

RESEARCH PAPER

Interviewer effects in a survey examining pain intensity and pain interference in nursing home residents

PATRICK KUTSCHAR¹, JUERGEN OSTERBRINK¹, MARTIN WEICHBOLD²

¹Institute of Nursing Science and Practice, Paracelsus Medical University, A-5020 Salzburg, Austria

²Department of Sociology, Paris Lodron University Salzburg, A-5020 Salzburg, Austria

Address correspondence to: Patrick Kutschar, Institute of Nursing Science and Practice, Paracelsus Medical University, Strubergasse 21, A-5020 Salzburg, Austria. Tel: (+43) 662 80 344; Fax: (+43) 662 2420 80 009. Email: patrick.kutschar@pmu.ac.at

Abstract

Introduction: Face-to-face surveys are applied frequently when conducting research in older populations. Interviewers play a decisive role in data quality, may affect measurement and influence results. This study uses survey data about pain in nursing home residents and analyses, whether affiliation-of-interviewer (internal vs. external to nursing home) and gender-of-interviewer affect residents' responses in terms of interviewer variance and systematically varying pain reports.

Methods: Overall, 258 nursing home residents with up to moderate cognitive impairment were examined by 61 interviewers about pain intensity and interference applying the Brief Pain Inventory. Interviewer variance was measured using intra-interviewer correlation coefficients (ρ). Two-factorial covariance analysis was applied to analyse whether pain intensity and interference scores differ by interviewer characteristics.

Results: Interviewer heterogeneity accounts for almost one quarter of total variance on average. Interviewer variance is higher for internal and male interviewers than for external and female interviewers. Covariance analyses show significant effects of interviewer characteristics on pain reports. Average pain intensity and interference scores vary considerably by interviewer gender and affiliation. Highest pain intensity was reported towards female internal and male external interviewers; highest pain interference was reported towards male external interviewers.

Conclusion: Residents' answers substantially differ in relation to who is assessing pain. There is a risk of imprecise and biased survey estimates on sensitive topics like pain in nursing homes. Interviewer gender and affiliation seem to evoke gender-specific and status-related expectations and attributions which influence residents' response process. Interviewer effects pose a considerable threat to survey data quality in institutionalised older populations.

Keywords: interviewer effects, survey data quality, nursing home residents, pain reports, older population, older people

Key Points

- Interviewer effects are of special concern for survey data quality in nursing home residents.
- Interviewers have an influence on older respondents' response processes.
- Pain reports are at risk to be erroneous or biased due to interviewer effects.
- Gender- and affiliation-of-interviewer increase interviewer variance and elicit systematically varying pain estimates.
- Residents' gender- and status-related expectations are assumed to influence pain assessment outcomes in pain management practice.

Introduction

Standardised face-to-face interviews are frequently applied in age research to assess attitudes, health states or quality of life of older people [1–3] and are often indispensable to measure health-related outcomes in geriatric settings [4, 5]. Interviewing older persons poses several methodological challenges in terms of recruitment or measurement [1, 6–9]. This becomes especially evident relating to nursing home residents (NHRs), where cognitive impairment (CI), physical limitations or limited audio-visual abilities are rather common [10–12]. Given these population-based constraints, the question–answer processes [13, 14] might be challenging especially for older respondents and introduce specific threats to survey data quality [1, 14, 15]. Here, interviewers play a crucial role, may influence respondents' provided answers and are a decisive factor for the quality of the collected data [16, 17].

Interviewer effects refer to errors of measurement in terms of deviations between an observed and true value of the phenomena of interest and are attributed to certain interviewer characteristics (e.g. gender, age, behaviour) [17–20]. Two types of interviewer effects are distinguished [18, 21, 22]. Interviewer bias refers to a systematic under- or overestimation of the measures due to errors rooted in the survey design, and thus, leading to similarly biased measures across all interviewers. Interviewer variance refers to differences in the extent to which individual, or a set of, interviewers cause specific errors resulting in an increased variability of responses and reduced precision of estimates. The latter may also cause biased survey outcomes due to systematic differences between interviewers.

An extensive body of literature reports on various interviewer effects [19]. Although older respondents might be particularly susceptible to interviewer effects in health-related surveys, there is still rather limited empirical evidence about this population segment [2, 18, 23]. Early studies found more pronounced interviewer variance with increasing age of interviewer and in male interviewers [24]. Others suggested interviewer experience as a mediating factor for more valid data and higher reporting of psychological symptoms [25]. More recently, Winters *et al.* [23] found that NHR ratings varied by the interviewer's previous experience or general knowledge of health care, while neither gender nor age had a significant effect. Beullens *et al.* [26] found increased interviewer variance in older respondents and that more interviewer assistance was necessary. Other studies observed that older persons reported significantly more health problems towards expert rather than lay interviewers [27] and that study outcomes would have been significantly underestimated if only lay interviewer data had been used [28].

Rationale And Objectives

Interviewer effects are of special concern in surveys with NHRs, especially when sensitive topics are examined, such as pain reports. Pain poses a significant burden in daily living and represents a sensitive topic with extensive subjective relevance [29–31]. A part of a parental study examining pain management in nursing homes (NHs), present research explores the influence of observed interviewer characteristics, i.e. affiliation-of-interviewer (AOI, NH internal vs. external) and gender-of-interviewer (GOI, male vs. female), on pain outcomes. AOI and GOI might elicit role- or status-related perceptions and induce socially desirable responses, i.e. respondents match their responses to what they anticipate the interviewers want to hear [32], especially with regard to sensitive questions relating to cultural norms or expectations [14, 17, 19, 32]. For instance, familiarity between nurses and residents may enhance social desirability. NHRs might attribute internal nurses an 'expert'-status while stereotyping external interviewers as 'lays', eventually affecting reports of health-related information [27, 28]. Gender stereotypes are widely present in Western societies [33] and presumably more pronounced in older populations. Health care institutions, such as NHs, tend to be shaped by hegemonic masculinity that affects interactions, gender constructions and role-related expectations [34, 35]. Such expectations and stereotypes form pain perception and communication, and thereby, might structure the respondents' editing of pain-related answers [35].

The objectives of this observational study were to examine interviewer effects on the reporting of pain intensity and interference based on self-reports of NHRs and to analyse whether AOI and GOI affect NHR responses in terms of (i) interviewer variance and (ii) biased pain outcomes.

Methods

Survey data come from a registered cluster-randomised controlled trial (c-RCT; UTN: U111-1187-3174) in 15 German NHs, where effects of pain management nursing education interventions on NHR pain situation were evaluated in a sample of 879 NHRs with no, mild, moderate and severe CI. Details and results from the c-RCT have been published elsewhere [36].

Participants

Eligible participants had to be 60 years and older, living in the participating NH, having sufficient German language abilities and providing informed consent. To estimate the ability to self-report in the survey situation, NHRs were initially screened using the Mini-Mental State Examination (MMSE) [37]. NHRs with up to moderate CI (MMSE score 10–30) were interviewed applying standardised interviews ($n = 598$). We included only those

residents who self-reported presence of pain, since their pain intensity and interference were assessed ($n = 258$) (Supplementary Table S1).

Data collection

Interviewers were health care professionals (e.g. social work, psychiatric nursing), students of nursing or health sciences (i.e. external to the NH) and nurses from the participating NH (i.e. internal to the NH). All interviewers received the same full-day comprehensive interviewer training (e.g. theory, exercises regarding interviewing older persons, procedures, interviewer behaviour) and were supervised during data collection. Interviewers were pragmatically assigned to NHR and study location—i.e. external interviewers were deployed in several NH, and internal nurses in their respective NH. Standardised, tablet-assisted, face-to-face interviewing was applied. Survey questions and answer categories were read aloud and additionally presented as paper versions.

Variables, instruments and outcomes

Residents’ gender, age and documented dementia diagnosis were extracted from the electronic care record system. While ‘dementia diagnosis’ provides general information if any kind of dementia is present, MMSE screening was applied by interviewers to assess current cognitive function, resulting in scores between 0 and 30 points, whereby higher scores represent less cognitive decline [37]. Pain intensity and interference were measured by the validated German version of the Brief Pain Inventory (BPI) [38]. Building on the initial question if any pain was present during the previous 24 hours, BPI measures pain intensity (four items assessing maximum, least, average and actual pain; 0—no pain to 10—worst pain) and interference (seven items assessing the impact of pain on general activities, mood, walking ability, resilience, relations with others, sleep and enjoyment of life; 0—does not interfere to 10—completely interferes) along two scales. BPI items were analysed item by item as well as by computing pain intensity and interference scales representing respondent’s mean scores across the original items. Interviewer affiliation and gender were protocolled in the survey tablet system.

Statistical analysis

All analyses were conducted using IBM SPSS 27. Absolute numbers (n) and percentages (%) of non-missing values were reported for categorical variables; arithmetic means (AM), standard deviation (SD), median (MD) and minimum and maximum (Min, Max) were used to describe metric variables.

Interviewer variance was measured by intra-interviewer correlation coefficients (IIC or ρ_{INT}) drawing on Kish’s analysis-of-variance (ANOVA) framework [39, 40]. In general, ρ expresses the ratio of between-interviewer variance to the total variance in the variable of interest [17, 22]. The larger this coefficient, the more similar are the responses per

Table 1. Interviewer and nursing home resident characteristics

Sample characteristics	Total sample	Subsample ($\geq 3_{int}$)
Interviewer level	BPI, $n = 61$	BPI ($\geq 3_{int}$), $n = 36$
Gender		
Female	45 (73.8%)	29 (80.5%)
Male	16 (26.2%)	7 (19.5%)
Affiliation		
External	42 (68.9%)	25 (69.4%)
Internal	19 (31.1%)	11 (30.6%)
Interviewer workload (number _{int})		
AM (SD)	4.23 (2.8)	6.03 (2.7)
MD, min-max	3.5, 1–13	5.5, 3–13
Resident level	BPI, $n = 258$	BPI ($\geq 3_{int}$), $n = 217$
Gender		
Female	166 (64.3%)	151 (69.6%)
Male	92 (35.7%)	66 (30.4%)
MMSE (score)		
AM (SD)	22.65 (5.1)	21.72 (5.3)
MD, min-max	23.0, 10–30	22.0, 10–30
Age (years)		
AM (SD)	81.29 (9.8)	82.00 (9.5)
MD, min-max	83.0, 60–101	84.0, 60–101
Documented dementia		
Yes	59 (22.9%)	56 (25.8%)
No	185 (71.7%)	149 (68.7%)

Notes: Total sample depicts characteristics based on total resident data, Subsample depicts characteristics based on resident data collected by interviewers with a workload of at least three interviews; n , number (% percentage of non-missing values); BPI, Brief Pain Inventory; MMSE, Mini-Mental State Examination; int, number of interviews; AM, arithmetic mean; SD, standard deviation; MD, median; Min, minimum; Max, maximum.

interviewer, and hence, the more response behaviour can be attributed to differences between interviewers. Interviewer variance cannot be prevented, but most ρ -coefficients rarely exceed 0.05 [41]. To reach sufficient variability of responses, ρ_{INT} were calculated for interviewers who conducted at least three interviews (i.e. interviewer workload ≥ 3). ρ -Coefficients derived from fixed-effects ANOVA are displayed for all BPI items and are summarised in boxplots. Results are presented for total sample and separately for interviewer’s subsamples.

To analyse whether mean scores of BPI scales differ by AOI and GOI, two-factorial covariance analyses (ANCOVA) for each pain outcome were conducted. AOI and GOI were included as main effects as well as interaction terms (i.e. affiliation \times gender) and adjusted by NHR’s MMSE, gender, age and pre- vs. post-test (i.e. before vs. after parental study’s intervention). Eta² (η^2) was calculated as a measure of association and type 1 error was set at $\alpha = 0.05$ (two-sided). Profile plots are used to visualise the estimated, adjusted marginal means.

Results

Interviewer and resident characteristics

Table 1 presents both the interviewer- and resident-level characteristics for the total sample as well as for the subsample data collected by all interviewers with an interviewer

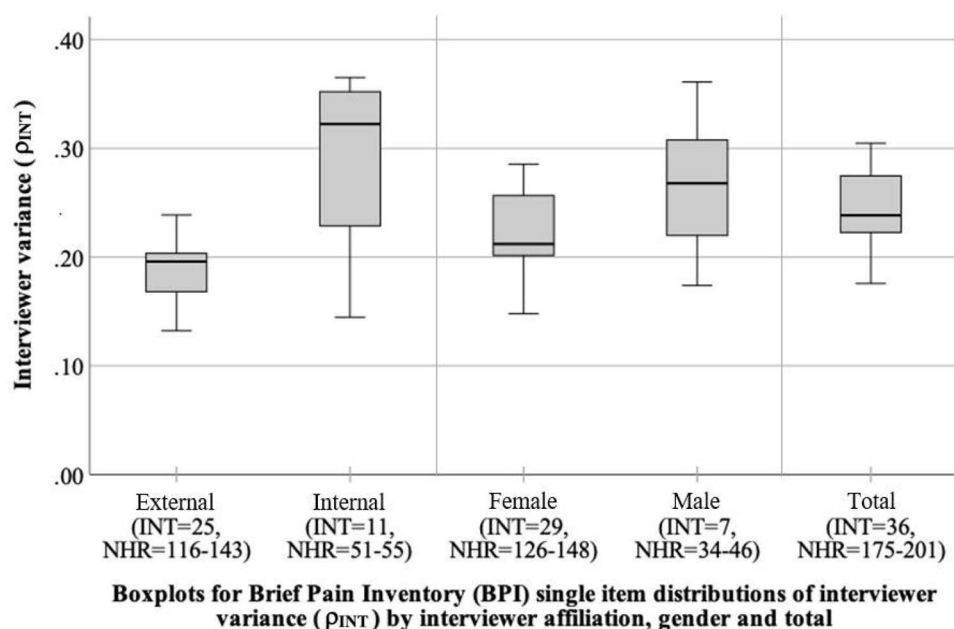


Figure 1. Boxplots for BPI items' interviewer variance (ρ_{INT}).

workload of $n \geq 3$. In total, pain intensity and interference of 258 NHRs were assessed by 61 interviewers. Three quarters were female and one-third was nurses from the respective NH. The average interviewer workload was 4.3 (SD = 2.8). Two-thirds of interviewed NHRs were female, and documented dementia was present in about 20% of NHRs. On average, NHRs were 81.3 (SD = 9.8) years of age and exhibited an MMSE score of 22.6 points (SD = 5.1).

An interviewer workload $n \geq 3$ was conducted by 36 interviewers, resulting in a sample of 217 NHRs. The average number of interviews was 6.0 per interviewer (SD = 2.7). Besides that, interviewer- and resident-level characteristics were basically similar.

Subsample sizes by external, internal, female and male interviewers varied between 7 and 44 interviewers and 57 and 198 examined NHRs, respectively. While most sample characteristics proved to be comparable, NHR's gender and documented dementia proportions differed between the subsamples (Supplementary Table S2).

Interviewer effects in terms of interviewer variance in BPI items

Interviewer variance (ρ_{INT}) was computed for the NHR's data collected by interviewers with an interviewer workload of $n \geq 3$. Details for all single items of BPI scales for the total sample as well as for AOI and GOI subsamples are displayed in Supplementary Table S3. In the total sample, the proportions of variance explained by differences between interviewers ranged between 17.6 and 30.5% for the 11 BPI items, exhibiting a mean ρ_{INT} of 0.242 (MD = 0.238). Figure 1 depicts this distribution while also presenting boxplots for the AOI and GOI subsamples.

Interviewer variance varied between external and internal as well as female and male interviewers but was especially pronounced in terms of AOI (Figure 1, Supplementary Table S3). All but one single items' ρ_{INT} were clearly lower within NHRs interviewed by external interviewers (AM = 0.190, MD = 0.196) as compared to NHRs interviewed by internal (AM = 0.288, MD = 0.322) interviewers.

GOI caused less pronounced differences in BPI items' ρ_{INT} coefficients. While the variation of NHR responses due to interviewer heterogeneity was explained by an average 26.5% (MD = 26.8%) within male interviewers, this amounted to an average of 22.2% (MD = 21.2%) for female interviewers. Relating to comparisons of ρ_{INT} between the single BPI items, almost all coefficients were higher in the NHR subsamples interviewed by male interviewers.

Interviewer effects in terms of biased BPI intensity and interference mean scores

Two-factorial ANCOVA (Table 2a) shows a significant disordinal interaction between interviewer gender and affiliation on NHR average pain intensity ($\eta^2 = 5.3\%$, $F = 12.2$, $P < 0.001$, $n = 227$). GOI and AOI exhibit a joint and counteracting effect on pain intensity score: NHRs reported the highest average pain intensity when being interviewed by external male (AM = 5.33) or internal female (AM = 5.28) interviewers. The lowest pain intensities were stated when NHRs were interviewed by external women (AM = 4.24) or internal men (AM = 4.29).

A combined effect of GOI and AOI on pain interference was found and indicates a weak but significant, disordinal interaction effect ($\eta^2 = 2.2\%$, $F = 4.9$, $P = 0.028$, $n = 245$) (Table 2b). While the mean pain interference score

Table 2. ANCOVA for BPI scores by affiliation- and gender-of-interviewer

ANCOVA BPI scores by interviewer gender and affiliation			
(a) BPI pain intensity score (0–10)		(b) BPI pain interference score (0–10)	
INT gender		INT gender	
Female AM (SE), <i>n</i>	4.76 (0.2), 173	Female AM (SE), <i>n</i>	4.23 (0.2), 189
Male AM (SE), <i>n</i>	4.81 (0.3), 54	Male AM (SE), <i>n</i>	4.19 (0.3), 56
INT affiliation		INT affiliation	
External AM (SE), <i>n</i>	4.78 (0.2), 159	External AM (SE), <i>n</i>	4.61 (0.2), 176
Internal AM (SE), <i>n</i>	4.77 (0.2), 68	Internal AM (SE), <i>n</i>	3.81 (0.3), 69
Gender × affiliation*		Gender × affiliation*	
INT (f) × (ex) AM (SE), <i>n</i>	4.24 (0.2), 130	INT (f) × (ex) AM (SE), <i>n</i>	4.22 (0.2), 144
INT (f) × (in) AM (SE), <i>n</i>	5.28 (0.3), 43	INT (f) × (in) AM (SE), <i>n</i>	4.42 (0.4), 45
INT (m) × (ex) AM (SE), <i>n</i>	5.33 (0.3), 29	INT (m) × (ex) AM (SE), <i>n</i>	5.00 (0.4), 32
INT (m) × (in) AM (SE), <i>n</i>	4.29 (0.4), 25	INT (m) × (in) AM (SE), <i>n</i>	3.38 (0.5), 56
Between subject effects		Between subject effects	
<i>Factors eta² (F), P</i>		<i>Factors eta² (F), P</i>	
INT gender	0.0% (0.03), <i>P</i> = 0.862	INT gender	0.0% (0.01), <i>P</i> = 0.914
INT affiliation	0.0% (0.01), <i>P</i> = 0.991	INT affiliation	1.84% (4.51), <i>P</i> = 0.036
INT gender × affiliation	5.26% (12.2), <i>P</i> < 0.001	INT gender × affiliation	2.01% (4.88), <i>P</i> = 0.028
<i>Covariates eta² (F), P</i>		<i>Covariates eta² (F), P</i>	
NHR gender	1.32% (2.94), <i>P</i> = 0.088	NHR gender	4.10% (10.15), <i>P</i> = 0.002
NHR age	0.35% (0.77), <i>P</i> = 0.381	NHR age	1.19% (2.84), <i>P</i> = 0.093
NHR MMSE score	0.89% (1.97), <i>P</i> = 0.162	NHR MMSE score	0.86% (2.06), <i>P</i> = 0.152
Sample pre/post	2.77% (6.24), <i>P</i> = 0.013	Sample pre/post	1.59% (0.38), <i>P</i> = 0.540
R² (R²_{corr}), <i>n</i>	9.9% (7.0%), 227	R² (R²_{corr}), <i>n</i>	6.7% (4.4%), 245

Notes: ANCOVA, analysis of covariance; BPI, Brief Pain Inventory; INT, interviewer; NHR, nursing home resident; MMSE, Mini-Mental State Examination; f, female; m, male; ex, external; in, internal; AM, arithmetic marginal mean; SE, standard error; *n*, sample size; eta², partial eta-square; F, F-value; *P*, error probability; check of ANCOVA assumptions demonstrated no relevant model violations; * partially small subgroup sample size.

was almost identical for external (AM = 4.22) and internal (AM = 4.42) female interviewers, marginal means differed between external (AM = 5.00) and internal male interviewers (AM = 3.38).

Details for pairwise comparisons of simple main effects are provided in [Supplementary Table S4](#). [Figure 2](#) depicts the estimated, adjusted marginal means by interviewer affiliation and gender for BPI pain intensity ([Figure 2a](#)) and BPI pain interference ([Figure 2b](#)).

Discussion

This study analysed whether affiliation- and gender-of-interviewer influence NHR responses in reporting pain intensity and interference. It was found that a considerable amount of response variability is due to interviewers' affiliation and gender, and that pain reports vary systematically by joint effects of AOI and GOI. Our findings give rise to potential risks of imprecise and biased survey estimates in NH populations.

Overall, the share of variance explained by interviewer heterogeneity amounts to an average of 24.2%, which represents a comparatively high proportion in general terms [23, 25, 40, 42]. Interviewer variance was especially pronounced in NHR responses interviewed by NH-internal interviewers. Observed interviewer effects might be explained by specific role- and status-related interpretations [28, 33, 35] on the part of NHR, that is, residents attribute

certain expectations and meanings to the interviewer. NHRs are dependent on caregivers for their care and may fear negative repercussions if they express undesirable information about perceived care provision [43, 44]. NHR responses might reflect such attributed expectations towards familiar, internal nurses leading to more heterogeneous responses. Furthermore, compared to external interviewers ('lays'), NHRs might have attributed internal nurses with a higher professional competence ('experts') [27, 28, 45], and provided more differentiated but person-specific, hence, varying pain reports. Finally, higher response variability in internal interviewers may suggest that they deviate more often from the predefined standardised interview script [46]. This is plausible insofar as interviews represent a communication opportunity for older persons [9]. The risk of digressing into the usual everyday communication is higher in the case of familiar nurses. Also, nurses usually know 'their' residents relatively well, may have individually adapted forms of communication per NHR, and thus, inducing higher variability of responses.

Our findings also show that pain scores differ according to who conducts the interviews. Pain intensity was highest when internal women or external men interviewed and lowest with external women and internal men. Highest pain interference was reported towards external men and lowest towards internal men but did not vary within female interviewers. Again, gender- and status-related expectations and stereotypes seem to elicit socially desirable and

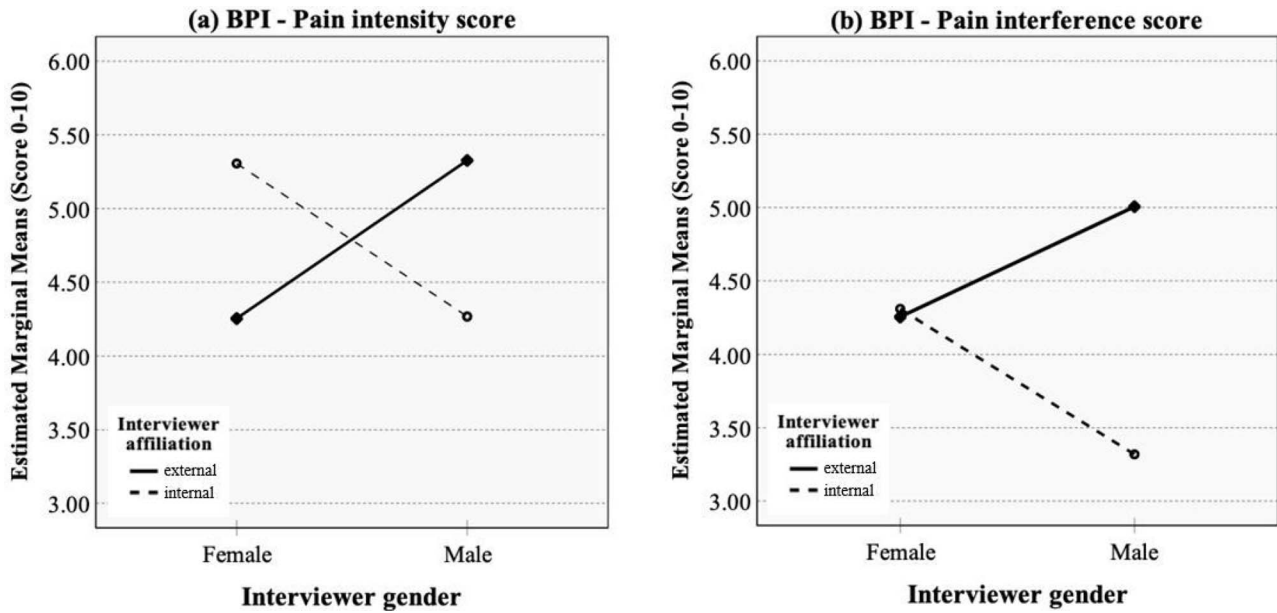


Figure 2. Profile plots for BPI scores by affiliation- and gender-of-interviewer.

role-dependent responses. Our results are consistent with other findings that demonstrate gender-specific communication with physicians and nurses [35], biased observer pain estimations due to gender role stereotypes [47], and varying pain levels by experimenter gender [48]. Individuals who consider themselves to be more masculine show higher pain thresholds and tolerances when providing pain information [49]. At the same time, response patterns of health-related information differ according to whether interviewers are attributed a certain status and topic-related competencies [27, 28, 45]. Hence, male internal nurses may collect lower pain scores as NHRs anticipate higher pain reports to be interpreted as dissatisfaction with received care. External male interviewers may be attributed as representatives of the official research institution and NHRs may be more likely to report high pain scores hoping to initiate improvements in care provision. For female interviewers, these patterns seem to be exactly opposite. Internal female interviewers seem to represent trustworthy and empathetic contact persons within the NH structure, all the while incorporating rather male positions of power, such as physicians or facility managers [34, 35]. External female interviewers may be seen as less trustworthy with low expert status and less power to effect any changes in care provision against the background of gender-induced role expectations.

Our study faces several limitations. This study's procedures were determined by the overarching project's study design. Interviewers were chosen pragmatically, do not represent a random sample from a population of interviewers and were not randomly assigned to respondents to avoid confounding of area effects, interviewer and respondent characteristics [17, 18]. For fieldwork feasibility, it was not possible to conduct such an optimal interpenetrated survey design. Some interviewer subsample numbers were low and

imbalanced in terms of gender and affiliation. Multi-level analyses to account for clustering of respondents by interviewers including random effects were not feasible for several items due to small sample sizes. This also limits further analyses in ANCOVA models to examine possible moderating factors, such as adding three-way interaction terms using NHR gender or testing gender-related (mis)matches, i.e. if female or male residents are interviewed by female or male interviewers [50]. Besides, age differences between interviewers and respondents might influence responses in interviews [2]. Interviewer age and characteristics like experience, skills, personality traits or attitudes towards pain or gender roles were not assessed but exhibit further explanatory potential to shed more light on the 'black box' [26] of the interview process. Although our findings do not suggest that the parental study's nursing education interventions affected our internal interviewers in applying the standardised survey, we cannot rule out that these might have introduced an additional difference between internal and external interviewers. While we do not expect our results to be biased because we used the MMSE as the current estimate of NHR cognitive function, the proportion of NHR with documented dementia varied between interviewer subsamples, possibly indicating differing interviewer strategies in selecting to-be-interviewed residents. Future studies should also focus on possible interactions of participant selection procedures with interviewer and respondent characteristics.

Given the paucity of methodological research in this specific field, we believe that our findings are of high relevance for both research and practice. While interviewer variance increases standard errors and decreases the precision of survey estimates, reported pain outcomes substantially differ by interactions between GOI and AOI. Anticipating the risk of imprecise and even biased outcomes, interviewer

effects should be considered when designing, analysing and interpreting surveys with older persons. Our results also give rise to important implications for pain management in NHs. Medical and nursing decisions rely on regular pain assessment. Depending on the assessor, pain may be over- or underestimated and care provision is unlikely to build upon a resident's 'true' perceived health state. Hence, future studies should also focus on the risk of inadequate care of NHR with pain as an unwitting but maybe far-reaching consequence from potentially biased pain assessment results.

Conclusion

NHR responses are influenced by interviewer characteristics. Self-reports on pain vary according to GOI and AOI. Interviewer characteristics are assumed to evoke gender- and authority-related attributions and expectations and substantially affect study outcomes of self-reported pain intensity and interference. The risk of interviewer effects in face-to-face interviews with older respondents seems to be high and threatens survey data quality in NHRs.

Supplementary Data: Supplementary data mentioned in the text are available to subscribers in *Age and Ageing* online.

Declaration of Conflicts of Interest: None.

Declaration of Sources of Funding: The parental study was funded by German nursing home operator Curanum AG and the province of Salzburg, Austria. Present methodological study does not relate to parental funding.

References

- Knäuper B, Carriere K, Chamandy M *et al.* How aging affects self-reports. *Eur J Ageing* 2016; 13: 185–93.
- Vidovičová L, Doseděl T. Who should interview older people? The Effect of Interviewer and Interviewee Characteristics in Surveys of Older People and Aging Topics *Sociológia* 2018; 50: 760–81.
- Schanze J-L. Response behavior and quality of survey data: comparing elderly respondents in institutions and private households. *Sociological Methods & Research* 2021; 50. <https://doi.org/10.1177%2F0049124121995534>.
- Fox MT, Sidani S, Streiner D. Using standardized survey items with older adults hospitalized for chronic illness. *Res Nurs Health* 2007; 30: 468–81.
- Bisla J, Calem M, Begum A *et al.* Have we forgotten about dementia in care homes? The importance of maintaining survey research in this sector. *Age Ageing* 2011; 40: 5–6.
- Tyler DA, Shield RR, Rosenthal M *et al.* How valid are the responses to nursing home survey questions? Some Issues and Concerns *The Gerontologist* 2010; 51: 201–11.
- Fisher SE, Burgio LD, Thorn BE *et al.* Obtaining self-report data from cognitively impaired elders: methodological issues and clinical implications for nursing home pain assessment. *Gerontologist* 2006; 46: 81–8.
- Kelfve S, Thorslund M, Lennartsson C. Sampling and non-response bias on health-outcomes in surveys of the oldest old. *Eur J Ageing* 2013; 10: 237–45.
- Perfect D, Griffiths AW, Vasconcelos Da Silva M *et al.* Collecting self-report research data with people with dementia within care home clinical trials: benefits, challenges and best practice. *Dementia (London)* 2021; 20: 148–60.
- Livingston G, Sommerlad A, Orgeta V *et al.* Dementia prevention, intervention, and care. *The Lancet* 2017; 390: 2673–734.
- Weatherhead I, Courtney C. Assessing the signs of dementia. *Practice Nursing* 2012; 23: 114–8.
- Lam HR, Chow S, Taylor K *et al.* Challenges of conducting research in long-term care facilities: a systematic review. *BMC Geriatr* 2018; 18: 242–53.
- Tourangeau R, Rips LJ, Rasinski K. *The Psychology of Survey Response*. Cambridge: Cambridge University Press, 2012.
- Bradburn NM. Understanding the question-answer process. *Statistics Canada* 2004; 30: 5–15.
- Tourangeau R. Defining hard-to-survey population. In: Tourangeau R, Edwards E, Johnson TP *et al.*, eds. *Hard-to-Survey Populations*. Cambridge: Cambridge University Press, 2014; 3–20.
- Randall S, Coast E, Compaore N *et al.* The power of the interviewer. *Demographic Research* 2013; 28: 763–92.
- Groves RM. *Survey Errors and Survey Costs*. New York: Wiley, 2004.
- Davis RE, Couper MP, Janz NK *et al.* Interviewer effects in public health surveys. *Health Educ Res* 2010; 25: 14–26.
- West BT, Blom AG. Explaining interviewer effects: a research synthesis. *Journal of Survey Statistics and Methodology* 2017; 5: 175–211.
- Huddy L, Billig J, Bracciodieta J *et al.* The effect of interviewer gender on the survey response. *Political Behavior* 1997; 19: 197–220.
- Ackermann-Piek D, Schröder J, Kluge R *et al.* Interviewer Effects in Standardized Surveys. Mannheim: GESIS - Leibniz Institute for Social Sciences (GESIS - Survey Guidelines), 2019.
- Loosveldt G, Wuyts C. A comparison of different approaches to examining whether interviewer effects tend to vary across different subgroups of respondents. In: Olson A, Smyth JD, Dykema J *et al.*, eds. *The Past, Present, and Future of Research on Interviewer Effects*. Boca Raton, London, New York: CRC Press - Taylor & Francis Group, 2020; 311–22.
- Winters S, Strating MH, Klazinga NS *et al.* Determining the interviewer effect on CQ index outcomes: a multilevel approach. *BMC Med Res Methodol* 2010; 10: 75. <https://doi.org/10.1186/1471-2288-10-75>.
- Freeman J, Butler EW. Some sources of interviewer variance in surveys. *Public Opin Q* 1976; 40: 79–91.
- Cleary PD, Mechanic D, Weiss N. The effect of interviewer characteristics on responses to a mental health interview. *Journal of Health and Social Behavior* 1981; 22: 183–93.
- Beullens K, Loosveldt G, Vandenplas C. Interviewer effects among older respondents in the European Social Survey. *Journal of Public Opinion Research* 2018; 31: 609–25.
- Smeeth L, Fletcher AE, Stirling S *et al.* Randomised comparison of three methods of administering a screening questionnaire to elderly people: findings from the MRC trial of the assessment and management of older people in the community. *BMJ* 2001; 323: 1403–7.

28. Tzourio C, Gagnière B, El Amrani M *et al.* Lay versus expert interviewers for the diagnosis of migraine in a large sample of elderly people. *J Neurol Neurosurg Psychiatry* 2003; 74: 238–41.
29. Hadjistavropoulos H, Craig KD, Fuchs-Lacelle S. Social influences and the communication of pain. In: Hadjistavropoulos H, Craig KD, eds. *Pain. Psychological Perspectives*. New Jersey: Mahwah, 2004; 87–112.
30. Takai Y, Yamamoto-Mitani N, Okamoto Y *et al.* Literature review of pain prevalence among older residents of nursing homes. *Pain Manag Nurs* 2010; 11: 209–23.
31. Zwakhalen S, Docking RE, Gnass I *et al.* Pain in older adults with dementia: a survey across Europe on current practices, use of assessment tools, guidelines and policies. *Schmerz* 2018; 32: 364–73.
32. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant* 2013; 47: 2025–47.
33. Lipps O, Lutz G. Gender of interviewer effects in a multi-topic centralized CATI panel survey. *Methods, Data, Analyses* 2016; 11: 67–86.
34. Reitinger E, Lehner E, Pichler B *et al.* "Doing Gender" im Altenpflegeheim. Perspektiven von Mitarbeitenden und Führungskräften. *Z Gerontol Geriatr* 2016; 8: 700–5.
35. Samulowitz A, Gremyr I, Eriksson E *et al.* "Brave men" and "emotional women": a theory-guided literature review on gender bias in health care and gendered norms towards patients with chronic pain. *Pain Res Manag* 2018; 2018: 6358624. <https://doi.org/10.1155/2018/6358624>.
36. Kutschar P, Berger S, Brandauer A *et al.* Nursing education intervention effects on pain intensity of nursing home residents with different levels of cognitive impairment: a cluster-randomized controlled trial. *J Pain Res* 2020; 13: 633–48.
37. Folstein MF, Folstein SE, Fanjiang G. *MMSE Mini-Mental State Examination Clinical Guide*. Odessa: Psychological Assessment Resources, 2001.
38. Budnick A, Kuhnert R, Könnner F *et al.* Validation of a modified German version of the brief pain inventory for use in nursing home residents with chronic pain. *J Pain* 2016; 17: 248–56.
39. Kish L, Slater CW. Studies of interviewer variance for attitudinal variables. *J Am Stat Assoc* 1962; 57: 92–115.
40. Brunton-Smith I, Sturgis P, Leckie G. Detecting and understanding interviewer effects on survey data using a cross-classified mixed-effects location scale model. In: National Centre for Research Methods (ed). *NCRM Working Paper: NCRM*, 2016.
41. Beullens K, Loosveldt G. Interviewer effects in the European Social Survey. *Survey Research Methods* 2016; 10: 103–18.
42. West BT, Olson K. How much of interviewer variance is really nonresponse error variance? *Public Opin Q* 2010; 74: 1004–26.
43. Knäuper B, Schwarz N, Park D. Selbstberichte im alter. In: Motel-Klingebiel A, Kelle U, eds. *Perspektiven der empirischen Alter(n)ssoziologie*. Wiesbaden: Springer, 2002; 75–98.
44. Kelle U, Niggemann C. "Weil ich doch vor zwei Jahren schon einmal verhört worde bin..." - Methodische Probleme bei der Befragung von Heimbewohnern. In: Motel-Klingebiel A, Kelle U, eds. *Perspektiven der empirischen Alter(n)ssoziologie*. Wiesbaden: Springer, 2002; 99–132.
45. Kohler-Riessman C. Interviewer effects in psychiatric epidemiology: a study of medical and lay interviewers and their impact on reported symptoms. *Am J Public Health* 1979; 69: 485–91.
46. van der Zouwen J, Smit JH, van der Horst MHL. Reporting the Frequency and Duration of Household Tasks by Elderly Respondents: The Effect of Different Interview Strategies on Data Quality. In: *Annual Meeting of the American Association for Public Opinion Association (Full paper)*. Miami Beach, FL: Fontainebleau Resort, 2005.
47. Zhang L, Losin EAR, Ashar YK *et al.* Gender biases in estimation of others' pain. *J Pain* 2021; 22: 1048–59.
48. Chapman CD, Benedict C, Schiöth HB. Experimenter gender and replicability in science. *Sci Adv* 2018; 4: e1701427. <https://doi.org/10.1126/sciadv.1701427>.
49. Alabas OA, Tashani OA, Tabasam G *et al.* Gender role affects experimental pain responses: a systematic review with meta-analysis. *Eur J Pain* 2012; 16: 1211–23.
50. Vercruyssen A, Wuyts C, Loosveldt G. The effect of sociodemographic (mis)match between interviewers and respondents on unit and item nonresponse in Belgium. *Soc Sci Res* 2017; 67: 229–38.

Received 2 July 2021; editorial decision 28 November 2021