How Valid Are the Responses to Nursing Home Survey Questions? Some Issues and Concerns

Denise A. Tyler, PhD, Renée R. Shield, PhD, Marsha Rosenthal, PhD, Susan C. Miller, PhD, Terrie Wetle, PhD, and Melissa A. Clark, PhD

1Department of Community Health, Center for Gerontology and Health Care Research, Brown University, Providence, Rhode Island.
2Institute for Health, Health Care Policy, and Aging Research, Center for State Health Policy, Rutgers University, New Brunswick, New Jersey.

*Address correspondence to Denise A. Tyler, PhD, Department of Community Health, Center for Gerontology and Health Care Research, Brown University, Box G-5121-6, Providence, RI 02912. E-mail: denise_tyler@brown.edu

Received July 4, 2010; Accepted October 14, 2010

Decision Editor: Nancy Schoenberg, PhD

Purpose: Although surveys are usually piloted before fielding, cognitive-based testing of surveys is not standard practice in nursing home (NH) research. Many terms used in the literature do not have standard definitions and may be interpreted differently by researchers, respondents, and policy makers. The purpose of this study was to ensure that survey respondents understood questions as intended, determine whether the Nursing Home Administrator (NHA) or the Director of Nursing (DON) was better able to answer questions on certain topics, and to inform the answer choices provided for questions. Methods: Using existing survey questions and input from experts, we developed surveys to be administered to DONs and NHAs. Cognitive-based interviews were conducted with 45 participants. We took detailed notes during all interviews, and two researchers independently coded these notes for key themes. Results: Many terms and concepts routinely used by NH researchers and policy makers, such as “direct-care workers” and “palliative care,” were not uniformly interpreted by those managing NHs. For example, respondents’ definitions of direct-care workers ranged from nursing assistants to broader categories of clinical and other staff members, including nurses, activities staff, and social workers. We also found NHAs and DONs, at times, did not possess or have access to information the researchers expected them to. Implications: Our results may help explain discrepant findings across NH studies. They also underscore the necessity of cognitive-based testing for survey development and have important implications for policy decisions.

Key Words: Cognitive-based testing, Survey development, nursing homes, Long-term care policy

Although it is a well-recognized standard in high-quality survey development, cognitive-based testing of surveys has not been standard practice in long-term care research or has not been commonly reported in research publications. Although focus groups and pilot testing have been routinely utilized, few nursing home (NH) studies have reported including cognitive testing as part of survey development (e.g., see Frank, Flynn, & Rothman, 2001; Housen et al., 2008; Katz et al., 2009). Cognitive-based testing is the systematic testing of survey questions intended to determine how respondents understand the question and the thought process involved in providing an answer (Presser et al., 2004). This form of testing is used to establish how respondents understand and interpret questions,
to determine how and why they choose specific responses, and to identify questions that are difficult for respondents for a variety of reasons. The main objectives of cognitive-based testing are to reduce measurement error, increase validity, and improve response rates (Jobe & Mingay, 1989) by identifying instances where researchers have different interpretations of terms or concepts from those of respondents. Unlike pilot testing, cognitive-based testing examines the assumptions researchers may make about how terms are defined and understood without realizing that these assumptions may not be shared. Cognitive-based testing is also useful for determining those instances where respondents may simply not possess the information the researcher is seeking to collect.

Identifying differing interpretations of terms and concepts regarding philosophies and systems of care may be especially important when new trends emerge and then evolve within the field, such as NH “culture change” and “palliative care.” For example, the Consumer Assessment of Health Providers and Systems survey was originally designed to include measures of “coordination of care” and “shared decision making,” but these constructs ultimately were not included in the final survey because uniformly understood questions could not be developed (Levine, Fowler, & Brown, 2005). Furthermore, researchers, clinicians, administrators, and consumers may have different perspectives or understanding of such complex concepts. A recent study found that those in the long-term care field who were not directly involved in the provision or study of palliative care knew little about its philosophy or strategies (Lima, Miller, & Shield, 2009). Cognitive-based testing can play an important role in developing surveys that more accurately measure practitioners’ understanding of and response to emerging concepts.

Determining where mismatches occur between the interpretation of terms and concepts by researchers and those of long-term care professionals may also have important implications for future research and policy and may explain discrepant findings in past research. For example, numerous studies have examined turnover among NH employees and found widely varying results (Barry, Kemper, & Brannon, 2008). A review of the turnover literature by Castle (2006) revealed a wide range of reported turnover rates for nurse aides (14% to 346%), licensed practical nurses (LPNs, 8% to 103%), and registered nurses (RNs, 19% to 64%). This may be related in part to discrepant interpretations of the concept of “turnover” between those designing surveys and those responding to surveys or may be related to other measurement issues not previously discovered.

The purpose of this study was to cognitively test two instruments, one designed to be administered to Nursing Home Administrators (NHAs) and one to directors of nursing (DONs) in a national study. We sought to ensure that survey respondents understood questions as intended, to determine whether the NHA or the DON was best able to answer questions on certain topics, and to inform the answer choices provided for questions. In doing so, we also identified a number of issues that raise concerns about the validity of previous questionnaires and the assumptions implicit in some long-term care policy.

**Methods**

This research is part of a larger National Institutes of Health program project involving several studies relying on primary data collection and the building of a database of secondary NH administrative data (e.g., Online Survey, Certification, and Reporting [OSCAR] data collected during each NH’s annual survey inspection and reported to the Centers for Medicare and Medicaid Services [CMS]; Minimum Data Set [MDS] data, which includes demographics, conditions, and psychosocial information, collected at least quarterly for all residents in Medicare and Medicaid certified NHs and reported to CMS; and Medicare billing data). Multiple projects will utilize this secondary database and link with data from the NHA and DON surveys. The program project is housed within a Center for Gerontology that includes access to a survey center with staff specially trained to collect survey data by e-mail, telephone, and web-based administration. Because each project is focused on a different area within long-term care, the surveys of the NHAs and DONs are designed to collect information on a variety of different topics, such as hospice and palliative care, NH staffing, and culture change.

We began instrument development with a careful review of the literature to identify and gather previously existing surveys on each project’s topic. When available, we selected only instruments and individual questions that had been previously tested, with good validity and reliability reported. We also consulted with experts in each topic area to seek advice about which specific questions from
existing instruments to use and where to develop new questions. We drafted separate questionnaires to be administered to NHAs and DONs. Some questions were included in both questionnaires.

Once first drafts of the questionnaires had been created, the cognitive-based interview process began. As described in detail subsequently, we first conducted in-person interviews in our state and then in the wider region and followed these with telephone interviews nationwide. In all cases, the interviews were structured and focused probing techniques were used. Using this method, respondents were first asked the survey question and were then asked follow-up questions about how they responded to the question and came to an answer or understood the question. This “thinking out loud” process allows the researcher to determine how respondents interpret terms, use strategies to answer questions (Jobe & Mingay, 1989), and arrive at the responses they give. For example, one of our survey questions asked, “What proportion of the rooms in your nursing home are private rooms”? After answering this question, respondents were asked the focused probe, “What is considered a private room in your nursing home”? Responses to this question varied based primarily on whether a private room included its own bathroom, a consideration we had not initially contemplated.

As shown in Table 1, participating facilities varied in terms of size, profit status, chain membership, and hospital affiliation. Figure 1 depicts an overview of the phases of the cognitive-based testing. During all phases of the cognitive-based testing, each participant was compensated 50 dollars. Because no personal information was collected, this study was determined by the Institutional Review Board to be exempt from review.

Initially, varying two-member combinations of a three-person team conducted each of four in-person pilot interviews with NHAs and DONs in Rhode Island (RI), debriefing with each other following each interview. During each interview, one team member conducted the interview and took notes whereas the other team member observed and took notes. After ensuring that all interviewers were following a uniform approach, the three

Table 1. Characteristics of Participating Nursing Homes (NHs)

<table>
<thead>
<tr>
<th></th>
<th>In-person interviews (13 NHs)</th>
<th>Telephone interviews (11 NHs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>For-profit</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Nonprofit</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Government</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Part of chain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>No</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Hospital based</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>50-150</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>&gt;150</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 1. Cognitive-based interview process.

* Results reported elsewhere.
interviewers then individually conducted a total of four more interviews in RI.

When these were completed, the three-person team met with the survey methodologist to review and compare preliminary findings regarding NHA and DON responses and comprehension. These preliminary findings were next discussed by the larger research team, and questions were refined accordingly. In many cases, questions were shown to be understood as we had intended them and were therefore not revised. Some questions were shortened and/or simplified. This occurred when questions included more explanation than necessary because respondents were found to consistently interpret terms as expected based on the literature and experts and therefore did not need to have terms defined within the question. For some questions asked of both the NHA and DON, a determination was made about who was the more appropriate respondent.

Following these first revisions to the questionnaires, the OSCAR system data were used to randomly select NHs throughout New England for further in-person cognitive-based interviews. We oversampled hospital-based facilities to ensure these were included in our interviews.

For these 17 in-person interviews, researchers worked in teams, with one researcher interviewing the NHA while the other researcher interviewed the DON. The interviews ranged between 60 and 90 min in length. The interviewers wrote detailed notes directly on the questionnaire form during the interviews to preserve as much of the verbatim responses as possible. Interviews were taped as a backup to and clarification of note taking. Responses were entered on NHA- and DON-specific response grids to allow comparison among them. Detailed scrutiny of these grids by the interview team provided the basis for identifying problematic question construction and determining who the most suitable target (NHA or DON) for each question should be. Interviews were conducted until saturation had been reached in that no new additional information was being garnered through the interviews (Kuzel, 1999).

After this second round of in-person interviews, problematic questions whose meanings were not uniformly understood or apparent to respondents were again identified by the interview team and survey methodologist, and the surveys were modified in consultation with the larger research team. An audit trail of decisions and survey modifications was kept throughout the process.

Next, a national sample of NHs was drawn from the OSCAR data, and 20 additional interviews of NHAs and DONs were conducted by telephone. The purpose of telephone interviews was twofold: (a) this allowed us to conduct cognitive-based testing in a national sample of NHs and (b) the feasibility of telephone survey administration was assessed. Conducting cognitive-based interviews nationwide was important because we needed to ensure that survey questions could be understood by a national sample of respondents and did not include regionally biased terminology or concepts. However, because budget restrictions did not allow for nationwide travel to conduct in-person interviews, telephone interviews were conducted. These were carried out in the same manner as the in-person interviews in that respondents were first asked the survey question and then asked focused probes. In addition, because the full-scale survey administration would include telephone interviews of mail nonresponders, we needed to ensure that the surveys could be administered by telephone and that visual prompts (i.e., answer choices) were not necessary.

During this round, members of the three-person team each conducted both the NHA and DON interviews for their assigned facilities, and in two cases just the NHA or DON was interviewed instead of both (hence 11 facilities yielded 20 interviews). Interview responses continued to be entered into the response grids throughout this process to inform survey refinement decisions. Telephone interviews were conducted until saturation had been achieved. Based on the results of the telephone interviews, the surveys were further modified and a field test was conducted (results to be reported elsewhere).

For the purposes of this paper, two researchers analyzed the response grids and the notes taken during all rounds of cognitive-based interviewing. They independently identified areas of discrepancy in the NHAs’ and DONs’ interpretation of common concepts and terms used in long-term care research and policy, identified questions that respondents had difficulty answering due to a lack of knowledge or information on the subject, and coded these for important themes. The larger research group then refined and clarified the scope of these themes.

**Results**

A number of questions needed no modification because they were understood by respondents as
Table 2. Revision of Palliative Care Questions

<table>
<thead>
<tr>
<th>Initial palliative care question</th>
<th>Revision(s)</th>
<th>Final resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHA Question:</td>
<td>Same question except, “Does your facility have . . .”</td>
<td>Question was eliminated from NHA survey.</td>
</tr>
<tr>
<td>Please tell me if this facility has a special program that has specially trained personnel dedicated to the program for anything listed below. This does not include special training that is provided to all personnel.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Hospice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Palliative care/end-of-life (end state/terminal condition—not hospice)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Pain management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DON Questions:</td>
<td>Question split into 2 questions:</td>
<td>Question was eliminated from the DON survey and replaced with:</td>
</tr>
<tr>
<td>Please indicate whether terminally ill residents at your nursing home can access the professionals listed below?</td>
<td>1) Are there nurses at your facility who have been certified by the Hospice and Palliative Nurse Association as having palliative care expertise?</td>
<td>Do you collaborate with at least one hospice provider?</td>
</tr>
<tr>
<td>a. Nursing home nurses who have received continuing education in symptom management or palliative care</td>
<td>a. No</td>
<td>a. No</td>
</tr>
<tr>
<td>b. Nursing home nurses who have been certified as having palliative care expertise</td>
<td>b. Yes</td>
<td>b. Yes</td>
</tr>
<tr>
<td>c. Hospice nurses (if the resident is enrolled in hospice)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Other pain and symptom management experts, such as through a palliative care consultation, a pain management department or clinic</td>
<td>2) On a scale from 1 to 5 with 1 being very insufficient and 5 being very sufficient, how would you describe the ACCESS to palliative care professionals at your facility?</td>
<td>On a scale from 1 to 5 with 1 being very insufficient and 5 being very sufficient, please CIRCLE the number that best describes the ACCESS to palliative care expertise at your facility?</td>
</tr>
<tr>
<td>Very insufficient...................................................................................................................</td>
<td>Very sufficient 1 2 3 4 5</td>
<td>Very insufficient...................................................................................................................</td>
</tr>
</tbody>
</table>
the researchers intended. However, three basic problems in some of the questions were identified via this cognitive-based interviewing strategy: (a) respondents were unable to answer questions as originally worded because they did not possess the necessary information, (b) the question did not provide enough information for the respondent to understand what was being asked, or (c) respondents had interpreted terms or concepts differently from the researchers’ intention. Although we expected respondents to have difficulty with questions about newly emerging concepts and practices, such as “palliative care” and “culture change,” we also found that we had made assumptions about respondents’ ability to answer questions about operational terms like “nurses” and “direct-care staff.” The issues related to these conceptual and operational constructs follow subsequently.

Conceptual Issues

Palliative Care.—Through the NHA and DON surveys, we attempted to identify which NH palliative care programs and staffing were available to address residents’ end-of-life care needs. To do this, we slightly modified a question from the 2004 National Nursing Home Survey (NNHS, Centers for Disease Control & Prevention, 2004) to ask NHAs about their NHs’ palliative care programs (see NHA question in Table 2). We also constructed a question to ask DONs about the availability of palliative care nurse expertise and pain and symptom management expertise (see DON question in Table 2).

National experts have agreed that palliative care is “both a philosophy of care and an organized, highly structured system for delivering care”; it is “operationalized through effective management of pain and other distressing symptoms, while incorporating psychosocial and spiritual care according to patient/family needs, values, beliefs and culture(s)” (National Consensus Project for Quality Palliative Care, 2004). In the 2004 NNHS, an estimated 17% of U.S. NHs reported having a palliative care/end of life care “special program that has specially trained personnel dedicated to the program” (Miller & Han, 2008). However, these NNHS data provided no information on what NHs meant when they reported having a special program for palliative/end-of-life care. Considering this and given that previous research showed interview participants (e.g., health care providers, policy makers, others) could readily discuss hospice care, but were often unclear as to what was meant when asked about palliative care (Miller & Shield, 2007), cognitive interviewing relating to the palliative care program and staffing questions was important.

We found the NHA question on palliative care programs to be confusing to all respondents, with respondents focusing their responses on training rather than on whether the facility had a program in place. In fact, when NHAs were asked to “Please describe what you were thinking about when I asked you about NON-Medicare HOSPICE palliative care programs”? most administrators conveyed uncertainty. Responses such as “Not sure; volunteer group while patient is declining”? or “Not really; don’t know what you’re referring to” were common. Other NHAs described palliative care by its nonmedical components, such as one NHA who said, “I’m familiar with palliative care programs with clergy and respite care and volunteers.” Even with some revisions, this question remained problematic, with the validity of responses questionable. Therefore, the question was ultimately eliminated from the questionnaire.

The initial DON palliative care question asked about terminally ill residents’ access to NH nurses who received training in symptom management or palliative care, nurses with certified palliative care expertise, hospice nurses, and other pain and/or symptom management experts (Table 2). Almost all DONs answered “yes” to having nurses with training in symptom management or palliative care. When asked “What types of continuing education have these nurses received”? answers were quite diverse. One DON mentioned “In-services by hospice” whereas another mentioned “Conferences, in-services,” and yet another stated “One of our nurses worked for hospice prior.” Most DONs also answered “yes” to having access to pain and symptom management experts. Similarly, when asked “What types of individuals were you considering when answering this question”? responses were diverse. One DON answered that there are “alternative therapies” but provided no specifics and another answered “Vendors, doctors, NP” whereas two DONs mentioned hospice services, and another noted that the facility’s “In-house medical director is a palliative care expert.” We concluded that we would not obtain adequately discrete data from responses to these questions and decided to eliminate them.
During the cognitive-based interviews, alternative questions were suggested to DONs, and a question that seemed to be more meaningful to them was “Are there any times when you feel that you do not have ENOUGH access to palliative care professionals”? Therefore, a rating question relating to sufficiency of palliative care access was added to the survey and was revised slightly for the final version of the survey (Table 2).

Finally, the question about access to hospice nurses was only answered “no” when DONs reported that their facility did not have a hospice contract. So, we substituted a question asking if the facility collaborated with at least one hospice provider (Table 2).

**Culture Change.**—Through cognitive-based testing, we learned that most respondents reported their facilities were involved in culture change activities, and when asked “How do you define culture change”? gave examples compatible with culture change constructs and core values that are widely accepted among researchers that were read to them. However, when further probing occurred, answers on the extent of a facility’s culture change were not always consistent with reports of activities undertaken. For example, when a DON that reported culture change “has completely changed the way we care for residents in some or all areas of the organization” was asked to “Please provide a few examples of culture change activities,” she said “Communication is first and foremost and giving staff power to do it, educating the staff and setting the standard with staff. The culture of the building was not positive.” In response to the subsequent probe, “How many years has your NH been involved in culture change activities”? she responded “less than a year” and when asked “What would you say was your facility’s first culture change activity”? she reported “Bringing in a new administrative team and seeing that they stayed longer than six months.” Whereas these reported activities are no doubt aimed at improving quality, most are not consistent with culture change constructs or with a facility that is considered to have “completely” implemented culture change.

As illustrated previously, respondents’ examples of culture change activities did not always elicit examples in agreement with researchers’ common understanding of culture change practices. Also, even when there was agreement, respondents still often reported that they adopted culture change or resident-centered care in superficial or piecemeal fashion. For example, a DON qualified that her facility had adopted consistent assignment of nursing assistants, but then noted “This is depending on the unit manager; some rotate and some don’t. Even the units that have consistent assignment will rotate if the staff gets lackadaisical.” However, an NHA seemed to recognize that changing the names of wings was a superficial change that did not necessarily constitute a culture change practice: “What we really have is two different wings and one is called ‘Maple Street’ and one is called ‘Vineyard Court.’ That relates to the idea of culture change, but that’s all we’ve done.”

In response to these findings, the research team decided that a facility’s participation in culture change could not be measured solely on the response to the global question but needed to be further measured by their answers to specific practice questions. Preliminary quantitative analyses have also found the response to the global question did not adequately distinguish facilities as having little versus much culture change adoption. In particular, the facilities reporting “Culture change has partially changed the way we care for residents . . .” varied widely on the accepted culture change practices reported to be in place. Because of this finding, we have begun to consider the responses to the global question together with responses to the individual practice questions to classify NHs as having little or much culture change adoption.

**Operational Issues**

**Staffing.**—Although we had expected that cognitive-based testing would expose difficulties respondents had in answering questions about complex concepts and processes, such as palliative care and culture change, we were surprised to find that respondents also had difficulty answering questions about operational terms, such as nurses and direct-care staff. In some cases, cognitive-based testing revealed discrepant meanings of terms between the researchers and respondents or among respondents. In other instances, the testing revealed that respondents did not possess, or have access to, basic information about facility staff. The difficulties respondents had in understanding and in accessing appropriate information in these instances raise questions about the validity of survey data reported in previous studies.

Several survey questions asked about direct-care staff. During cognitive-based interviewing,
NHAs and DONs were asked to report which staff members they were thinking of when they heard the term direct-care staff. A wide variety of answers were given. For example, one DON said “RNs, LPNs, and CNAs,” whereas another said “Nursing staff and head nurses, LNAs (licensed nurse aides), and activities staff.” NHAs were even more inclusive. For example, one stated that direct-care workers were “anybody on the nursing staff, social workers, anybody who documents in the record.”

Our response to this finding was to revise questions so as to be specific as to which type of staff we were referring, such as using the terms “nursing assistants” or “nurses (RNs and LPNs)”.

We also found that many respondents did not always distinguish between RNs and LPNs on staff and, therefore, had difficulty reporting separately the number of RNs and LPNs employed. For example, one DON said, “We have 120 nurses and about a fourth are RNs.” Similarly, respondents were unsure of whether to include nurses in management positions in their counts. As an NHA stated, “Do you mean just direct-care because we have an RN who is the MDS coordinator”? Again, our response to these issues was to revise questions so that it was clear who should be included, such as in the following example:

For the next three questions, staff refers to all NON-MANAGEMENT employees of the nursing home in all departments.

Respondents had similar difficulties when asked about the physicians who provide care in their facilities. Several NHAs reported being unable to give precise numbers. For example, one NHA stated, “I can’t tell exactly how many come in because there will be the oddball here or there.” Another stated, “Three or four; there are two main physicians, the medical director, and one who works with her, and a few others have patients.” And another said, “I think we have five, three main and two stragglers.”

**Turnover and Retention.**—We also asked about staff turnover and retention. With each revision to the questionnaires, we made a series of refinements intended to inform us about the way in which NHAs and DONs retrieved information and responded to questions about turnover and retention. In an effort to calculate turnover and retention rates, we had hoped to collect data on the exact number of employees in several categories, as well as the number who had left these positions in the last year and the number who had remained in their positions for at least one year. However, during the in-person interviews, we observed respondents looking up information (e.g., counting from the schedule), and we timed how long this took (3-5 min). Because the need to look up information when responding to surveys has been shown to decrease response rates, we then guided respondents not to look up information. This resulted in respondents saying they “had to guess.” Responses included, “maybe 15, I’m guessing,” “I’ll guess around 25,” “Less than 5,” “maybe 60,” and “probably 20.”

We next asked respondents about their turnover more generally, including how they gathered that information (if they did) and what it was used for. We asked about turnover and retention separately for each category of staff: RNs, LPNs, and CNAs. We encountered a number of differing interpretations to these questions, both in the meaning of the term turnover and in deciding who should be counted in the response. One respondent showed the difficulty in interpretation of turnover by asking “does that include fired.” Several of the respondents could not categorize turnover by staff type, having only summary data for all employees, or for all nursing employees. Some of the respondents used a formula based on positions open and filled, rather than individual separations (e.g., “if we have 100 positions and 30 of those people have come in the last year, we say we retained 70% and our turnover is 30%, even if some of the positions were filled more than once”).

We next offered respondents a choice of response categories for the turnover questions. Some participants responded to the categories, and some directly responded with a numerical estimate. We noted the answers as given, because part of our purpose was to identify how respondents would answer the questions. This produced variations in the range of responses. For example, when asked for the overall “yearly turnover rate,” responses ranged from 14% to 20%. However, when we asked about the percentage of RNs, LPNs, and CNAs who had left in the last 12 months, the responses ranged from 0 to 40% for RNs, from 0 to 60% for LPNs, and from “less than 5%” to “more than 50%” for CNAs.

In the NHAs’ responses about retention, we again encountered differing interpretations of the term “retention” as well as different ideas about which staff should be counted. For example, in responding to the questions about the number of
RNs, LPNs, and CNAs who had worked in the NH for at least 12 months, some NHAs gave answers in numbers, some in percentages, and some in general phrases, such as “almost all.” Similarly, in telling us the average length of time that staff members have worked in the facility, the NHAs gave a range of estimates for RNs (e.g., “8–9 years”; “more than 5 years”; “probably over a year”). They also judged the LPNs’ length of service based on that of RNs (“what did I say for the RNs? For LPNs, longer . . .”). Some NHAs told us they had to guess, and two said they did not know.

Based on this information, the research team decided that reliable and valid exact data on turnover and retention, which could be transformed into continuous variables, could not be successfully gathered. We decided to gather turnover and retention data using questions with categorical answers. For example:

Please think about the NURSING ASSISTANTS who were employed at any time during the PAST 12 MONTHS. About what percent of these Nursing Assistants left your employment in the LAST 12 MONTHS? 0 to 20%; 21 to 40%; 41 to 60%; 61 to 90%; 91 to 100%.

Discussion

There is growing emphasis on evidence-based policy and practice at federal, state, and clinical levels. With increased availability of data to provide such evidence, including information from research and administrative databases, it is important to have a clear and commonly shared understanding of how respondents interpret concepts and report data and a clear understanding of the validity of data reported. This study illustrates some reasons why cognitive-based interviewing should be considered as part of the survey development process and that research results based on surveys that have not been cognitively tested should be approached with caution. Our findings demonstrate that not only it is important to verify that respondents understand survey questions as intended, but also it is important to determine if respondents possess the information needed to validly answer questions.

Research Implications

The work reported in this paper indicates that terms commonly used in long-term care may not have shared meanings among respondents to research questionnaires or administrative data collection instruments. We found differences in interpretation of long-term care concepts, and even seemingly simple terms that have been operationalized in research instruments and administrative data reporting formats. These differences have several potentially negative consequences. First, differences in interpretation of concepts and related questions may contribute to the observed discrepancies in the scientific literature. For example, our findings highlight the difficulty in collecting valid and reliable data on staff turnover and retention and could partially explain the widely varied rates reported in previous studies (Castle, 2006). As our results show, most NHAs do not have data on number of staff or turnover at their fingertips and their information management systems vary widely. Whereas some facilities had systems that tracked turnover by staffing category, most only had aggregated information for all employees or all nursing employees. Even the most carefully defined and painstakingly operationalized concepts cannot overcome issues of validity when survey respondents simply do not possess the necessary information.

Through cognitive-based testing, we determined that trade-offs had to be made between the level of specificity we wanted from respondents and the amount of time we could devote to any group of questions. Each question that required NHAs or DONs to look up information on computers or in written files allowed less time for additional questions and also risked having respondents abandon the survey altogether, resulting in reduced response rates. This is an important concern for any survey addressing this or similar topics.

We also found that many concepts explored in this study are dynamic and some, such as “culture change,” have an aspect of social desirability, resulting in reports of higher levels of culture change than supported by more detailed questions. It is, therefore, critical that we clearly define terms for respondents and not assume that researchers and providers have the same understanding of even commonly used terms. For example, an unexpected finding was the range of interpretations of the term direct-care staff, the meaning of which we initially considered to be well understood and shared.

Further, the industry and long-term care researchers should be striving for a more collective understanding of these concepts because until a consensus is reached on defining emerging concepts, like culture change and palliative care, the
validity of the data and its reliability across studies may be compromised. More widely agreed upon operational definitions of these terms would go a long way toward alleviating some of the difficulties uncovered in this study. In addition, it would be helpful to train new long-term care researchers and other evaluators in conducting cognitive interviews as the emergence of new terms and concepts is sure to continue as the field evolves.

Policy Implications

Based on our findings, policy makers should also use caution when interpreting research results from NH studies. Differences in interpretation of concepts between researchers, respondents, and policy makers may result in mistaken understanding of the meaning of data and result in inappropriate policy or reimbursement decisions. For example, a state may be using research results in deciding whether to implement a reimbursement policy of incentives for additional direct-care staff in NHs. For these policy implementations, a clear and mutual understanding of the meaning of the term direct-care staff is crucial.

Attributing quality outcomes to processes, staffing levels, or interventions also requires clear understanding of the interpretation and meaning of these concepts. For example, there has been a good deal of attention paid to culture change practices in NHs in recent years. Determining whether culture change is associated with improved quality of care or resident satisfaction is highly dependent upon the interpretation of the concept of culture change among responding NHAs and/or staff. Just considering the physical environment, the range of interpretations observed in this project included full-scale construction efforts aimed at remaking the lived space to very modest changes in décor. Several states are currently considering altering reimbursement, bed construction moratoria, and other long-term care policies based on the culture change concept, making the shared understanding of culture change by researchers, providers, and policy makers all the more important.

Similar to previous research (Miller & Shield, 2007), we found NHAs and DONs could readily discuss hospice care but had varying interpretations of palliative care and palliative care expertise. Because of this, we ultimately could not develop questions about the presence of palliative care programs or staffing. Considering this experience, we believe results of surveys asking NH staff about the presence of palliative care programs should be viewed cautiously. Policy recommendations for supplemental NH reimbursement for provision of palliative care have surfaced (Huscamp, Stevenson, Chernew, & Newhouse, 2010), but concerns remain about the readiness of NHs to provide palliative care (Meier, Lim, & Carlson, 2010). When attempting to assess such readiness and/or considering supplemental payments for palliative care provision, it is critical that the meaning and scope of palliative care be consistently understood. Without this understanding, NHs may be credited with and/or paid for providing more comprehensive palliative care than is the case.

Limitations

Cognitive-based testing is a labor intensive undertaking, and some research teams may not have the personnel or budget to support it at the scale reported herein. However, performing at least some cognitive-based testing is better than none because although standard pilot testing may help identify some problems with survey construction such as double-barreled questions or missing response options, only cognitive-based testing can identify problematic assumptions made by the researcher, such as an assumed understanding of terms and concepts or the assumption that potential respondents possess the information sought. Conducting testing locally or by telephone, as we have done, may help reduce costs.

In addition, cognitive-based testing takes several forms. We have utilized the focused-probe method; however, others prefer the think-aloud method. Debate continues as to the optimal procedures and protocols (Willis, 2004). Training new long-term care researchers in cognitive testing techniques would also help them determine the methods most appropriate for their survey development.

Finally, as we have noted, several of the terms and concepts our questionnaires focused on are newly emerging and lack consensus even among experts. It is not surprising, therefore, that NH professionals did not agree on definitions. In some cases, we were forced to eliminate questions because uniformly understood questions could not be developed. In these cases, cognitive-based testing was useful in identifying these problematic questions, but could not overcome this lack of consensus. We were, therefore, forced to acknowledge that valid survey data could not be collected on these topics.
Conclusions

As our cognitive-based testing demonstrates, much work must be done to understand the interpretation of concepts, structure questions to be as clear as possible, and minimize the range of interpretations. It is also important to determine if potential survey respondents possess the information necessary to provide valid responses. It is troubling to note that many of the questions used in structuring our initial questionnaires were drawn from other large scale, widely publicized and cited studies, and yet these questions posed a variety of problems for respondents. The findings of this study underscore the importance of systematic cognitive-based testing of interview and survey instruments, to assure that researchers and respondents share common interpretations of key concepts. Moreover, this strategy helps to ensure that the content of the question the researcher believes he or she is asking is consistent with the interpretation of the question by the respondent and that this interpretation is stable across individual respondents. Fidelity of interpretation of core concepts and terms needs to be ensured. Without such safeguards and rigor, the validity of NH research implications may be in considerable jeopardy.

Funding

This research was made possible in part by the Shaping Long Term Care in American Project funded by the National Institute on Aging (P01AG027296).

References


