



This is a repository copy of *Face validity of four preference-weighted quality-of-life measures in residential aged care: a think-aloud study.*

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/207930/>

Version: Published Version

Article:

Engel, L. orcid.org/0000-0002-7959-3149, Kosowicz, L., Bogatyreva, E. et al. (7 more authors) (2023) Face validity of four preference-weighted quality-of-life measures in residential aged care: a think-aloud study. *The Patient - Patient-Centered Outcomes Research*, 16 (6). pp. 655-666. ISSN 1178-1653

<https://doi.org/10.1007/s40271-023-00647-6>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



Face Validity of Four Preference-Weighted Quality-of-Life Measures in Residential Aged Care: A Think-Aloud Study

Lidia Engel¹ · Leona Kosowicz² · Ekaterina Bogatyreva^{1,3} · Frances Batchelor^{2,4} · Nancy Devlin⁴ · Briony Dow^{2,4} · Andrew S. Gilbert^{2,5} · Brendan Mulhern⁶ · Tessa Peasgood⁴ · Rosalie Viney⁶

Accepted: 28 August 2023 / Published online: 6 October 2023
© The Author(s) 2023

Abstract

Objective There is an increased use of preference-weighted quality-of-life measures in residential aged care to guide resource allocation decisions or for quality-of-care assessments. However, little is known about their face validity (i.e., how understandable, appropriate and relevant the measures are ‘on their face’ when respondents complete them). The aim of this study was to assess the face validity of four preference-weighted measures (i.e., EQ-5D-5L, EQ-HWB, ASCOT, QOL-ACC) in older people living in residential aged care.

Methods Qualitative cognitive think-aloud interviews were conducted using both concurrent and retrospective think-aloud techniques. To reduce burden, each resident completed two measures, with the four measures randomised across participants. Audio recordings were transcribed and framework analysis was used for data analysis, based on an existing framework derived from the Tourangeau four-stage response model.

Results In total, 24 interviews were conducted with residents living across three residential aged care facilities in Melbourne, Australia. Response issues were identified across all four measures, often related to comprehension and difficulty selecting a response level due to double-barrelled and ambiguous items that have different meanings in the residential aged care context. We also identified issues related to understanding instructions, non-adherence to the recall period, and noted positive responding that requires attention when interpreting the data.

Conclusions Our findings provide further evidence on the appropriateness of existing measures, indicating numerous response issues that require further research to guide the selection process for research and practice.

1 Introduction

By 2030, one in six people in the world will be aged 60 years and over [1]. The complex health and care needs of

an ageing population mean the demand for residential aged care (i.e., nursing homes or long-term institutional care) is expected to rise internationally. Currently an average of 1.5% of gross domestic product is spent on long-term care in OECD countries and this is estimated to double by 2050 [2]. This will increase pressure on government budgets, requiring greater efficiency in the design and delivery of services to meet the demand of an ageing population.

To ensure high quality service provision, there is an increasing awareness amongst policy makers of the need to link quality-of-care assessments to older persons’ quality of life (QoL), which represents an important person-centred indicator of quality of care. For example, the Australian Government has recently expanded the National Aged Care Mandatory Quality Indicator Program, adding six new indicators, including QoL, on top of the previous five mandatory clinical indicators [3]. From April 2023 onwards, residential aged care providers will be required to report the percentage of care recipients who report ‘good’ or ‘excellent’ QoL

✉ Lidia Engel
Lidia.engel@monash.edu

¹ Monash University Health Economics Group, School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia

² National Ageing Research Institute, Parkville, VIC, Australia

³ School of Public Health and Social Development, Deakin University, Burwood, VIC, Australia

⁴ The University of Melbourne, Parkville, VIC, Australia

⁵ Department of Social Inquiry, La Trobe University, Bundoora, VIC, Australia

⁶ The Centre for Health Economics Research and Evaluation, University of Technology Sydney, Sydney, NSW, Australia

Key Points for Decision Makers

While there is increased interest in the use of preference-weighted measure in residential aged care to guide resource allocation decisions and for quality assessment, little is known about their face validity in this setting.

This study sought to investigate how understandable, appropriate and relevant the EQ-5D-5L, ASCOT, EQ-HWB and QOL-ACC are 'on its face' when residents complete them.

Our study identified a number of response issues with existing measures and provides recommendations for potential adaptation that could improve face validity, including improvements to instruments' instructions, clarification around how to factor the use of health aids and assistance from others into the responses, provision of examples for activities of daily living that are more appropriate for older adults, and the avoidance of double-barrelled and ambiguous questions.

using the Quality-of-Life Aged Care Consumers (QOL-ACC) instrument [4]. Similarly, in the United Kingdom, the Care Quality Commission introduced quality ratings of care homes in 2017, based on how well the home keeps people safe, meets their health and care needs, responds to concerns, is caring and well-led [5]. Although QoL is currently not integrated into the UK quality ratings, a previous study examined the relationship between residents' QoL and quality ratings in care homes, showing that better care ratings were linked with higher QoL using the Adult Social Care Outcomes Toolkit (ASCOT) tool [6]. The ASCOT has also been tested in the Australian context to measure QoL outcomes, but only of community aged care programmes [7].

The standardisation of routine QoL measurement in residential aged care facilities requires the careful selection of QoL measures that capture important aspects of older people's QoL while being easily understood by residents and their proxies (i.e., family members, health professionals or aged care staff) who may be required to complete the QoL measure on behalf of residents due to cognitive or other health impairments. The QOL-ACC was co-designed from its inception with older people using aged care services, and is a relatively new measure with some emerging evidence demonstrating its psychometric validity as well as being meaningful, relevant and acceptable to aged care consumers and providers [8–10]. The ASCOT is a measure of a person's social care-related QoL [11], which was specifically developed to measure aspects of QoL that can be reasonably attributed to social care services [11]. The ASCOT is

a suite of measures available for self-completion (SCT4), interviewer administration (INT4), and for care home settings (CH4). Previous studies have provided evidence for the construct validity of the self-completion ASCOT tool (i.e., SCT4) [12] and demonstrated the acceptability and feasibility of the CH4 tool to inform practice in care homes [13].

An important characteristic of both the QOL-ACC and the ASCOT is that they are preference weighted, reflecting the relative importance of different domains of QoL. Preference-weighted measures are used in economic evaluation, where outcomes are expressed as quality-adjusted life-years (QALY) [14]. A QALY is a unit that combines both length and QoL and is recommended by many health technology assessment agencies around the world to assess the cost effectiveness of healthcare interventions [15–17]. To date, QALYs have most commonly been derived using the EQ-5D [18], which is also the most commonly used instrument to measure outcomes in economic evaluation within aged care [19, 20]. The EQ-5D measures health-related QoL with three (EQ-5D-3L) or five (EQ-5D-5L) corresponding response levels [21, 22]. However, recognising that services for older people often have broader benefits that may go beyond health (e.g., safety or independence) [23], a new preference-weighted measure, the EQ Health and Wellbeing (EQ-HWB), has been developed by the Euro-QoL group [24, 25]. The EQ-HWB extends the descriptive system beyond health, capturing broader aspects of wellbeing. Although the development of the measure involved older people and social care users [26, 27], more evidence is needed regarding its performance among older people.

While there is increased interest in the use of preference-weighted measures in residential aged care to guide resource allocation decisions and for quality assessment, little is known about their face validity in this setting. Face validity is an important aspect of content validity and refers to the extent to which a measure appears, on the surface, to measure what it is intended to measure [28]. Face validity is concerned with how understandable, appropriate and relevant the items of a particular instrument are 'on its face' when participants complete it [29], which is often assessed using appropriate cognitive debriefing methods [30]. The aim of this study was, therefore, to assess the face validity of the four preference-weighted measures (i.e., EQ-5D-5L, EQ-HWB, ASCOT, QOL-ACC) in older people living in residential aged care using a think-aloud approach accompanied by qualitative analysis.

2 Methods

2.1 Sample and Recruitment

Residents were recruited across three sites from a not-for-profit residential aged care provider in Melbourne, Australia,

in October and November 2022. Each aged care facility nominated residents based on cognitive functioning and the ability to speak and understand the English language. A staff member from the facility then introduced the interviewer to the resident. All nominated residents received an explanation of the study's purpose and process, the risks and benefits of participating, as well as an information sheet. Nominated residents were also invited to ask questions before written consent was obtained, prior to the interview. To ensure consent was informed, the researcher asked residents to summarise the study in their own words. This study was approved by the Monash University Human Research Ethics Committee (Project ID: 32170) and participants received reimbursement in the form of a gift voucher.

2.2 Measures

All questionnaires were original paper-based versions provided by the instrument developers. Self-report versions were used, given that self-completion is the recommended approach by the National Aged Care Mandatory Quality Indicator Program in Australia [3].

EQ-5D-5L The EQ-5D-5L comprises five dimensions (mobility, personal care, usual activities, pain/discomfort, and anxiety/depression) with five response options (no problems, slight problems, moderate problems, severe problems, extreme problems/unable) [22]. The five questions are accompanied by the EQ VAS, a vertical scale numbered 0–100, with zero denoting ‘the worst health you can imagine’ and 100 representing ‘the best health you can imagine’. The EQ-5D-5L instructs respondents to describe their health ‘TODAY’. The Australian English-language version of the EQ-5D-5L was used.

QOL-ACC The QOL-ACC is a new older-person-specific QoL instrument designed for application in quality assessment and economic evaluation in aged care, including in-home and residential care settings [4]. It was co-designed from its inception with older people using aged care services and includes six dimensions (independence, mobility, pain management, emotional wellbeing, social connection, and activities). Five response levels are attached to each dimension, ranging from the best level ‘all of the time’ to the worst level ‘none of the time’ [8]. The QOL-ACC asks respondents to describe how they feel about their ‘current situation’.

ASCOT The ASCOT is a measure of a person's social care-related QoL and includes eight domains (personal comfort and cleanliness, personal safety, food and drink, occupation, control over daily life, social participation, home cleanliness and comfort, and dignity) with four levels (ideal state, no needs, some needs, and high needs) [11]. The ASCOT toolkit includes a number of different versions

based on mode of completion, inclusion of observation, proxy responses and whether data is collected on the difference between QoL with current care levels and expected QoL without current care levels [31]. Despite the availability of the ASCOT care homes toolkit (CH4), which differentiates between current and expected social care-related quality of life, the self-report four-level version (SCT4) was used in this study [32], which only measures current social care-related quality of life, ensuring consistency with the other measures used in this study. While the interviewer-led ASCOT (INT4) prompts respondents to ‘*think about your situation at the moment*’, the self-report SCT4 version of the ASCOT does not specify a particular recall period in the instructions, as it was found that it made the questions difficult to follow and because many people had conditions that fluctuated, leading respondents to ignore the recall period [11].

EQ-HWB The EQ-HWB is a new measure of health and wellbeing, suitable for use across health care and social care [24]. The development of the EQ-HWB was based on extensive review of qualitative studies and field work including face validity and psychometric assessment across six countries (Argentina, Australia, China, Germany, the United Kingdom, and the United States) to inform final item selection [26, 27]. Two experimental versions are currently available (a 25-item long measure and a 9-item short-version [EQ-HWB-S]). Although the EQ-HWB-S forms the descriptive system for which a pilot value set has been developed [25], in practise, it is also possible to derive preference-weighted scores from the 25-item EQ-HWB. Thus, in this study the 25-item EQ-HWB English version for Australia was used, which has identical wording to the UK version. The recall period of the EQ-HWB refers to the ‘last 7 days’.

2.3 Procedure

After participants provided written consent, interviews were conducted by one of two interviewers (LE and LK), both experienced in qualitative research with older people. Participants were able to nominate their preferred location for the interview (resident's room or a common area) and privacy from staff members and other residents was ensured. Participants were first asked to complete a brief demographic questionnaire. To reduce cognitive burden, participants were then asked to complete only two of the four QoL questionnaires, with random assignment across the four instruments (EQ-5D-5L, ASCOT, QOL-ACC and EQ-HWB). The interview used think-aloud methods, where participants were asked to speak out loud to articulate their thoughts about the questionnaires while completing them [33]. The think-aloud method aids in understanding the cognitive processing

step—from presenting questions to the respondent to the selection of responses. If participants became silent, they were prompted to continue thinking aloud. Both concurrent think-aloud techniques (i.e., asking participants to think aloud while they are completing the questionnaire) and retrospective think-aloud techniques (i.e., asking participants to describe how they arrived at responses) were used during interviews [34].

Once the first questionnaire was completed, participants were asked to reflect on whether the measure was clear, easy to complete, comprehensive (i.e., no missing QoL domains) and of acceptable length. After completing the second questionnaire, participants were also asked which measure they preferred and why. All interviews followed an interview guide that included instructions and prompts, which was designed to keep the interview length under one hour to prevent participant fatigue. The interview guide (see Supplementary file 1 in the electronic supplementary material [ESM]) was designed based on recommendations for cognitive interviewing [34, 35]. All sessions were audio recorded and transcribed verbatim.

2.4 Analysis

Analysis of qualitative data was undertaken independently by the two interviewers. First, the transcripts were read repeatedly to ensure familiarity. One person (LK) extracted participants' responses concerning each item of the corresponding measure into an Excel spreadsheet and noted the selected response option for the corresponding item. Two people (LE and LK) coded the qualitative text using framework analysis by adapting the coding framework from a previous study, which examined the content validity and feasibility of the EQ-5D-3L, ICECAP-O and ASCOT in older adults [36]. This coding framework developed by van Leeuwen et al. was based on Tourangeau's model of survey response, which comprises four components: comprehension (i.e., understanding of instructions, questions, or response options), retrieval (i.e., retrieving appropriate information from memory), judgement (i.e., judging how the recalled information should be used to answer) and response mapping (i.e. formatting information into a valid response) [37]. The framework developed by van Leeuwen et al. was adapted to the current study by dropping categories that were not applicable (e.g., disagreement with order of response options) and adding new categories that emerged during the analysis (e.g., item inappropriate). During analysis, disagreements were resolved via discussions between the two coders, which also involved reflection on potential personal biases. Data saturation (i.e., the point at which no new categories emerged from the data) was discussed regularly,

Table 1 Characteristics of study participants

	N = 24
Gender, female: <i>n</i> (%)	15 (63%)
Mean age (SD), min–max	82 (8.6), 68–95
Country of birth, Australia: <i>n</i> (%)	16 (67%)
English first language, yes: <i>n</i> (%)	22 (92%)
Education: <i>n</i> (%)	
Year 10 or less	10 (42%)
Year 11/12	3 (13%)
Certificate	3 (13%)
Diploma	2 (8%)
University degree	6 (25%)
Living in residential aged care: <i>n</i> (%)	
Less than half a year	3 (13%)
Half to 1 year	6 (25%)
1–2 years	3 (13%)
3–4 years	6 (25%)
6–10 years	3 (13%)
More than 10 years	2 (8%)
Unsure	1 (4%)

and was reached after 18 interviews, but it was agreed that six additional interviews should be conducted as they were already scheduled with residents.

3 Results

In total, 24 residents participated in the study. Characteristics of study participants are outlined in Table 1. All four measures were explored in 12 interviews and appeared equally often as the first or second measure. Details of randomisation are provided in Supplementary file 2 (see ESM). The final response issue framework used for analysis is presented in Table 2. The matrix of response issues identified across the four measures is presented in Table 3. For ease of reporting, the 25 EQ-HWB items were grouped under domains [38]. Response issues are discussed in the following sections for each measure separately.

3.1 EQ-5D-5L

Not all participants understood how to complete the questionnaire, including that only one response per dimension was required. Participants often selected multiple response options within the same dimension, reflecting the variation of their health that they experienced, thereby ignoring the recall period *TODAY*. One participant also commented on the fact that the EQ-5D-5L does not ask 'proper questions':

“There’s no real question to this. It’s just got usual activities, it doesn’t say, do you have problems with your usual activities or anything like that.” The layout consisting of multiple boxes also made it more difficult for participants to follow the order of the questions: “Where am I up to? I’ve lost my place.” Participants sometimes felt the need to provide further details about their health. For example, for mobility, participants felt the need to clarify whether they were using mobility aids, which was added as text next to the box: “I have no problems with walking around. I’ll tick it but I’ll put walker.”

Interpretation issues mainly arose due to uncertainty around the inclusion of mobility aids, such as use of a wheelchair or walker: “I’m not able to walk around without a walker. So, I don’t really know how to answer that.” Therefore, participants wanted separate response options reflecting their mobility with the use of mobility aids and without. The personal care dimension was also problematic for one resident who required different levels of assistance: “I can dress myself. I can’t shower myself.” One resident referred to severe discomfort because of sitting in a wheelchair but only experienced moderate pain; yet the resident selected ‘moderate’ on the pain/discomfort dimension. This was also a problem for the anxiety/depression dimension, which residents thought would represent two separate constructs yet allowed only one response. There were also interpretation issues in terms of whether the pain/discomfort and anxiety/depression questions should be answered as if they were hypothetically not taking medication or whether they should consider the use of medications.

Item inappropriateness was especially an issue for the usual activities dimension, where examples provided (e.g., work, study, housework, family or leisure activities) did not reflect the types of usual activities in residential

aged care. As a result, one resident skipped the question. When prompted to describe which usual activities would be more relevant, residents often referred to aspects of personal care (e.g., showering) or walking, despite previously answering questions around mobility and personal care in the EQ-5D-5L.

Recall issues mainly arose for the pain/discomfort dimension, where participants often described how their level of pain fluctuates: “At the present minute, I have no pain. But if you had come yesterday, the ankle was hurting a little bit”. One resident selected ‘moderate pain/discomfort’ yet used the term ‘occasionally’, suggesting that a frequency scale would be more relevant to them. Similarly, one resident did not tick a box for the anxiety/depression dimension but rather wrote the note ‘sometimes’. Participants were also unclear about the difference between response options: “Now tell me, what’s the difference between slight and moderate? I really thought slight and moderate would be the same thing.”

One resident spoke about how pain interferes with their walking, yet, ticked ‘no problems’ on the mobility dimension, indicating positive responding, hence giving an answer that is better than an outsider may expect. Positive responding was also an issue for anxiety/depression, where one resident took offense at the question and stated that they have “Nothing to do with that.” Another resident compared themselves to other residents: “No, I’m generally pretty happy. Some of the others in here will be depressed.”

The EQ VAS was problematic for several participants who did not fully understand the instructions: “I’d have to put zero or 100 on this scale? Yes? Or, do I put it down here? No. I put the number here from this scale. See that’s clunky.” One resident questioned its usefulness: “How can

Table 2 Response issue framework

Comprehension	
Unclear instructions ^a	Unclear how to complete the questionnaire
Difficult interpretation	Not understanding what is meant by the item
Item inappropriate*	Item was perceived as inappropriate for residential aged care
Retrieval	
Recall issue ^a	Using the wrong time frame for retrieval
Judgement	
Positive responding	Choosing a more positive answer than an outsider would expect
Reference point issue ^a	Assessment was relative to other people of same age or residing in residential aged care
Response mapping	
Different answers to different aspects	Different response options apply to different aspects of the question
Similar response option	Two response options are similar
No appropriate response option ^a	Available response options did not fully reflect the person’s state or situation

^aThis category was added to the original framework by van Leeuwen et al. [36]

Table 3 Matrix of response issues across the four measures

	Comprehension			Retrieval	Judgment		Response mapping		
	Unclear instructions	Difficult/narrow interpretation	Item inappropriate	Recall	Positive responding	Reference	Different answers to different aspects	Similar response option	No appropriate response option
EQ-5D-5L (<i>n</i> = 12)									
Mobility	////	////			/		/	/	/
Personal care	///						/		/
Usual activities	/		//// //						/
Pain/discomfort	/	//		//			/		//
Anxiety/depression	//	/		/	/	/	//		//
EQ VAS	////		/	//	/	//			
QOL-ACC (<i>n</i> = 12)									
Mobility		//	/						//
Pain management		///							////
Emotional wellbeing	/			//		/			
Independence	/			/		/		/	
Social connection							/		
Activities	/	/							
ASCOT (<i>n</i> = 12)									
Control	/	/				//		/	
Personal comfort and cleanliness				//	/	//			
Food and drink	///	////			/	/			//// //
Personal safety		/		//		/			
Social participation		////	//	//					
Occupation	/		/	/				/	
Home cleanliness		///			/				/
Dignity (having help)		///				//			//
Dignity (treated)		///							//
EQ-HWB (<i>n</i> = 12)*									
Feeling and emotions (12,13,14,15,16)	//	/	//	//// /	///	/	//	/	
Cognition (10,11)				//					/
Self-identity (20)	/	/							
Coping, control, (17,18)	//			//				/	/
Social connections (8,9,19)	//	//		/				/	
Physical sensation (6,7,22,23,24,25)		/		/					
Activity (1,2,3,4,5,21)	////	//// //	///	//// //		//	////	/	//// //

/ indicates issue experienced by one participant

*Numbers in brackets refer to the corresponding questions in the EQ-HWB. Numbers in bold refer to the EQ-HWB-S descriptive system

you measure it in numbers?" and instead suggested to just "ask how their health feels".

3.2 QOL-ACC

The instructions for completing the QOL-ACC were generally clear to all residents, although one resident added text to provide context about why they were not feeling happy or experienced low levels of independence. The mobility

question of the QOL-ACC makes explicit reference to the use of mobility aids (i.e., with the use of mobility aids e.g. wheelchair, walker, stick if you use them). However, two participants who did not use any mobility aids were unclear how to respond to that question: "I am able to get around as much as I like with the use of a mobility aid and wheelchair, walker or stick. I don't have any of them. None of the time. None of the time, that's a matching question for me I suppose, all of the time." As a result, they felt that response

options that exclude reference to mobility aids were missing. Similar problems arose for the pain management question (i.e., *When I experience pain, it is well managed*). Despite the fact that the QOL-ACC refers to pain management rather than pain severity, participants who did not experience pain were unsure how to respond to the question: *"I don't know what to put here...I haven't had any pain, I suppose. I'll put most of the time, I suppose. I don't know how."* One resident selected 'none of the time' referring to levels of pain, rather than whether their pain is managed, which would indicate the worst response option on the QOL-ACC. Another resident noted: *"I haven't got physical pain I've got mental pain; they're not bringing in mental pain"*, suggesting uncertainty in terms of whether the question only refers to physical pain.

The emotional wellbeing question elicited a strong response for two residents who were considering past events, such as moving into residential aged care that left them with feelings of unhappiness. Judgement issues arose for the emotional wellbeing and independence questions, where one resident with dementia was reluctant to tick the best level stating: *"If you've got dementia, you can't be happy all the time"*. No issues were generally observed related to response levels, although one resident struggled: *"Phrases like some, a little and most and stuff [...] people haven't got anything to hang the answer on to, if you know what I mean. What's most of the time?"* For the social connection question that refers to having good social relationships with family and friends, one resident exclusively referred to family, indicating that different response options might be required for relationships with friends compared with family. When prompted to describe the things they associated with activities/hobbies, participants referred to walking or listening to music and reflected on the activities offered inside the residential aged care home. However, one resident spoke about wanting to have more appointments with the physiotherapist.

3.3 ASCOT

Compared with other measures, the questions of the ASCOT are implicitly embedded in the response options. However, this caused response issues where one participant interpreted them as separate questions: *"I don't have as much control as I'd like. That's the same question again. I have as much control over my life as I want. And this one, I have adequate control over my life. I've just said I haven't."* There was also ambiguity in terms of the interpretation of this item, with two residents associating 'control over life' with independent toileting and control over bowel movements. In this context, other residents were often used as reference point: *"I'm lucky, I can go to the toilet when I want. A lot of them can't and it's very sad."*

There were also some double-barrelled questions that required different answers to different aspects of the

question. For example, one resident stated: *"I feel clean but I don't think I can present myself the way I'd like. I wouldn't like to walk out with shirts like this and they don't iron them half the time and not very well when they do."* Similar comment was made for the food and drink question: *"I would cross the drink out. We get coffee and that. But the food is cold."* As a result, participants could not identify an appropriate response option. Also, for the dignity question some participants were unclear how to respond to that question if they were not receiving help, suggesting potential ambiguity and narrow interpretation of the term 'help' that was possibly associated only with help with personal care by some residents.

One resident had problems interpreting the direction of the response options: *"The way I am helped and treated does not affect the way I feel about myself. Now is this meant to be in a bad way?"* Some participants struggled with the dignity question in general and some could not clearly see the difference between question eight and nine (i.e., *how having help to do things makes you think and feel about yourself* versus *the way you are helped and treated makes you think and feel about yourself*). Interpretation issues arose for the food and drink question, where many participants spoke about the quality rather than quantity, including the accommodation of dietary requirements and appropriateness of scheduled meal times, which was not directly reflected in the question. For the social participation question, it was not always clear to participants whether the question was referring to social contacts within the facility or outside, including contacts with family and friends.

Ambiguous interpretation was also noted for the home cleanliness question, with some residents only considering their own room, while others talked about common areas. Responses to some questions were also highly dependent on service provision. Two participants talked about the activities offered in the facility, which influence resident's social situation. This was also an issue for the occupation item, making it hard for residents to distinguish between their ability to do certain things and the services provided that would influence their responses. Judgemental issues arose for questions related to personal comfort and cleanliness, food and drink, and home cleanliness with one resident saying: *"I mean you don't expect your life to be perfect in a place like this, do you?"*

3.4 EQ-HWB

Instructions for the EQ-HWB were generally clear to all residents, although the presence of empty spaces around the tick-boxes invited residents to add written text to provide further context. Despite the fact that the vision item refers to the use of glasses or contact lenses, one resident who normally wears glasses rated their vision without the use of

glasses, indicating the need for separate response options with and without the use of glasses. For the item related to hearing, residents sometimes rated their hearing from the perspective of their family members rather than their own perception of hearing: *“The kids say I had difficulty in hearing. I said, ‘Because you mumble’”*. One resident also admitted to not wearing hearing aids due to concerns about losing them.

Although mobility was perceived as important, the item was problematic for some residents who reported variability in their walking ability depending on whether they were walking outside or inside and which mobility aids they were using. In responding to the question, most respondents took a combined view of their mobility (i.e., inside and outside) but some residents focused naturally on the inside area where they would walk more frequently. Although walking stick and wheelchair are noted as examples, one resident suggested the addition of a ‘walker’ as a mobility aid option: *“I use a walker to get around. It’s not actually a wheelchair, it’s a walker. Do you want me to put that down?”*. There were also different interpretations of what is considered outside and inside, with some residents associating inside with their room and others with the entire facility.

The day-to-day activities item that referred to working, shopping and housework caused response issues where many residents noted the inappropriateness of the examples in the residential aged care context. Participants struggled to think of other day-to-day activities and were unclear how to answer the question as they were no longer doing the activities included as examples: *“I don’t do any of those, shopping or work. No, I don’t do any of that. Should I put not working? Not applicable or something”*. This then brought up the issue of not adhering to the recall period, as residents started to reflect back on the past when they used to do those activities. When residents were asked to comment on the recall period in general, one resident perceived the 7-day recall period as too long: *“I think the last seven days in a place like this, most people can’t think back one day”*, whereas another resident suggested ‘the last month’ as more appropriate.

Problems were also noted for the self-care question that asked for the level of difficulty with washing, using the toilet, getting dressed, eating or caring for appearance, where the level of difficulty varied for many residents across these activities: *“I dress myself. I do have difficulty getting my socks and shoes on [...] Showering, I have to have someone with me to shower me and I’ve got a seat to sit on”*. While some answered the question assuming that assistance of aged care staff is included in the perceived level of difficulty, others assumed the question referred to the level of difficulty if they were to undertake the tasks on their own. The question regarding the perceived support from other people created ambiguity, with some residents only considering the staff, others spoke about family members and some residents also

referred to other residents. Participants also struggled with the safety question, which refers to the examples fear of falling, physical harm, and abuse. Although fear of falling was often stated as relevant, many felt that abuse was not applicable to them, which made responding to this item difficult as there was only one response option: *“Yeah, see this one it’s fear of falling which is physical harm. I believe that this should be separated. If I cross out abuse and go sometimes”*.

One participant also noted that they perceived sadness and depression as different constructs and would answer them differently: *“I think they’re different for me. I often feel sad because I look around and I look at all the photos. I don’t know whether it’s depression”*. Although the EQ-HWB has separate questions for pain and discomfort, when answering the pain question, one resident spoke about itchiness, which is explicitly mentioned in the discomfort item but was considered under pain. Other examples for discomfort in the EQ-HWB include feeling sick or breathless, yet, when prompted to describe what else residents associate with the word discomfort, they referred to being uncomfortable in bed, not feeling comfortable with yourself, or not feeling yourself. The latter two references tap into emotional discomfort despite the reference to ‘physical discomfort’ in the EQ-HWB. One participant queried: *“Wouldn’t physical discomfort mean pain?”*. In terms of the response options, one resident was unclear about the difference between ‘slight difficulty’ and ‘some difficulty’; another resident did not see the difference between ‘only occasionally’ and ‘sometimes’.

4 Discussion

This study aimed to examine the face validity of four preference-weighted QoL measures in older people living in residential aged care. We identified response issues across all measures. The most common issues related to comprehension and response mapping issues, often due to double-barrelled and ambiguous items that may be subject to different interpretations in a residential aged care context. We also identified issues related to understanding instructions, including adherence to the recall period, and noted positive responding that requires attention when interpreting the data. These issues will impact responses and the validity of measures, leading to potentially biased scores.

Compared with the other measures, the QOL-ACC had fewer response issues, which may reflect that it was co-designed from its inception with older people using aged care services. Response issues mainly arose for the mobility question and the pain management question when participants were not using any mobility aids or did not experience pain, resulting in uncertainty about how to respond to those questions. During the development of the QOL-ACC, a previous study examined the face validity of candidate

items to inform the exclusion of items that were flagged as problematic [39]. Although the final six domains included in the QOL-ACC were found to be acceptable, our study identified some issues that require further research and potential modifications to the mobility and pain management questions to improve clarity.

Two previous qualitative think-aloud studies with older people found that the EQ-5D-3L and EQ-5D-5L had fewer response issues compared with other measures (i.e., ASCOT, ICECAP-O, WEMWBS SF-12 and ONS-4) [36, 40]. Our study noted response issues especially related to uncertainty around the use of mobility aids, varying need for assistance with personal care, inappropriate examples for usual activities resulting in recall issues, use of drugs for pain/discomfort and anxiety/depression, and the conflation of two concepts within one question. Issues arising from the double-barrelled EQ-5D questions have previously been noted [41, 42]. While similar response issues were also noted in the previous think-aloud studies, we also found that the instructions and the layout were not always clear to participants. This was especially an issue for the EQ VAS. A recent systematic review of the feasibility of the EQ-5D measures in the older adult population found that the EQ VAS resulted in more missing values compared with the five dimension-based questions, indicating that it is more difficult to comprehend and poses a higher burden on older participants [43]. This raises the question as to whether there is a need to expand the suite of EQ-5D instruments by developing older-adult-specific versions of the EQ-5D instruments, similarly to the development of child-friendly versions (i.e., EQ-5D-Y [-3L or -5L] [44]) to improve appropriateness for use in older adults. However, any modifications to the original instruments would require more research and further validation studies, as well as new health state valuation studies, which are costly and labour intensive. Although it may not be necessary to have preference weights accompanying the instrument if the purpose is not for use within an economic evaluation, there are still remaining issues around comparability and challenges around transitions between age-specific instruments.

Relative to the length of the measures (i.e., 5 items versus 25 items), the EQ-HWB resulted in fewer response issues than the EQ-5D-5L, suggesting better face validity in residents living in residential aged care. However, the activity domain was most problematic and could be improved by (i) separating questions around mobility inside and outside and defining those terms; (ii) providing more appropriate examples for day-to-day activities for residents living in aged care; and (iii) separating self-care questions, including clarification around how to factor in assistance for self-care. Further attention is also required to improve clarity of the four pain and discomfort questions that measure frequency and severity separately, ensuring that participants interpret pain and

discomfort as two separate constructs and consider only physical discomfort. The inclusion of a large number of questions related to feelings and emotions is important but could result in positive responding as found in this study. Given that the EQ-HWB is currently designed as an experimental version, it is worth exploring further modifications to the instrument to improve appropriateness and relevance for older adults in residential aged care. Alternatively, there is scope for developing a residential-specific version of the instrument.

With regard to the ASCOT, we also identified a number of response issues, especially related to double-barrelled questions, which were previously found in another qualitative study of people living with dementia and their carers [29]. There was also some ambiguity around certain words, such as social contacts, which could either refer to social contacts outside or inside the facility. Although this could reflect intentional ambiguity by the instrument's developers to allow interpretation based on personal preferences and context [32], it becomes problematic if a resident's responses are not reflective of their true situation due to narrow or different interpretation to those interpreting the data. Additionally, we found that many participants struggled with the dignity questions. This aligns with findings from a previous study, which found that the question was poorly understood, as participants indicated they did not understand what was meant by the questions and did not see a connection between receiving help and the way they feel and think about themselves [36]. Although the ASCOT was designed to capture the effect of service provision on people's lives, we found that residents struggled with some questions and were unable to differentiate between their functional abilities to do certain things and the opportunities offered to them. This was observed, for example, for the occupation question, which could have been caused by the inconsistent use of the term 'able' in response options (level 4 'I don't do anything I value or enjoy with my time' vs level 1 'I'm able to spend my time as I want, doing things I value or enjoy'), where the term 'able' possibly led some respondents thinking about what they would be able to do on their own, rather than thinking about their capabilities (their ability to function in a particular way, whether or not they choose to do so) as intended by the ASCOT measure. While this could have been made clearer in the instructions and response options in the ASCOT, measuring capabilities through questionnaires remains an ongoing challenge and an area for further research [45, 46]. We also noted positive responding to some ASCOT items that are directly linked to quality of care, such as food and drink, as well as home cleanliness and safety, which could be due to fear of repercussions if negative responses are given. The use of the ASCOT in residential aged care, therefore, may warrant additional attention to this potential bias.

Positive responding has been observed as an issue across all measures. Participants often assessed their health relative to other people of their age or residing in the aged care facility who were worse off, thereby lowering their benchmark for good QoL and responding more positively. This was found in two previous qualitative think-aloud studies of older people [36, 40]. Positive responding was particularly observed for psychological aspects of QoL, which is in line with a previous study that found that higher levels of psychological distress (e.g., depressive symptoms and anxiety) were reported by younger patients with heart failure when compared with older people, due to the ability of older patients to reconceptualise or change their expectations for QoL in the context of heart failure [47]. This concept of response shift is also commonly observed in older people responding to patient-reported outcome measures, impacting the validity of responses and comparisons to other groups [48, 49].

A somewhat surprising finding was the high proportion of response issues related to instructions, which were also driven by the layout of the respective instruments. Instrument development has largely focused on content validity and respective psychometric performance of the items. However, exploring respondents' understanding of instructions is equally important, which could otherwise result in missing or invalid responses. Participants in this study were asked to self-complete the measures. These issues could possibly be avoided via interviewer administration. However, in view of the recent introduction of the National Aged Care Mandatory Quality Indicator Program in Australia, where self-completion of the QOL-ACC is recommended by all care recipients with capacity (e.g. care recipients with no or mild cognitive impairment) [3], further exploration and improvements are needed in this area.

4.1 Limitations

Our sample mainly comprised residents who reported English as their primary language; further exploration of these measures in culturally and linguistically diverse samples is warranted. All participants also had sufficient cognitive capacity to provide an informed consent and future studies could explore the association between cognitive status and response issues. Although all residents were asked to compare the two measures after the interview, only a small sample of participants were able to state which measures they preferred. When prompted to elaborate on their choice, participants could not justify their choice, recognising the complexity of this task. The use of randomisation of measures also means that all four measures were not explored by the same group of residents. The measures explored in this study were selected based

on their relevance for the Australian aged care context; other measures exist that were not explored in the study, including the ICECAP-O [50], which cannot be used to generate QALYs, or the more recently developed WOOP [51]. The study was conducted in Australia; results may or may not be generalisable to use of these instruments in other countries (even if English speaking). Our analysis focused on the face validity of the measures; we did not convert participant's responses into preference-weighted utility scores. Future studies are warranted exploring the resulting utility scores, including the psychometric assessment of the four measures in terms of convergent validity, reliability and responsiveness to change.

5 Conclusion

The increased demand for preference-weighted measures in residential aged care to guide resource allocation decisions and quality-of-care assessments requires further evidence on instruments' acceptability and performance. This study provided findings on the face validity of the EQ-5D-5L, EQ-HWB, ASCOT and QOL-ACC, indicating numerous response issues that require further attention. To improve validity of those measures and avoid confusion or interpretation issues, our study findings provide a number of avenues where modifications to existing measures can be made. This includes improvements to the instructions, clarification around how to factor the use of health aids and assistance from others into the responses, provision of examples for activities of daily living that are more appropriate for older adults, and the avoidance of double-barrelled and ambiguous questions. Our results serve also as a guide to inform the development of new measures for use in older people residing in aged care homes.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40271-023-00647-6>.

Acknowledgments We would like to acknowledge the helpful and constructive comments received by two anonymous reviewers on a previous version of our manuscript. We are also very grateful to the study participants and the three participating residential aged care facilities.

Declarations

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Funding for this study was obtained from the EuroQol Research Foundation (150-RA); Lidia Engel is currently supported by a National Health and Medical Research Council (NHMRC) Investigator EL1 grant.

Conflicts of Interest Authors Lidia Engel, Nancy Devlin, Brendan Mulhern, Tessa Peasgood and Rosalie Viney are members of the Eu-

roQoL group that developed the EQ-5D-5L and EQ-HWB discussed in the study.

Data Availability The data generated in the current study are not publicly available but are available from the corresponding author on reasonable request.

Ethics and Consent This study was approved by the Monash University Human Research Ethics Committee (Project ID: 32170) and all participants provided a written consent prior to the interviews.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by LE and LK. The first draft of the manuscript was written by Lidia Engel and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Consent to participate All participants provided a written consent prior to participating in the interviews.

Consent for publication All participants were informed in the Plain Language Statement that findings of this study will be published in a peer-reviewed journal article.

Code availability Not applicable.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

1. World Health Organization. Ageing and health, Retrieved 28.08.2023, from. 2022. [https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=By%202030%2C%201%20in%206,will%20double%20\(2.1%20billion\)](https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=By%202030%2C%201%20in%206,will%20double%20(2.1%20billion)). Accessed 28 Aug 2023.
2. Organization for Economic Cooperation and Development (n.d.). *Ageing and Long-term Care*, Retrieved 28.08.2023, from. <https://www.oecd.org/els/health-systems/long-term-care.htm>. Accessed 28 Aug 2023.
3. Aged Care Quality and Safety Commission. National Aged Care Mandatory Quality Indicator Program, Retrieved 28.08.2023, from. 2023. <https://www.agedcarequality.gov.au/providers/national-aged-care-mandatory-quality-indicator-program>. Accessed 28 Aug 2023.
4. Ratcliffe J, et al. Valuing the Quality-of-Life Aged Care Consumers (QOL-ACC) Instrument for Quality Assessment and Economic Evaluation. *Pharmacoeconomics*. 2022;40(11):1069–79.
5. Care quality commission. A new start. Responses to our consultation on changes to the way CQC regulates, inspects and monitors care services., Care Quality Commission, Retrieved 28.08.2023, from. 2013. https://www.cqc.org.uk/sites/default/files/documents/cqc_consultation_2013_tagged_0.pdf. Accessed 28 Aug 2023.
6. Towers AM, et al. A cross-sectional study exploring the relationship between regulator quality ratings and care home residents' quality of life in England. *Health Qual Life Outcomes*. 2019;17(1):22.
7. Cardona B. Measuring outcomes of community aged care programs: challenges, opportunities and the Australian Community Outcomes Measurement ACCOM tool. *Health Qual Life Outcomes*. 2018;16(1):104.
8. Hutchinson C, et al. The integration of mixed methods data to develop the quality of life - aged care consumers (QOL-ACC) instrument. *BMC Geriatr*. 2021;21(1):702.
9. Khadka J, et al. Assessing feasibility, construct validity, and reliability of a new aged care-specific preference-based quality of life instrument: evidence from older Australians in residential aged care. *Health Qual Life Outcomes*. 2022;20(1):159.
10. Khadka J, et al. Assessing the construct validity of the Quality-of-Life-Aged Care Consumers (QOL-ACC): an aged care-specific quality-of-life measure. *Qual Life Res*. 2022;31(9):2849–65.
11. Netten A, et al. Outcomes of social care for adults: developing a preference-weighted measure. *Health Technol Assess*. 2012;16(16):1–166.
12. Malley JN, et al. An assessment of the construct validity of the ASCOT measure of social care-related quality of life with older people. *Health Qual Life Outcomes*. 2012;10:21.
13. Towers AM, et al. The acceptability and feasibility of using the Adult Social Care Outcomes Toolkit (ASCOT) to inform practice in care homes. *BMC Health Serv Res*. 2016;16(1):523.
14. Brazier J, et al. A review of generic preference-based measures for use in cost-effectiveness models. *Pharmacoeconomics*. 2017;35(Suppl 1):21–31.
15. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013, Retrieved 28.08.2023, from. 2013. <https://tinyurl.com/bkex73sw>. Accessed 28 Aug 2023.
16. Pharmaceutical Benefits Advisory Committee. Guidelines for preparing a submission to the PBAC, Version 5.0, Commonwealth of Australia Department of Health, Retrieved 28.08.2023, from. 2016. <https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf>. Accessed 28 Aug 2023.
17. Zorginstituut Nederland (2016). Guideline for economic evaluations in healthcare, Zorginstituut Nederland, Retrieved 28 Aug 2023, from. 2016. <https://tinyurl.com/rd8kksnc>. Accessed 28 Aug 2023.
18. Wisloff T, et al. Estimating QALY gains in applied studies: a review of cost-utility analyses published in 2010. *Pharmacoeconomics*. 2014;32(4):367–75.
19. Bulamu NB, et al. A systematic review of instruments for measuring outcomes in economic evaluation within aged care. *Health Qual Life Outcomes*. 2015;13:179.
20. Cleland J, et al. A review of the development and application of generic preference-based instruments with the older population. *Appl Health Econ Health Policy*. 2019;17(6):781–801.
21. Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;37(1):53–72.
22. Herdman M, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–36.
23. Milte CM, et al. How important is health status in defining quality of life for older people? An exploratory study of the views of older South Australians. *Appl Health Econ Health Policy*. 2014;12(1):73–84.

24. Brazier J, et al. The EQ health and wellbeing: overview of the development of a measure of health and wellbeing and key results. *Value Health*. 2022;25(4):482–91.
25. Mukuria C, et al. Valuing the EQ Health and Wellbeing Short (EQ-HWB-S) using time trade-off and a discrete choice experiment: A feasibility study. *Value Health*. 2023;26(7):1073–84.
26. Carlton J, et al. Generation, selection, and face validation of items for a new generic measure of quality of life: the EQ-HWB. *Value Health*. 2022;25(4):512–24.
27. Peasgood T, et al. Developing a new generic health and wellbeing measure: psychometric survey results for the EQ-HWB. *Value Health*. 2022;25(4):525–33.
28. Mokkink LB, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737–45.
29. Engel L, et al. A qualitative exploration of the content and face validity of preference-based measures within the context of dementia. *Health Qual Life Outcomes*. 2020;18(1):178.
30. Terwee CB, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27(5):1159–70.
31. Personal Social Services Research Unit (University of Kent). Introduction to ASCOT, Retrieved 18.08.2023, from. 2018. <https://www.pssru.ac.uk/ascot/>. Accessed 28 Aug 2023.
32. Netten A, et al. Quality of life outcomes for residents and quality ratings of care homes: is there a relationship? *Age Ageing*. 2012;41(4):512–7.
33. Willis GB, Artino AR Jr. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *J Grad Med Educ*. 2013;5(3):353–6.
34. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res*. 2003;12(3):229–38.
35. Miller K, et al. *Cognitive interviewing methodology : a sociological approach for survey question evaluation*. Somerset: Wiley, Incorporated; 2014.
36. van Leeuwen KM, et al. Exploration of the content validity and feasibility of the EQ-5D-3L, ICECAP-O and ASCOT in older adults. *BMC Health Serv Res*. 2015;15:201.
37. Tourangeau R, et al. *The Psychology of Survey Response*. Cambridge: Cambridge University Press; 2000.
38. Mukuria C, et al. Qualitative review on domains of quality of life important for patients, social care users, and informal carers to inform the development of the EQ-HWB. *Value Health*. 2022;25(4):492–511.
39. Cleland J, et al. From the ground up: assessing the face validity of the Quality of Life—Aged Care Consumers (QOL-ACC) measure with older Australians. *Quality in Ageing and Older Adults*. 2023;24(1/2):3–18.
40. Penton H, et al. A qualitative investigation of older adults' conceptualization of quality of life and a think-aloud content Validation of the EQ-5D-5L, SF-12v2, Warwick Edinburgh Mental Wellbeing Scale, and Office of National Statistics-4. *Value Health*. 2022;25(12):2017–27.
41. Belay YB, et al. Examining the psychometric properties of a split version of the EQ-5D-5L anxiety/depression dimension in patients with anxiety and/or depression. *Qual Life Res*. 2023;32(7):2025–36.
42. Engel L, et al. What is measured by the composite, single-item pain/discomfort dimension of the EQ-5D-5L? An exploratory analysis. *Qual Life Res*. 2022;32(4):1175–86.
43. Marten O, et al. Feasibility of the EQ-5D in the elderly population: a systematic review of the literature. *Qual Life Res*. 2022;31(6):1621–37.
44. Devlin N, et al. EQ-5D-Y-3L value sets, valuation methods and conceptual questions. *Pharmacoeconomics*. 2022;40(Suppl 2):123–7.
45. Karimi M, et al. The capability approach: a critical review of its application in health economics. *Value Health*. 2016;19(6):795–9.
46. van Loon MS, et al. Quality of life in a broader perspective: Does ASCOT reflect the capability approach? *Qual Life Res*. 2018;27(5):1181–9.
47. Moser DK, et al. "It could be worse ... lot's worse!" Why health-related quality of life is better in older compared with younger individuals with heart failure. *Age Ageing*. 2013;42(5):626–32.
48. Ubel PA, et al. What is perfect health to an 85-year-old?: evidence for scale recalibration in subjective health ratings. *Med Care*. 2005;43(10):1054–7.
49. Vanier A, et al. Response shift in patient-reported outcomes: definition, theory, and a revised model. *Qual Life Res*. 2021;30(12):3309–22.
50. Coast J, et al. Valuing the ICECAP capability index for older people. *Soc Sci Med*. 2008;67(5):874–82.
51. Himmler S, et al. Estimating an anchored utility tariff for the wellbeing of older people measure (WOOP) for the Netherlands. *Soc Sci Med*. 2022;301: 114901.