated, “. . . it [24 spaces] made me reach for people on the fringes but I know I can call.”

Similarly, but to a lesser extent, all participants completing the NSSQ-R.1 reported being influenced by the 12-page length of the instrument, which presents two nominations per page, and asked if they should continue adding network members ( $M = 7.78$ ;  $SD = 5.30$ ). Some noted that the 12 pages made them feel that they should list more members, but all agreed they based their decision to stop adding members on the context of the support items. When subsequently shown the original NSSQ, all noted that, if they had not seen the support items, they would have been heavily influenced by the 24 spaces to list more people. Some commented that the revised format was more transparent; they knew what was expected before nominating network members. One participant noted, “I would have listed more people if I did not see all the questions . . . I would have had a shock . . . I would have felt duped. What do I do now after I said they were important [to me]?”

#### ***Support Rating Difficulties (NSSQ vs. NSSQ-R.1)***

*Brief Summary.* Both NSSQ and NSSQ-R.1 respondents noted similar bias/sensitivity and response category difficulties that a not applicable response category for support items should ameliorate. However, this difficulty was more pronounced in the original NSSQ group. Bias/sensitivity problems will be presented in the following text first, followed by the unanticipated comprehension difficulty both groups showed with the wording of the Aid items. Finally, we cover minor logical processing problems.

*Bias/Sensitivity and Response Category.* The absence of a not applicable (n/a) response option appeared to bias all participants' responses. Although more apparent in the original NSSQ group, the majority in both groups noted that lack of an "n/a" response option caused them to give some network members a socially acceptable but inaccurately high rating. All agreed that not applicable lacks the negative connotation that 0 (*none at all*) or a low rating 1 (*a little*) and 2 (*moderately*) has.

All seven original NSSQ participants indicated that they were reluctant to rate network nominees lowly and told us that they were giving some inaccurately high ratings because of this. Notably, in these cases, some respondents even gave ratings of 3 (*quite a bit*) or 4 (*a great deal*). This was more often the case for Affect and Affirmation items, perhaps because of a social desirability bias. Participants were less reluctant to give low ratings and even 0's for Aid items. Participants completing the original NSSQ noted that an n/a option would have helped them to accurately rate all network members; one said, "I probably would have put n/a for most of it and left [rated] my kids and husband . . . n/a would really help out . . . you don't have to feel bad. It doesn't make them bad people or me." "Instead of giving a 2, n/a has a different connotation. If you would not expect to confide in that person you could put n/a. It would help if you do not want to give 0."

Participants completing the NSSQ-R.1 agreed, "It is better for me to put n/a. None at all [0] underlines my negative . . . it means that we have no relationship," and "There has to be another [response] category for those relationships and for those who live far away." In addition, four of the nine NSSQ-R.1 participants said that they would have listed more people if they had the n/a option: "n/a could be used for neighbors and coworkers," and "n/a would help tremendously. I have professional friends . . . I would have put more people."

*Comprehension.* No participant in either phase had a problem understanding the first four items measuring Affect and Affirmation (see Table 1). However, the next two items, which measure perception of tangible help, presented difficulty to most participants. Norbeck and her colleagues (1981) embedded examples in these Aid items. The first item concerns immediate short term help: "If you needed to borrow \$10, a ride to the doctor, or some other immediate help, how much could this person usually help you?" The second item concerns sustained help for a circumscribed period of time: "If you were confined to bed for several weeks, how much could this person help you?"

Unexpectedly, regardless of format (NSSQ or NSSQ-R.1), no participant appeared to understand that the mention of \$10 or of a ride were merely *examples* of immediate help. All participants focused on the examples as guides for the type of help to be given. That is, they focused on the issues of whether they would ask for money, a network member's ability to drive, that the member could not drive because of geographical distance, and so forth. Still, others demonstrated difficulty deciding how to rate a member if that member could provide one type of aid but not the other: "Now you are asking two different things . . . sometimes you are not available to give a ride but money is not a problem." This same participant also noted that she only saw the examples: "I didn't even see [the phrase] 'some other immediate help.' Leave out the examples and say 'if you needed immediate help.'"

Likewise, the second Aid item ("If you were confined to bed for several weeks . . .") posed comprehension problems for most participants. Interestingly, this item is not meant to determine how much help one would have during a medical crisis but rather if a network member could provide sustained, tangible help. Perhaps this is because the embedded example connotes bedside assistance and thus physical ability as well as geographic proximity. This confusion is evident in participants' remarks: "Physical proximity is important

when you choose support people”; “Geography influenced my answers”; and “If I needed physical help, she would be unable to provide it so I gave her a 2.” Only one participant seemed to understand that sustained help takes many forms and is applicable across multiple contexts, noting, “A significant person is when you are in trouble and you can call them and ask for help.”

*Logical Processing.* Several participants displayed difficulty rating some network members on the Affect and Affirmation items. For example, one participant noted that her daughter usually likes/respects/agrees with her and she can usually confide in her, but she said that they just had a fight and she was reconsidering. Another participant noted that it was hard to rate network members regarding confiding in them because it was she who chose not to confide. Still another questioned whether a member not agreeing with the respondent is a bad thing if they have one’s best interests at heart. In these instances, an n/a option would not be appropriate; participants eventually indicated that they settled on a rating they were pleased with. One noted that “it did not take a long time just sort of the process of averaging.” Logical processing may have also been a problem for the Aid items, but this is difficult to evaluate because all participants focused so heavily on the examples.

*General Format Observations.* All participants were asked their opinion of the alternate NSSQ-R.1 format. Four of the five original NSSQ participants preferred the NSSQ-R.1 format noting that they liked rating one network member on all questions before nominating and rating the next network member. The one participant who did not like the revised format stated that the scoring bubbles looked too much like a test. Likewise, eight of the nine NSSQ-R.1 participants preferred the revised format, expressing sentiments such as “I do not like an old one. It is tedious. You have to keep looking up [to find the rating on top]. The new one is easier to use because you can see all questions and easily can fill in answers—boom, boom, boom.”

*Conclusions.* Bias/sensitivity issues related to lack of an n/a response option presented regardless of format type. Because of the negative connotation of a *not at all* (0) rating, some original NSSQ participants inflated their support scores, giving inaccurately high ratings. This occurred most often with Affect and Affirmation items. In contrast, lower ratings and some 0s were given for tangible Aid items. However, it is likely that these lower Aid ratings were the combined result of miscomprehending the intent of the aid examples as well as the lack of an n/a response option.

Although some inaccurately high ratings did occur among the NSSQ-R.1 participants, these were less frequent likely because they had not nominated network members who could not or would not be expected to support them in that way. Nevertheless, some NSSQ-R.1 participants appear to have had inaccurately low support scores because they would have nominated more network members if n/a was an option. It is recommended adding n/a response option to administrations of the original NSSQ format. Regarding the aid examples, these caused comprehension problems for most participants. Miscomprehension is likely a major source of measurement error in the Aid items. We therefore eliminated the examples from the Aid items.

## Phase Two

Based on results of Phase 1, the NSSQ-R.2 was created by modifying the NSSQ-R.1 format in two ways: adding an n/a response option and removing the examples from the Aid items. These Aid items now read: “If you needed immediate help, how much could this person usually help?” and “If you needed help for several weeks, how much could this

person help you?" We then conducted a second round of cognitive interviews to identify problems with these revisions or further problems that were the result of these revisions. Support rating difficulties were the only focus in this round. All participants in phase two completed the NSSQ-R.2.

### ***Support Rating Difficulties***

*Bias/Sensitivity and Response Category.* Interestingly, none of the six participants used the n/a response option. Some did not see it and suggested moving it away from the other responses to make it a "stand-alone" option. Most, however, noted that they saw it but did not need to use it because they had only nominated network members who could provide all three types of support. Notably, the removal of the aid examples may have decreased the need for the n/a option because participants could find some instance when a nominee could be expected to provide some form of immediate and/or sustained help or perhaps nominated fewer people who could not provide aid.

*Comprehension.* Regarding the first Aid item (immediate aid), all participants noted providing ratings based on their particular situations and relative to that particular network member. All seemed to focus on being able to "count on" a network member if they needed some form of immediate help. For example, one noted that her elderly grandfather could not come to help her with a flat tire, but she could call him and he would make sure to send someone who could provide aid: "I know I can count on him." Another participant, a young mother attending college, said, "I can call her [neighbor] and she will get my daughter off the bus if I am running late at school." Notably, core network members whom participants said could provide immediate help in all areas of their lives were, with few exceptions, rated highly. Other, more peripheral network members, whom participants noted could provide help in a more limited sphere were accurately rated lower. None were rated 0.

Regarding the second Aid item, concerning help for several weeks, participants again demonstrated tailoring their responses to their particular situations. When asked what type of help came to mind, each participant recounted a time when they needed sustained help. Examples included after losing their home to Hurricane Sandy, childcare needs for working parents, renovating a newly purchased home, and undergoing chemotherapy and radiation therapy. Support ratings were also varied based on the person's ability to provide the kind of sustained help they had needed: "I would say moderately [2] because she has her own family and just can't be that available." Again, no ratings of 0 were given.

## **CONCLUSIONS**

Based on these cognitive interviews, we retain and expand our initial hypotheses regarding erroneous network nominations, we revise our hypothesis regarding lack of an n/a response option, and we propose a new hypothesis regarding the effects of examples in the Aid items. We also recognize inherent limitations in this study's design.

### **Network Nomination**

We conclude that a portion of measurement error is indeed the result of extraneous network nominations because of the original format's 24 nomination spaces and concealment of the support items. These erroneous nominations result in errors in both network size/composition and support scores. NSSQ-R.1 participants noted feeling more comfortable

with their decision to stop nominating network members than did original NSSQ participants. Moreover, original NSSQ respondents who feel duped after completing network nominations may not complete the support ratings or return the survey at all and this would result in nonresponse error when trust is violated (Dillman, 2007).

### **Support Ratings**

Lack of an n/a response option does result in score restriction for averaged Aid subscores by keeping denominators (network size) erroneously high. In addition, we found that miscomprehension of the purpose of the embedded examples in the Aid items also contributes to score restriction by lowering ratings of tangible aid. Because support ratings are a component of support scores, this lowers the numerators of averaged Aid items.

Less clear are the effects of lack of an n/a response option on increased social desirability bias, and hence, score inflation for Affect and Affirmation items because the n/a option was not used when added to the NSSQ-R.2. This seems, in part, because of some participants not seeing the n/a option; others saw it but did not feel the need to use it because the revised nomination method allowed them to choose members who were likely to give all types of support. Although the n/a option was not used in this small sample, we expect it will be useful for larger administrations of the instrument with participants, for example, who may nominate more peripheral members who can give some types of support but not all. In addition, it may be that knowing there is an n/a option will let participants feel more comfortable nominating people who cannot provide all types of support, helping give a more accurate portrayal of their networks.

### **Study Limitations**

This study was designed as a think aloud followed by structured probes because the intent was to evaluate participants' decision processes when completing a questionnaire designed to be self-administered (Willis, 2005). However, as noted, most participants benefitted from frequent coaching such as asking, "How did you arrive at that answer?" This was judged a function of individual differences in talkativeness and/or educational level. Although Willis notes that such probing is often necessary and actually occurs in most cognitive interviews, he also notes that it has its drawbacks. These include possible participant reactivity and bias. For example, such coaching may have prompted participants to consider more carefully the network nomination or support rating at the time of the prompt, and this may influence subsequent nominations and ratings.

Likewise, with few exceptions, we needed to clarify the written network nomination directions by telling participants that they did not need to use all the network nomination spaces. Although this was necessary because participants asked, possible bias and reactivity are likely. We do not know if participants would have nominated more network members. However, after analyzing answers to the subsequent structured probe questions, we were satisfied that bias and reactivity were minimal.

### **Future Directions**

We plan a quantitative analysis to systematically investigate the effects of each hypothesized source of error on resultant NSSQ support. Effects of incremental revisions will be investigated to measure the specific effects of each on support scores. We will begin by

addressing bias that is localized to the Aid items by removing the aid examples in a version of the NSSQ that will retain the original instrument's format and network nomination/rating process.

The second revision will add an n/a response option to the instrument to address possible score inflation in the Affect and Affirmation subscales. We predict that this revision will remedy score restriction in averaged Aid subscores because denominators (network size) can be accurately adjusted by "dropping" a network member from that item's calculation. Furthermore, because the Aid examples are already removed, we expect that numerators (Aid support scores) will not be lowered by miscomprehension. We expect this will help generate two independent measures: network size and of level of social support. If further investigation supports the convergent validity of these two measures, then new areas of research using the NSSQ will emerge.

The final revision will be the most extensive. This version will resemble the NSSQ-R.2. While retaining the previous changes, it will also completely change the way network members are nominated and rated. The final format should result in fewer network nominations than the other three formats. In addition, respondents completing the final version should show better response rates and less missing support rating data than is found for the other three formats.

## REFERENCES

- Ajrouch, K. J., Antonucci, T. C., & Janevic, M. R. (2001). Social networks among blacks and whites: The interaction between race and age. *The Journals of Gerontology. Series B: Psychological Sciences and Social Sciences*, *56B*(2), S112–S118.
- Antonucci, T. C., & Akiyama, H. (1987). Social networks in adult life and a preliminary examination of the convoy model. *Journal of Gerontology*, *42*(5), 519–527. <http://dx.doi.org/10.1093/geronj/42.5.519>
- Byers, P. H., & Mullis, M. R. (1987). Reliability and validity of the Norbeck social support questionnaire in psychiatric inpatients. *Educational and Psychological Measurement*, *47*(2), 445–448.
- Dillman, D. A. (2007). *Mail and internet surveys the tailored design method* (2nd ed.). Hoboken, NJ: Wiley.
- Dunbar, R. I. M., & Spoor, M. (1995). Social networks, support cliques, and kinship. *Human Nature*, *6*(3), 273–290.
- Gigliotti, E. (2002). A confirmation of the factor structure of the Norbeck Social Support Questionnaire (NSSQ). *Nursing Research*, *51*(5), 276–284.
- Gigliotti, E. (2006). A confirmatory factor analysis of situation-specific Norbeck Social Support Questionnaire items. *Nursing Research*, *55*(3), 217–222.
- Gigliotti, E., & Samuels, W. E. (2011). Use of averaged Norbeck Social Support Questionnaire scores. *ISRN Nursing*, *2011*, 567280. <http://dx.doi.org/10.5402/2011/567280>
- House, J. S., Kahn, R. L., McLeod, J. D., & Williams, D. (1985). Measures and concepts of social support. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 83–108). San Diego, CA: Academic Press.
- Kahn, R. L. (1979). Aging and social support. In M. W. Riley (Ed.), *Aging from birth to death: Interdisciplinary perspectives* (pp. 77–91). Boulder, CO: Westview Press.
- Kahn, R. L., & Antonucci, T. C. (1980). Convoys over the life course: Attachment, roles, and social support. In P. Baltes & O. Brim (Eds.), *Life-span development and behavior* (pp. 253–286). New York, NY: Academic Press.
- Norbeck, J. S. (1985). Types and sources of social support for managing job stress in critical care nursing. *Nursing Research*, *34*(4), 225–230.
- Norbeck, J. S. (1995). *Scoring instructions for the Norbeck social support questionnaire (NSSQ)*. San Francisco, CA: University of California.

- Norbeck, J. S., Lindsey, A. M., & Carrieri, V. L. (1981). The development of an instrument to measure social support. *Nursing Research, 30*, 264–269.
- Norbeck, J. S., Lindsey, A. M., & Carrieri, V. L. (1983). Further development of the Norbeck Social Support Questionnaire: Normative data and validity testing. *Nursing Research, 32*(1), 4–9.
- Pollet, T. V., Roberts, S. G. B., & Dunbar, R. I. M. (2011). Extraverts have larger social network layers. *Journal of Individual Differences, 32*(3), 161–169.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.

Correspondence regarding this article should be directed to Eileen Gigliotti, PhD, RN, The College of Staten Island and The Graduate Center, The City University of New York. E-mail: Eileen.gigliotti@csi.cuny.edu