

Measuring quality in elderly care: possibilities and limitations of the vignette method

Bozena Poksinska and Peter Cronemyr

The self-archived version of this journal article is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-139665>

N.B.: When citing this work, cite the original publication.

This is an electronic version of an article published in:

Poksinska, B., Cronemyr, P., (2017), Measuring quality in elderly care: possibilities and limitations of the vignette method, *Total quality management and business excellence (Online)*, 28(9-10), 1194-1207.
<https://doi.org/10.1080/14783363.2017.1303875>

Original publication available at:

<https://doi.org/10.1080/14783363.2017.1303875>

Copyright: Taylor & Francis (Routledge): SSH Titles

<http://www.routledge.com/>



Measuring quality in elderly care: possibilities and limitations of the vignette method

*Bozena Bonnie Poksinska and Peter Cronemyr
Logistics and Quality Management, Linköping University, Sweden*

Abstract

Listening to citizens is seen as an important source of information about public service performance. In Sweden to secure the quality of elderly care the National Board of Health and Welfare (NBHW) conducts a yearly survey of in-home elderly care services and nursing homes. A central problem of the existing survey methodology is the interpersonal incomparability of survey responses due to differences in preferences and health conditions. One way to deal with this problem is to use the survey methodology with anchoring vignettes. The purpose of the paper is to investigate the possibilities and limitations of using anchoring vignettes as a general survey method and specifically to test the method for measuring elderly care quality. The vignettes were developed interactively with professionals working in elderly care and evaluated with 3,994 users of in-home elderly care services and nursing homes. The results showed that anchoring vignettes reduce the impact of respondents' personal characteristics on survey results. In general, anchoring vignettes give more robust answers that reduce the problem of incomparability. However, anchoring vignettes increase the complexity of the questionnaire and have limited value for elderly care. Our results indicate that the method might be applicable when using healthier and younger respondents.

Keywords – user and customer surveys, anchoring vignettes, quality of public services, elderly care.

Corresponding author. Email: bonnie.poksinska@liu.se

Introduction

Improving the quality and efficiency of public services has become a priority for the Swedish government, as well as for other welfare state governments (Cheng et al., 2005, Christensen, 2016). Public services that meet the citizens' needs in relation to employment, health, housing, education, social security and care are considered crucial for socio-economic development (Pillinger, 2001). Following the reforms anticipated by the proponents of New Public Management, there has been increasing pressure to assess the quality and efficiency of public services, especially from the users' perspective (Carvalho and Brito, 2012). National user surveys are frequently used as instrument to measure the performance of public services despite questions about their value (Bovaird and Löffler, 2003). Considerable resources are spent on conducting national user surveys, but they actually do not lead to the improvement of public services (Reeves and Seccombe, 2008).

In Sweden, the National Board of Health and Welfare (NBHW) conducts a yearly national survey of nursing homes and in-home elderly care services. The survey results are used to evaluate and control the elderly care providers, but the evaluation may include a bias when comparing across providers. The problem is that the answers given by different individuals are not always comparable since the questions may be interpreted and response scales used differently. For example, individuals may have their own definition of topics such as staff responsiveness, as well as a propensity to respond using only extreme or moderate scale categories (Rice et al., 2010). Though this is not a problem if the results are evaluated from the single respondent's perspective since a person's perception of the situation is his/her "truth", it appears when self-reported data are compared across institutional settings and population groups.

King et al. (2004) proposed using anchoring vignettes to improve the comparability of survey responses. The method works as a calibration tool to adjust the self-reported data to an independent standard. The purpose of the paper is to investigate and discuss the possibilities and limitations of using anchoring vignettes as a general survey method and specifically to test the method for measuring the quality in elderly care.

Theoretical framework

The vignettes were introduced into research in the 1970s (Alexander and Becker, 1978; Nosanchuk, 1972). Hughes (1998) defined vignettes as "*stories about individuals, situations and structures which can make reference to important points in the study of perceptions, beliefs and attitudes*". Vignettes have been widely used in social sciences to inform about a situational context and influential research topic variables (Hughes and Huby, 2012). Vignettes are considered useful for studying sensitive or subjective topics. The interviewees are invited to respond to hypothetical situations and thereby elucidate their own experience on the question under investigation (Barter and Renold, 1999).

This study deals with the application of vignettes to survey research. The method of anchoring vignettes was proposed by King et al. (2004) to improve the interpersonal comparability of survey responses when respondents understand questions in completely different ways, or differ in the manner in which they use response scale categories. The problem with different use of response scales is illustrated in Figure 1. A common assumption is that the response categories are of equal distance on a measurement scale, but the respondents may use them differently by, for example, actively avoiding using middle and low values. In the anchoring vignette method the traditional self-assessment questions are accompanied by vignettes describing different levels of the measured variable. In the example presented in Figure 1, respondents rate two vignettes along with a self-assessment question using the same scale. There are two key measurement assumptions associated with the vignette method. First, response consistency assumes that each respondent uses the self-

assessment and vignette categories in approximately the same way across questions. Second, vignette equivalence assumes that “the actual level of the variable represented in any one vignette is perceived by all respondents in the same way and on the same unidimensional scale, apart from random measurement error” (King et al., 2004, p. 194).

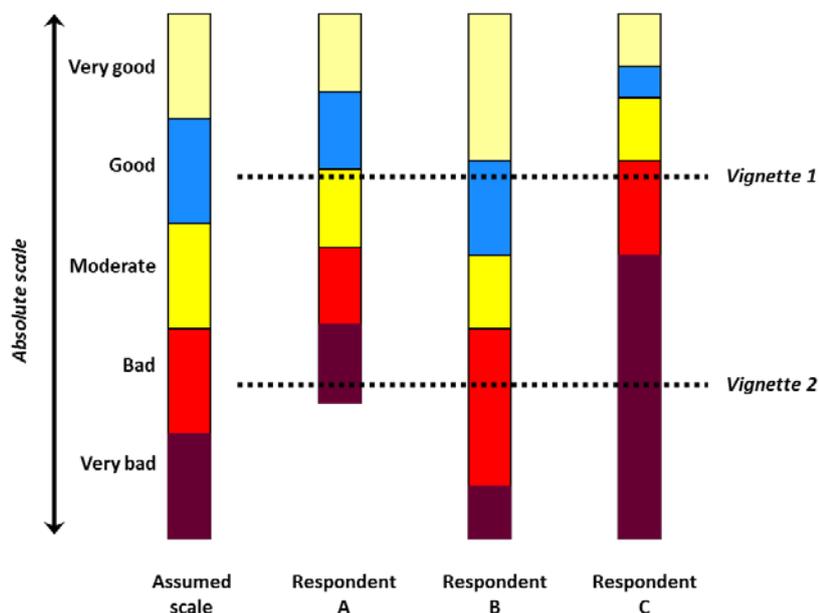


Figure 1: Use of response scales and vignette rating.

Research points out both advantages and concerns with anchoring vignettes. Results from survey validation studies suggest that vignettes may be sensitive to many factors that are outside the researchers’ control and that the method may not work as intended (d’Uva et al., 2011; Gupta et al., 2010; Jürges and Winter, 2013; Van Soest et al., 2007). The primary problem is the limited knowledge of the aspects associated with the vignette’s design, which may influence the respondents’ rating. Jürges and Winter (2013) provided evidence that respondents may rate vignettes differently depending on whether the persons described therein are women or men or are associated with a certain age. Van Soest et al.’s 2007 study of student drinking habits presented similar results and found that vignettes are rated differently depending on whether they contained a female or a male name (Van Soest et al., 2007). Further, there is also lack of knowledge of how vignettes are perceived. Though anchoring vignettes were used in several studies targeting the elderly, none of the studies investigated how well the method works for this respondent group (Bonsang and Van Soest, 2012; d’Uva et al., 2011). The research on questionnaire surveys in general shows that surveys involving elderly respondents have lower respondent accuracy. Elderly have more difficulties in interpreting and answering survey questions because of changes in memory associated with aging and lower motivation to participate in surveys (Enagonio, 2006; Taylor-Davis and Smiciklas-Wright, 1993).

Problems associated with measuring elderly care quality

Subjectivity characterizes elderly care quality measurement, which undermines the development of a representative set of quality indicators and suitable measurement instruments (Zimmerman, 2003). The study by Bauld et al. (2000) showed that health status affects how satisfied elderly people are with care services. Elderly people who perceive their health as poor are less likely to be satisfied with the services and vice versa. Another problem is that users of elderly care services usually receive them as an entitlement. They may be afraid of being perceived as demanding and may minimize their expectations and needs. This may also reduce their willingness to assess or criticize quality (Barnes, 1999). Other factors affecting satisfaction ratings in elderly care are reluctance to criticize individual workers and

fear of dependency or reprisal (Bauld et al., 2000). This creates a problem in service provider comparisons because it is not possible to see whether the differences in ratings depend on the individual provider's quality or on, for example, respondent health status, preferences, or different use of the response scale. In other words, comparing self-reported data across institutional settings may have serious consequences for single service providers, but could be based on incorrect or biased information.

Methodology

The project was initiated by the Ministry of Health and Social Affairs and commissioned by the Swedish Agency for Health and Care Services Analysis. The overall aim of the research was to investigate how the use of anchoring vignettes could increase the comparability of the results from the national survey on elderly care.

General model for vignette method

The research started by developing a general model for the use of vignettes in measuring elderly care quality. The model is presented in Figure 2 and described below.

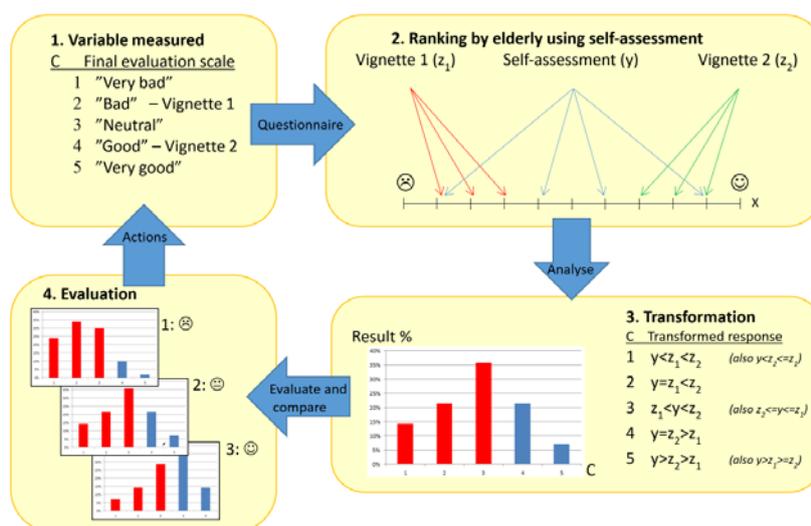


Figure 2: The general model for vignette method

Step 1: Five questions from the survey on elderly care have been chosen to test and evaluate the vignette method. These questions were considered problematic due to interpersonal incomparability and covered the following aspects: overall health, anxiety, mobility, staff responsiveness, and violation of personal integrity. The first three aspects refer to the elderly's perception of their health conditions, while the last two aspects refer to their opinions of being treated with respect and dignity. The final evaluation scale was the 5-point Likert scale, which is used in the original survey. Since one important requirement was to make the questionnaire as simple as possible, two vignettes for each of the five survey questions were enough to adjust the self-assessment questions. One vignette was intended to describe the situation for the measured variable as "rather bad", the other as "rather good".

Step 2: Respondents rate the two vignettes and make a self-assessment on the same absolute scale (X). The questionnaire consisted of five topics with one page per topic (see example in Figure 3). Each page included two vignettes and one self-assessment. The respondents rated each vignette and made self-assessments by putting an X on the scale between the ends saying "Very bad" and "Excellent" with smileys. Our experience on response scales suggested that the data transformation works better if continuous scale is used, that is, respondents do not need to select an integer, but can freely put the cross on the scale.

Overall health

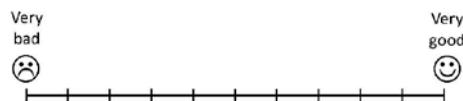
1. Anna is 75 years old. She has a chronic disease. She often feels sick and has a stomach ache. She rests in her bed for several hours per day. Anna often feels powerless and downhearted. She very seldom wants to have visitors since she is unsure whether she is going to cope.

How do you assess Anna's overall health?
Make a cross on the scale below.



2. Olle is 70 years old. He likes to take a walk in the forest and pick mushrooms. Unfortunately he cannot do it so often because he gets a backache. The ache normally ceases the following day. On general Olle is feeling good and can manage his daily life on his own at home.

How do you assess Olle's overall health?
Make a cross on the scale below.



3. How do you assess your overall health??
Make a cross on the scale below.

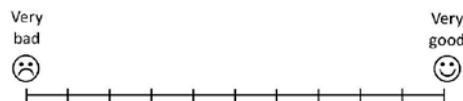


Figure 3: One page from the questionnaire (translated from Swedish)

Step 3: As recommended by King et al. (2004), the data are transformed into a relative response (C) by comparing the values for the self-assessment (y) with values for the vignettes (z_1 and z_2) according to the following rules:

- C=1 if $y < z_1 < z_2$ (also $y < z_2 \leq z_1$)
- C=2 if $y = z_1 < z_2$
- C=3 if $z_1 < y < z_2$ (also $z_2 \leq y \leq z_1$)
- C=4 if $y = z_2 > z_1$
- C=5 if $y > z_2 > z_1$ (also $y > z_1 \geq z_2$)

The transformed data can, for example, be presented in a histogram.

Step 4: Using the transformed data, comparisons can be made either over time or between different elderly care providers.

Development and test of the vignette method

Figure 4 illustrates the process of developing, testing and evaluating the vignette method for measuring quality in elderly care.

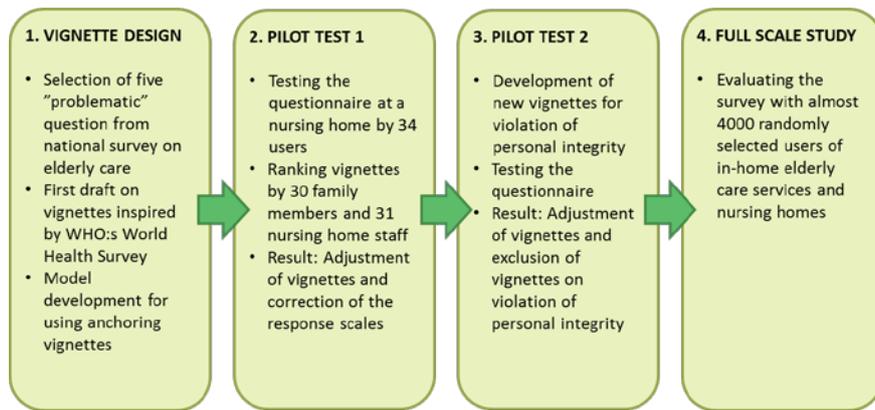


Figure 4: Process of developing, testing and evaluating the vignette method

The development of vignettes is an iterative approach that implies several redesigns and revisions. The first vignette is supposed to be “rather good”, while the second is supposed to be “rather bad”. The vignettes have been developed iteratively so it should be possible to put the answer to the self-grading question either below both vignettes, above both vignettes, close to one vignette, or between the vignettes. Hence the vignettes should not be “too bad” or “too good” and that is why several redesigns are required. The vignettes are considered good enough if the pass rate is at least 90 percent; that is, respondents ranked the vignettes in the intended order ($z_2 > z_1$).

The design of the initial vignettes was inspired by the World Health Survey (WHO, 2002–2004) and developed in collaboration with professionals working in elderly care. After several redesigns, two pilot tests at a nursing home were carried out. The development of personal integrity violation vignettes has been shown to be problematic, since people have very different thresholds to what they regard as violation of personal integrity. The topic is also difficult to describe by a vignette since gross violation depends not only on what is done or said (relatively easy to describe by a vignette) but to large extent on how insulting words are said, body language, facial expressions, and the relationship between staff and elderly person. After the reliability test consistently failed, the topic was excluded from the full-scale study.

The full-scale study was performed in autumn 2014. The questionnaire was sent to 3,994 randomly selected users of in-home elderly care services and nursing homes. In total, 1,600 respondents answered the questionnaire, producing a 43 percent response rate (corrected response rate after excluding undeliverable questionnaires). As a comparison, the response rate for the 2014 national elderly care survey was 63 percent. Not all the respondents answered all the questions. Hence some tables below say, for example, “Frequency=1,510”, meaning the number of answers for a specific question.

Findings

Topic: 'Overall Health'

Figure 5 shows the analysis of the answers vis-à-vis 'Overall Health', with 96 percent of respondents ranking the vignettes in the intended order, that is, the "low" vignette was rated lower than the "high" vignette.

In the left-hand graph, elderly people's self-assessments are presented on an absolute scale (X) together with the transformed responses (C) using the formulas presented in the general model (see Figure 2). The right-hand graph shows how the data were transformed. Bars placed on the main diagonal (indicated by a line) show responses that have not been changed by vignette transformation ($C=X$ is 56 percent). Bars closer to the viewer indicate responses that have been increased ($C>X$ is 40 percent), while bars further away indicate responses that have been decreased ($C<X$ is 14 percent). The mean value for the absolute responses ($X=2.81$) is significantly lower than that for the relative responses ($C=3.11$), which indicates that data transformation produces a more positive view of the elderly health status.

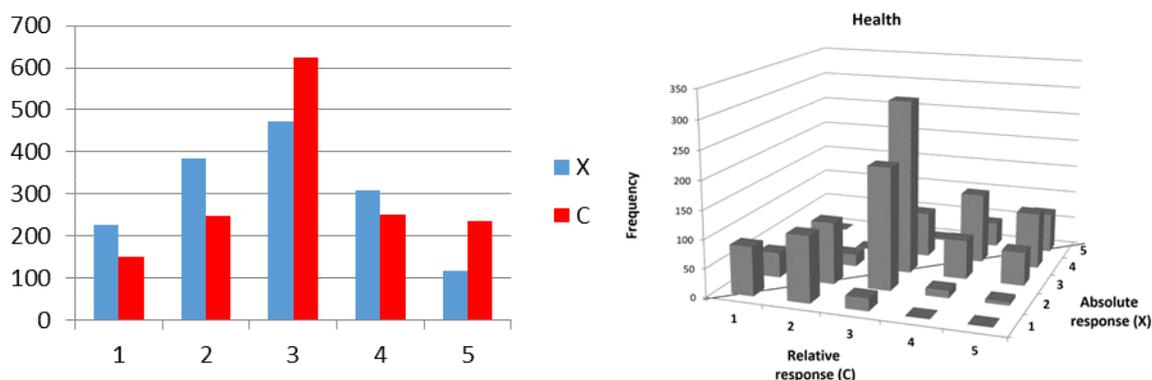
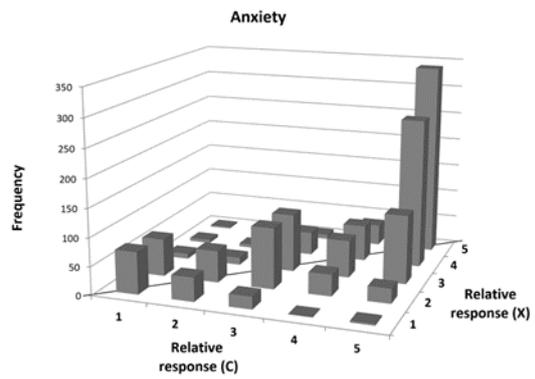
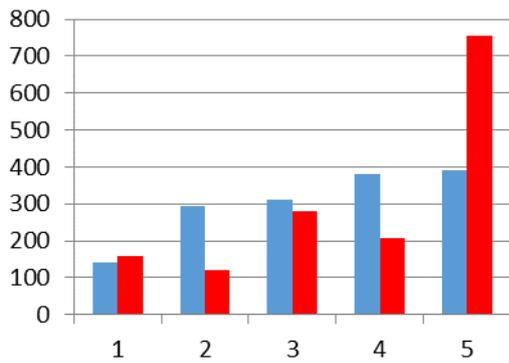


Figure 5: Overall health: absolute (X) and relative (C) responses

Topic: 'Anxiety'

The second topic is 'Anxiety', which is one of the three topics related to health and measures the extent to which elderly have trouble with anxiety, worry or fear. The pass rate for the vignettes is 86 percent, which is below the expected standard. The results have shown that respondents rated both the "low" and "high" vignette relatively low, which means that the relative responses are likely to end up higher after the transformation.

Figure 5 shows that 42 percent of responses are not changed after the transformation ($C=X$), 45 percent of absolute responses are transformed into higher values ($C>X$), and 13 percent to lower values. In particular, respondents that rated themselves as low ($x = 2, 3$ or 4) received high relative values ($C = 5$), suggesting that elderly people feeling anxiety actually had fewer problems than those described in the "fairly good" vignette. This is also shown by the mean value, which is significantly lower for the absolute responses ($X = 3.39$) than for the relative responses ($C = 3.84$). However, the results must be interpreted with caution since both vignettes are rated relatively low.



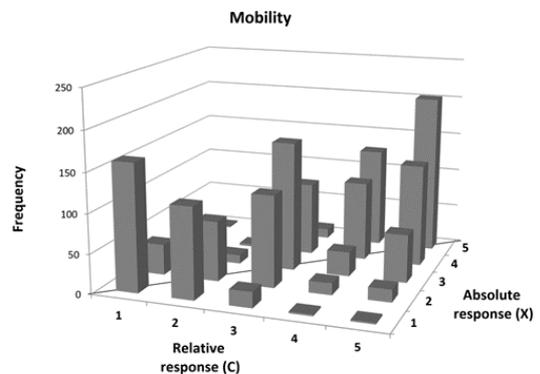
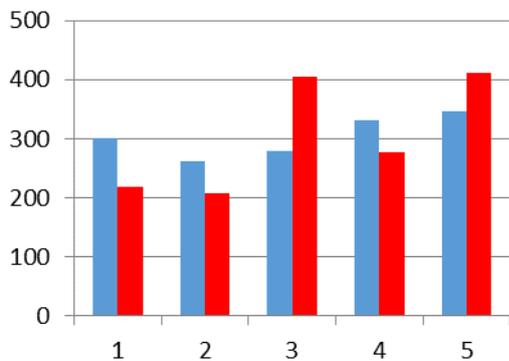
Anxiety	Frequency	Mean value	+/-	Proportion C<X	Proportion C>X
<i>Absolute response (X)</i>	1,519	3.39	0.03		
<i>Relative response (C)</i>	1,519	3.84	0.04	13 %	45 %

Figure 6: Anxiety: absolute (X) and relative (C) responses

Topic: 'Mobility'

The third health-related topic is 'Mobility', concerning the ability to move around indoors, with or without help from others. The pass rate is 92 percent.

Figure 7 shows that many respondents' self-assessed mobility (47 percent) is located on the main diagonal ($X = C$), 33 percent of absolute values are transformed to higher values, and 20 percent to lower values. The mean value of the absolute responses ($X = 3.10$) is significantly lower than the average value of the relative responses ($C = 3.30$), but the dispersion is high. The results indicate that the interpersonal differences are high for the mobility question and using vignettes makes the results more comparable (53 percent of absolute responses were transformed to lower and higher relative values). However, it cannot be stated that respondents on average have better (or worse) mobility than vignettes because there is no clear trend in which direction the absolute responses were transformed.



Mobility	Frequency	Mean value	+/-	Proportion C<X	Proportion C>X
<i>Absolute response (X)</i>	1,518	3.10	0.04		
<i>Relative response (C)</i>	1,518	3.30	0.04	20 %	33 %

Figure 7: Mobility: absolute (X) and relative (C) responses

Topic: 'Staff responsiveness'

Staff responsiveness measures the experience of being treated with dignity and respect. The pass rate for vignettes is 93 percent. It should be noted that the high vignette was sometimes rated a little higher than intended, and the low vignette a little lower than intended. In other words, there is a risk that proportionally high number of self-assessment responses are between vignettes, that is, $C = 3$.

As shown in Figure 8, there is no significant difference in the mean value for the absolute responses ($X = 4.37$) and the relative responses ($C = 4.31$). In contrast with what was expected from the discussion above, there is very small percentage of absolute responses

transformed into C = 3. Instead, most self-assessments have values 4 and 5, both in absolute (X) and relative scale (C). This means that very few respondents said they had experienced poor responsiveness from the staff. This is a positive result, but it could be also a sign that the elderly are reluctant to complain about care, for example, because of their dependence. The latter interpretation is supported by the relatively large proportion (38 percent) of absolute responses X = 5 (“always good responsiveness”) being transformed to C = 4, which corresponds to a “pretty good” vignette.

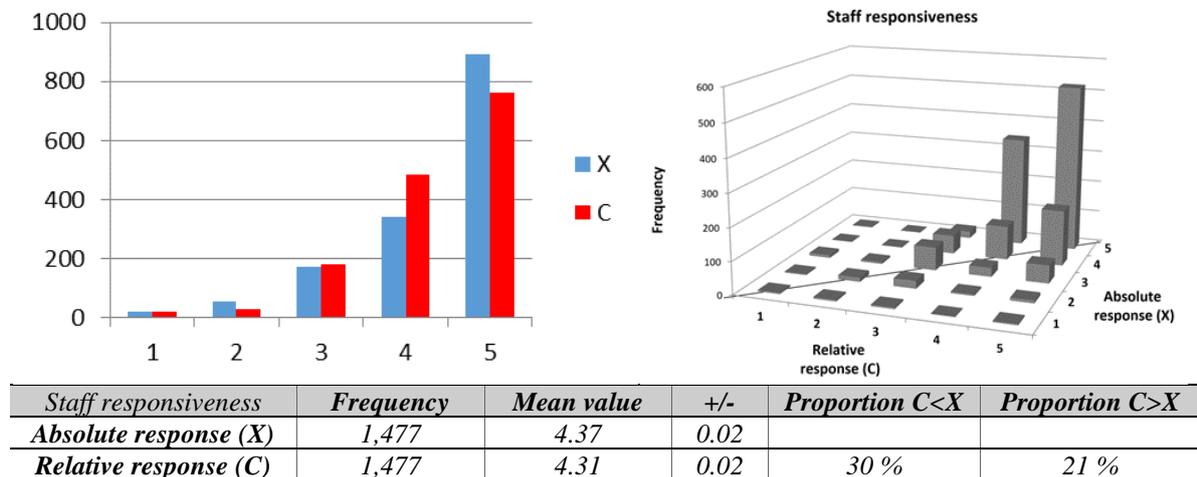


Figure 8: Staff responsiveness: absolute (X) and relative (C) responses 8

Relationship between staff responsiveness and health variables

Previous studies have shown that there is relationship between elderly health and the perceived quality of care (Bauld et al., 2000). Figure 9 shows respondents’ answers according to the 10-point scale (0–10), where staff responsiveness (y-axis) is shown as a function of the respective health variables: overall health, anxiety, and mobility (x-axis).

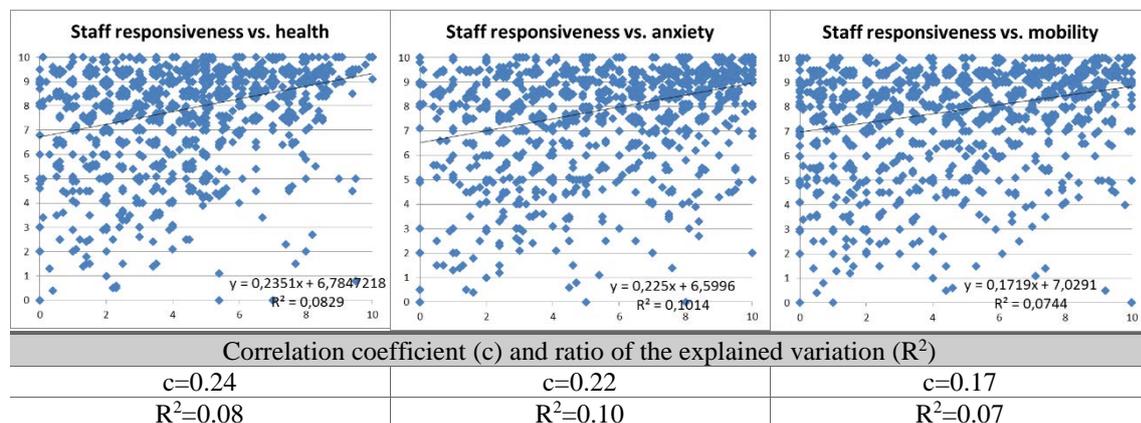


Figure 9: Staff responsiveness as a function of health variables using self-assessment responses (0–10)

Elderly respondents in various degrees of health rate the staff responsiveness as good (grades 8–10), but a perception of good staff responsiveness is even more common for those who also have good health (grades 8–10). The linear regression line inclination for staff responsiveness and overall health shows that the correlation coefficient between them is 0.24, which means that if the value for health increases by 1, the value for staff responsiveness increases on average by 0.24. A similar relationship applies for mobility and anxiety, where staff responsiveness increases by 0.22 and 0.17 respectively. The higher the self-assessed health, the higher the staff responsiveness rating, which is in line with previous research (Bauld et al., 2000).

It should be noted that the dispersion around the linear regression line for all health variables is large (the ratio of the explained variation (R^2) is very low), which means that the correlation cannot be used as a prediction model. The staff responsiveness cannot be predicted as a function of health variables since the responsiveness rating largely depends on other, unknown variables than health. However, the results indicate that care providers with less well elderly can be expected to have slightly lower staff responsiveness ratings than providers with healthier elderly respondents.

Whether the positive relationship is influenced by the use of vignettes can be examined by analysing and comparing the staff responsiveness as a function of health variables using the absolute and relative scale. The results presented in Table 1 show that the correlation does not disappear completely, but the coefficients are lower when using vignettes. The correlation in Table 1 is as expected similar for the self-assessment (0–10) and the absolute responses (X), since the values are just a “translation” of the self-assessment from a continuous scale (0–10) to a five-point Likert scale.

Correlation coefficient (c) and ratio of the explained variation (R^2)	Responsiveness vs. Health	Responsiveness vs. Anxiety	Responsiveness vs. Mobility
<i>Self-assessment (0–10)</i> <i>[see also figure 9]</i>	$c = 0.24$ $R^2=0.08$	$c = 0.22$ $R^2=0.10$	$c = 0.17$ $R^2=0.07$
<i>Absolute responses (X=1–5)</i>	$c = 0.22$ $R^2=0.07$	$c = 0.21$ $R^2=0.08$	$c = 0.17$ $R^2=0.07$
<i>Relative responses (C=1–5)</i>	$c = 0.14$ $R^2=0.03$	$c = 0.08$ $R^2=0.02$	$c = 0.07$ $R^2=0.01$

Table 1: Staff responsiveness as a function of health variables using different scales

Respondents’ reaction on the vignette method

The full-scale study aggregated a total of 631 comments, with the categorization presented in Table 2. The majority of comments (80 percent) concerned different aspects of elderly care and elderly people’s life situation. About 6.3 percent of comments imply that the questionnaire couldn’t be answered because of health, which is important reason of the nonresponse. The focus of this section is analysis of the comments on the vignette method.

Category	Count	Share %
Comments on different aspects of elderly care and elderly people’s life situation	507	80.3
Negative comments on the vignette method	52	8.2
Information that the questionnaire couldn’t be answered because of health condition	40	6.3
General comments on the questionnaire and improvement suggestions	28	4.4
Positive comments on the vignette method	4	0.6

Table 2: Categorization of comments from the full-scale study

There are three general issues:

1. Vignette method is more demanding

A major drawback of the vignette method is that the questionnaire becomes more complex and demanding. This is a strong argument against the use of the vignette method in elderly care since many potential respondents are too sick to answer the current traditional survey. In our study, 216 respondents, or 5 percent of all posted questionnaires, declined the survey with reference to health aspects such as dementia, Alzheimer's, or visual impairment.

Additionally, 20 percent of all questionnaires were answered by a relative or friend, and 30 percent were answered with help of a relative or friend. The comments indicated that the vignette method might not be suitable for elderly. Some examples follow below:

- *“Unfortunately, I can't communicate with my mother in a way that we can answer this survey”*
- *“Meaningless survey. Mom is 100 years old, blind and has dementia!”*
- *“Because of memory problems hard to consider the vignettes. Sincerely Daughter”*

2. Vignette method is perceived as strange

Surveys are a common method for measuring people's perceptions, opinions, and attitudes, and most people have answered many questionnaires. Respondents therefore have certain expectations of how questionnaires are designed. Since the vignette method is relatively uncommon, about 30 comments indicated that the purpose with vignettes is not understood. Some examples follow:

- *“Ridiculous survey, with too much text, if I get something similar in the future I do not care about it!”*
- *“It was a little confusing with the case descriptions. Next time just ask myself of how I feel about the in-home services. Thanks”*
- *“The survey was very difficult to answer. The first time such a strange structure”*
- *“Messy with the short stories. Why???”*
- *“It seems to be Reading Comprehension Test”*

3. The vignette method is perceived as condescending

Around 20 respondents commented that the questionnaire is perceived as condescending. These respondents believed that the survey was designed thusly because researchers think that the elderly have cognitive difficulties that undermine answering a traditional survey. Some examples of the comments follow:

- *“Not particularly intelligent questionnaire. Do you believe that we old are so senile that we can't respond to usual questions?”*
- *“The dumbest questions I've ever encountered!”*
- *“I don't think it's appropriate with 'smiles' in the survey. It's about old men with long life experience, not children. Feels degrading!”*

Conclusions

The method of anchoring vignettes showed a clear advantage for reducing the incomparability of survey results, but it has a limited value when used with elderly care service respondents.

Theoretical implications

The vignette method decreases the effect of interpersonal differences on the survey results. When transforming from absolute to relative answers, on average, about half of the responses were transformed into higher or lower values with the help of vignettes, which suggests that

there are big differences in how respondents interpret and answer questions. Thus, the vignette method improves the interpersonal comparability of the results.

One of the biggest comparative advantages of the vignette method is that the perception of quality in elderly care, for example, staff responsiveness, becomes less dependent on the respondents' health. Since the health status of the elderly may vary significantly between different types of elderly homes, as well as between investigations carried out at the same elderly home but on different occasions, the increased comparability of results is an important aspect of measuring provider quality. Unfortunately, it is not possible to develop a standardized formula that predicts the effect of respondents' health on the satisfaction with elderly care. The correlation between health and satisfaction varies between individuals, which means that the vignette method needs to be used for each individual respondent to have an effect.

However, the method still requires further development. The vignette method was perceived as strange and condescending by some respondents, possibly due to its unfamiliarity in Sweden. The purpose of vignette rating was in several cases misunderstood, creating discomfort and irritation among the respondents. Whether this is a question of method development or an intrinsic property of the vignettes is unfortunately not possible to conclude based on this study. However, based on the strengths of the method found in this paper, we believe that anchoring vignettes have a potential to be used among healthy and younger respondents. That has not yet been proven and therefore needs to be investigated in further research.

Practical implications

The result indicated that care providers with sicker elderly can be expected to have a slightly lower ratings of service quality than providers with healthier elderly. Despite the vignette method's clear advantages for comparing the results across providers, it nevertheless makes questionnaires more complex, mentally demanding, and time consuming. This is a strong argument against the use of the vignette method in elderly care, because many elderly people have already difficulty answering today's less complicated surveys.

References

- Alexander, C. S. & Becker, H. J. (1978) The use of vignettes in survey research. *Public Opinion Quarterly*, 42(1), pp. 93–104.
- Barnes, M. (1999) *Unequal partners: user groups and community care*, London, The Policy Press.
- Barter, C. & Renold, E. (1999) The use of vignettes in qualitative research. *Social Research Update*, 25(9), pp. 1–6.
- Bauld, L., Chesterman, J. & Judge, K. (2000) Measuring satisfaction with social care amongst older service users: issues from the literature. *Health & Social Care in the Community*, 8(5), pp. 316–324.
- Bonsang, E. & Van Soest, A. (2012) Satisfaction with social contacts of older Europeans. *Social Indicators Research*, 105(2), pp. 273–292.
- Bovaird, T. & Löffler, E. (2003) Evaluating the quality of public governance: indicators, models and methodologies. *International Review of Administrative Sciences*, 69(3), pp. 313–328.
- Carvalho, C. & Brito, C. (2012) Assessing Users' Perceptions on how to Improve Public Services Quality. *Public Management Review*, 14(4), pp. 451–472.
- Cheng, B.-W., Chang, C.-L. & Liu, I.-S. (2005) Enhancing care services quality of nursing homes using data mining. *Total Quality Management & Business Excellence*, 16(5), pp. 575–596.
- Christensen, J. M. (2016) Professional development in social work: elderly care in Sweden and Denmark. *Social Welfare: Interdisciplinary Approach* 1(6), pp. 62–76.

- D'uva, T. B., Lindeboom, M., O'donnell, O. & Van Doorslaer, E. (2011) Slipping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources*, 46(4), pp. 875–906.
- Enagonio, E. M. (2006) What does this question mean to you? Cognitive interviewing to pretest a questionnaire for older adults. Master thesis.
- Gupta, N. D., Kristensen, N. & Pozzoli, D. (2010) External validation of the use of vignettes in cross-country health studies. *Economic Modelling*, 27(4), pp. 854–865.
- Hughes, R. (1998) Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociology of Health & Illness*, 20(3), pp. 381–400.
- Hughes, R. & Huby, M. (2012) The construction and interpretation of vignettes in social research. *Social Work and Social Sciences Review*, 11(1), pp. 36–51.
- Jürges, H. & Winter, J. (2013) Are anchoring vignettes ratings sensitive to vignette age and sex? *Health Economics*, 22(1), pp. 1–13.
- King, G., Murray, C. J., Salomon, J. A. & Tandon, A. (2004) Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(01), pp. 191–207.
- Nosanchuk, T. A. (1972) The vignette as an experimental approach to the study of social status: An exploratory study. *Social Science Research*, 1(1), pp. 107–120.
- Pillinger, J. (2001) *Quality in social public services*, Office for Official Publications of the European Communities.
- Reeves, R. & Seccombe, I. (2008) Do patient surveys work? The influence of a national survey programme on local quality-improvement initiatives. *Quality and Safety in Health Care*, 17(6), pp. 437–441.
- Rice, N., Robone, S. & Smith, P. C. (2010) International comparison of public sector performance: the use of anchoring vignettes to adjust self-reported data. *Evaluation*, 16(1), pp. 81–101.
- Taylor-Davis, S. A. & Smiciklas-Wright, H. (1993) The quality of survey data obtained from elderly adults. *Journal of Nutrition for the Elderly*, 13(1), pp. 11–21.
- Van Soest, A., Delaney, L., Harmon, C. P., Kapteyn, A. & Smith, J. P. (2007) Validating the use of vignettes for subjective threshold scales. IZA Discussion Paper No. 2860.
- WHO (2002–2004) *World Health Survey (WHS)* [Online].
<http://apps.who.int/healthinfo/systems/surveydata/index.php/catalog/whs/about>.
[Accessed 22 May 2016].
- Zimmerman, D. R. (2003) Improving nursing home quality of care through outcomes data: the MDS quality indicators. *International Journal of Geriatric Psychiatry*, 18(3), pp. 250–257.