Challenges Conducting Cognitive Interviews in Low and Middle Income Countries: A Case Study with Older Adults in Lebanon

Survey Methods: Insights from the Field, Volume 3, issue 2 (2025)

Alexandra Abi Nassif | Mayssan Kabalan | Carlos Mendes de Leon | Julie de Jong | Frederick Conrad * | | * Corresponding author | Note: A. Abi Nassif & M. Kabalan contributed equally to this work.

How to cite this article : Abi Nassif, A., Kabalan M., de Jong J., Mendes de Leon C. & Conrad F. (2025). Challenges Conducting Cognitive Interviews in Low and Middle Income Countries: A Case Study with Older Adults in Lebanon. Survey Methods: Insights from the Field. Retrieved from https://surveyinsights.org/?p=21128

DOI : 10.13094/SMIF-2025-00009

Copyright : © the authors 2025. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

Abstract : The global increase in population ageing, especially in Low and Middle Income Countries (LMICs), has created greater need for population-based surveys that document the health and social requirements of older adults in this context. The surveys fielded in response to this need ask questions adapted from ongoing surveys in High Income Countries; hence it is important to pretest the adapted questions for interpretability and cultural appropriateness using methods such as cognitive interviewing. The aim of this paper is to highlight challenges in the administration of cognitive interviews in LMICs with older adults, based on our pretest of a new population survey study in Lebanon. We analysed 58 cognitive interviews conducted with older adults, and identified two sets of challenges: population-specific and ageing-specific challenges. The first set includes trouble understanding the purpose of cognitive interviewing, respondents' suspicion of institutions, reluctance to disclose sensitive information, and problems arising from diglossia, i.e., languages whose official, written form differs substantially from the colloquial, spoken form. The second set concerns age-related attributes such as hearing problems, fatigue, and decline in cognitive abilities. Based on our findings and experience, we offer recommendations and best practices for conducting cognitive interviews in LMICs with older adults.

1. Introduction

The population of older adults is growing worldwide, especially in low- and middle-income countries (LMICs) (Sudharsanan & Bloom, 2018; UNDESA, 2020). To document the health and social needs in these populations, an increasing number of population-level survey studies has been established, modelled after the Health and Retirement Study (HRS) in the United States, and connected with one another in the network of the Health and Retirement Studies Around the World (*HRS Around the World*, n.d.).

One such survey is a new population-based study of ageing and dementia in Lebanon: the Lebanon Study on Aging and HeAlth (LSAHA). Many of the questions and instruments in the LSAHA survey have been

adapted from HRS, making it necessary to assess the interpretability and cultural appropriateness of the draft questionnaire for this population (Willis, 2015). To that end, LSAHA designed and implemented cognitive interviews. Cognitive interviewing is one of the most widely used survey pretesting methods, helping researchers evaluate the extent to which respondents understand the questions as intended and are able to carry out the response tasks required by each question (Beatty & Willis, 2007; Willis, 2004). However, application of cognitive interviewing in LMICs has been limited to a narrow range of topics, such as nutrition and program participation (Andrew et al., 2022; Ashok et al., 2022; Namaste et al., 2024), and studies have noted the need for careful cultural adaptation of the cognitive interviewing process to ensure the task is well-understood among populations in LMICs (Martin et al., 2017; Melgar et al., 2022; Sarkar & de Jong, 2024) as well as among diverse populations in western contexts (Park et al., 2016; Park & Goerman, 2019). To date, there has been negligible use of cognitive interviewing to assess survey instruments with questions related to ageing and dementia: a recent meta-analysis (Lay et al., 2024) indicates that nearly all relevant studies occurred in western contexts and there is limited evidence the method has been used when adapting instruments to the cultural context of LMIC populations (Fasfous et al., 2017; Franzen et al., 2020; Zeinoun, 2022), despite the opportunities it offers for improvements to the quality of adapted instruments.

In the current article, we address the challenges of conducting cognitive interviews with older adults in LMICs. Researchers may be particularly likely to experience these challenges in population groups that are relatively unfamiliar with systematic survey research, or have low “survey literacy” (Massey, 2025), such as older adults in non-western countries where survey research has been rare. We focus on two sets of challenges in the administration of cognitive interviews in LMICs that we believe have received insufficient attention. Based on our findings, we offer some practical guidelines for conducting cognitive interviews with older, LMIC populations around the globe.

One set of challenges was *population-specific* in that they were related to the administration of the LSAHA draft questionnaire in Lebanon. The first of this set of challenges concerned the discrepancy between the official and everyday languages used in the respondent’s world. The official language in Lebanon is Arabic, but the colloquial Arabic spoken there and throughout the Arab world differs in many respects from official Arabic. In fact, many Lebanese, especially those with limited formal education, are not familiar with or comfortable communicating in formal Arabic. As Abou-Mrad et al. (2015) note, language and cognition can be highly interrelated (so-called linguistic relativity), making it essential to conduct cognitive interviews in the respondent’s everyday language to accurately measure the respondent’s thinking while they answer draft questions. Therefore, despite having formally translated the draft questionnaire prior to pretesting it (described below), we uncovered many residual translation issues in the cognitive interviews. A second population-specific challenge was respondents’ lack of survey literacy, i.e., familiarity with systematic survey research, let alone the reasons one would pretest a survey questionnaire. A third population-specific challenge involved a general sense of mistrust or suspicion of governmental agencies and activities (Alijla, 2016; Darwish, 2019), which may result in a reluctance or refusal to respond to sensitive questions, whether administered as part of a cognitive interview or in the actual survey. Note that this reluctance to disclose sensitive information differs from socially desirable responding (e.g., Tourangeau & Yan, 2007) which is usually attributed to the presence of an interviewer as opposed to governmental institutions.

The second set of challenges was *ageing-specific* in that they concerned older adults facing difficulties participating in cognitive interviews. Older adults often experience health changes that can affect their ability to participate in survey interviews (e.g., Schwarz et al., 1998) and presumably cognitive interviews, which rely on the respondent’s ability to reflect and elaborate upon their internal response processes (Willis, 2004). The first ageing-specific challenge we report is sensory, specifically hearing impairments

(Lin et al., 2011) although older adults are also at increased risk of visual impairments (Varma et al., 2016), both of which can complicate the administration of cognitive interviews. The second ageing-related challenge we discuss concerns cognitive decline. Compared to their younger counterparts, older adults are likely to experience reduced ability to recall information from long term (episodic) memory and to maintain information in working (short term) memory, as well as in overall executive functioning (Craik, 1999; Hartshorne, & Germine, 2015; Salthouse, 2019; Wingfield, 1999). They commonly experience difficulties in language understanding and speech production such as with word-retrieval (Burke & Shafto, 2004) and object naming (Harada et al., 2013; Kemper, Thompson, & Marquis, 2001). Communication difficulties can be exacerbated by hearing problems, reduced working memory capacity, and generally slower information processing (Martínez-Nicolás, Llorente, Ivanova, Martínez-Sánchez, & Meilán, 2022; Kim, & Oh, 2013; Yorkston, Bourgeois, & Baylor, 2010). Older adults may experience further challenges in comprehension and communication when they find themselves in situations that are stressful or cognitively demanding (Nadeau et al., 2019; Caruso, Muller, & Shadden, 1995). To compound these difficulties,there is evidence that older respondents may compensate for such challenges in cognitive interviews by providing an answer regardless of whether they really understood the questions or follow-up probes (Hall & Beatty, 2014; Mallinson et al., 2002). Finally, older adults tire more quickly (Avlund, 2010), which may affect participation throughout an entire cognitive interview even more than in actual survey interviews given the additional tasks beyond answering questions in the former, such as answering probe questions or being asked to paraphrase a question.

To document the nature and scope of these two sets of challenges, we performed a detailed review of the audio recordings of 58 cognitive interviews we conducted as part of the pretest of the LSAHA survey questionnaire.

2. Methods

2.1. Main Study

LSAHA is a population based survey designed for a sample of 3,000 adults aged 60 years and older, randomly sampled from two locations in Lebanon: Beirut (urban) and Beqaa (rural). The LSAHA survey instrument consists primarily of standardised questions and tests adapted from other studies, covering topics such as socio-economic circumstances, labor force and pension status, cognitive and physical functioning, medical history, mental health and functional health status, social networks and support, and other life course events such as political violence. Translation of the survey instrument from English to Arabic followed an iterative process with a team of local experts, consistent with the TRAPD framework (e.g., Behr, 2023).

2.2. Cognitive Interviewing

Between August and October 2022, 58 cognitive interviews were conducted with a convenience sample of respondents. The primary purpose of the pretest was to explore whether respondents understood the survey questions as intended by asking them to provide real time commentary on their thinking while responding. The fidelity of the translation was not the purpose of the pretest per se but a secondary purpose was to assess cross-cultural equivalence of the survey questions which had been developed for administration in HICs. This concerns, for example, whether questions might be more or less sensitive in Lebanon than in western countries.

Due to the large number of items in the LSAHA questionnaire, cognitive interview respondents were presented with one of two non-overlapping scripts (sets of questions), each based on around 50 questions (not counting the probes). Interviewers administered each script in person to roughly half (27 and 31, respectively) of the respondents (see Appendix), who received an incentive of 5 USD. The interviews lasted between 30 and 65 minutes and were audio-recorded for analysis. Both the questions and probes were written and administered in formal Arabic to ensure a standardized script across the two study areas in Lebanon which have distinct dialects, and to maintain consistency for broader use across Arab countries.

The interviewers had bachelor's degrees or advanced training in social sciences or public health and were familiar with the study objectives. The six interviewers were trained and certified by a survey methodologist from the University of Michigan with extensive experience in cognitive interviewing in different settings around the world.

Using a quota sampling approach, we stratified the large sample by age, sex and education, recruiting participants from both study areas (Beirut or Beqaa) so that the sample would reflect the diversity of the target population (see Table 1). While considerably smaller samples are the norm in cognitive interview pretests, larger samples have been shown to reveal proportionally more unique problems than smaller ones (Blair & Conrad, 2011).

2.3. Analysis

To classify problems for purposes of improving the questionnaire – not the focus of the current article – the research team developed a problem coding system based on three of the four stages of the survey response model proposed by Tourangeau et al. (2000): Comprehension, Retrieval, and Response, omitting the fourth response stage, *Judgment,* since problems at that stage are not easily detectable in cognitive interviewing (Conrad & Blair, 2009). We later added a fourth category, *Process*, to capture difficulties respondents experienced in understanding and engaging with the cognitive interview process. By listening to the audio recordings, each interviewer coded each problem that respondents seemed to experience, assigning it to one of the four categories, and provided a qualitative description of each problem. These were used to identify and address problems in the initial draft of the LSAHA questionnaire which was subsequently administered in the main survey interviews.

By reviewing the qualitative descriptions of problems for the current article, we observed two overarching themes or types of challenges: population-specific challenges and ageing-related challenges. Subsequently, one of the authors listened to the audio recordings to confirm that these challenges were evident and extract relevant examples. Note that in this article we do not analyze how the coded problems were addressed to improve question wording.

2.4. Characteristics of Respondents

The distribution of respondents by education and gender appear in the top panel of Table 1 and by education and age in the lower panel of the table. Overall, we achieved a reasonable diversity of background characteristics; the smaller percentage of respondents in the high education category corresponds in relative terms to their smaller proportion in the target population.

Table 1: Distribution of respondents by sex, age, and educational attainment

	Education			
	Low (grade 5 or less)	**Medium** (grades 6-12)	**High** (at least some university education)	**Total**
Total	23 (40%)	25 (43%)	10 (17%)	58
Gender				
Male	9	15	7	31 (53%)
Female	14	10	3	27 (47%)
Age				
< 70 (57-69)	12	17	7	36 (62%)
70+ (70-85)	11	8	3	22 (38%)

3. Results

The results are structured into two sets of methodological challenges: Population-specific challenges and Ageing-related challenges conducting cognitive interviews in LMICs with older adults.

3.1 Population-specific challenges

3.1.1. Problems understanding the purpose of cognitive interviewing

The focus here is on a phenomenon that is commonly observed, at least anecdotally, in cognitive interviews: respondents are unfamiliar with cognitive interviewing, especially tasks such as thinking aloud and paraphrasing questions (Willis, 2004), and thus do not understand the purpose of conducting cognitive interviews. Such lack of familiarity with the cognitive interviewing enterprise is of particular concern in the current study because respondents from an LMIC such as Lebanon often have no experience with survey research (lack survey literacy), let alone pretesting a questionnaire. Moreover the increased likelihood of cognitive decline in older adults may well exacerbate their ability to make sense of what they are being asked to do and why.

Although the interviewers explained the purpose of the interviews, many respondents seemed confused by what they were asked to do. For example, rather than reporting on their understanding of questions, a number of respondents just answered the questions, i.e., did not provide information about their comprehension of the questions or the process of coming up with answers. Consider two examples:

I: Has a doctor ever told you that you have high cholesterol?

R: No, I don’t.

I: Can you please let me know, using your own words, what is this question asking about?

R: About what? No, I don’t have cholesterol, I have ulcers.

I: When wearing glasses, is your eyesight good, very good, fair…

R: I used to wear glasses, however, I had surgery.

I: So it’s excellent now?

R: Yes. Thank God.

I: Can you please let me know using your own words, what is this question asking about?

R: For the eyesight, you mean?

I: Yes.

R: Maybe about my health or something, I don’t know. What does this question mean to you?

Similarly, when respondents were probed about their interpretation of a draft question, many of their responses contained mostly irrelevant content, suggesting that they did not understand that they were asked to reflect on how they came up with their answer:

I: What does the phrase “consider this their primary household” mean to you?

R: Us Lebanese, we have a lot, a lot of love for the family […] we help each other to a great extent. […] I have a lot of love for these customs and traditions, they are ingrained in my mind because of my parents and grandparents and the people in my village, they all love these things. Because of that, it means a lot to me.

I: What does the phrase “stooping, kneeling, or crouching” mean to you?

R: Right now, it means difficulty. But yesterday I was at the doctor’s and he said, “when you’re done with taking cortisone, you’ll be back to normal”.

[…]

I: Is there another word to explain [stooping or kneeling]?

R: It means I worry. I worry because of the condition I’m in.

Both respondents appear to have been well intentioned but seemed to misunderstand the purpose of the probe.

3.1.2. Problems arising from suspicion of institutions and reluctance to disclose sensitive information

Respondents exhibited suspicion about the purpose of the cognitive interview. While not unique to respondents from Lebanon, suspicion of institutions in LMICs is likely more common than in HICs which have more stable governments, higher survey literacy, and less worry about governmental agencies surveilling citizens. To the extent such suspicion is common in Lebanon, it would likely apply to many other countries with high levels of suspicion toward governmental institutions (Arab Center Washington DC, 2023). Suspicion of institutions also likely compounds the widely observed reluctance to disclose sensitive or stigmatised information in surveys, especially in in-person interviews (e.g., Schaeffer, 2000; Tourangeau & Smith, 1996; Tourangeau & Yan, 2007).

For example, we asked to see an ID to confirm respondents' age, but some suspected that this was done to identify their religion, since this can sometimes be deduced from the information on the ID. Religious identification has been a very sensitive issue for many decades in Lebanon. One respondent questioned why we wanted to see his ID and only complied when he was assured that it would not be shared with anyone.

Some respondents speculated that there was a covert goal or purpose in asking questions on topics such as religious attendance and identity, which are sensitive in high income, western countries, but potentially threatening in a country like Lebanon with such a fraught history of religious conflict (see Brenner, 2017). For instance, when asked to convey what they have understood from the following question "Apart from funerals, how often do you go to a mosque [church (for Christian respondents)] these days?", seven respondents were concerned that the question was intended to measure their level of faith; in fact, the question was one of several used to measure social participation, specifically using frequency of attending religious services. This approach was modeled on other HRS Around the World studies (Howrey & Hand, 2019), allowing us to maintain comparability between LSAHA survey data and those studies.

> **R1**: Why are you asking us if we went to the mosque or not?
>
> **I**: We're checking if you're engaged in any activity and leaving the house.
>
> **R2**: Is this question purely sectarian?
> **I**: No, with this question we're seeing if you're getting out of the house and meeting people. It's not about sectarianism.

In addition to questions about religion, questions about war and political violence, a topic often intertwined with religious sectarian conflicts in Lebanon were, unsurprisingly, sensitive.

> **I**: Did you lose a parent or a child because of the war/conflict/explosions up until today? [...] How did you feel answering this question?
>
> **R**: I wasn't comfortable.
>
> **I**: How do you think other people who suffered this loss would feel answering this question?
>
> **R**: Bitter, you know?

Another respondent whose parents died during the Lebanese civil war was emotional answering this same question, saying, “You made me relive the moment.” In addition to suggesting how the survey questions might be improved, such a response can indicate areas where respondents may relive trauma and where additional support resources should be offered at the end of the interview.

In Lebanon, where political violence and sectarian conflict have recurred for decades, questions about war and personal loss are sensitive for the entire population: it is impossible to live in Lebanon without being affected by collective trauma, political sensitivities, and concerns about openly discussing such topics. Asking respondents about these questions required careful wording and framing.

3.1.3. Problems arising from diglossia

Many languages, including Arabic, consist of a “high” version used in formal and ceremonial settings such as education, literature, public speaking, and media, and one or more “low” varieties used in everyday conversation. Linguists refer to such languages as “diglossic” (Ferguson, 2003). In Lebanon, formal Arabic differs substantially in vocabulary and grammar from colloquial, spoken Arabic which in turn varies substantially across Arab countries. Over and above inaccuracies introduced by translating from English into formal Arabic, the gulf between formal written and colloquial spoken forms of Arabic may cause problems for respondents.

The draft questions, translated from English into formal Arabic, were not well understood by many respondents who are used to communicating in a local dialect of Arabic. Some questions required the interviewer to reword or further explain the question, demonstrating that administering questions in the formal Arabic translation can harm standardisation:

> **I**: How often do you feel uncertain about how to care for him?
>
> **R**: What do you mean uncertain?
>
> **I**: You’re not sure how to help him.

The word “uncertain” used in formal Arabic here (transliteration: *aadam al-yakin*) does not exist in the lexicon of everyday Lebanese Arabic, which, instead, uses “*mesh meta’aked*”.

A similar example is this respondents’ lack of famliarity with the word “crouching” in formal Arabic (transliteration: *jouthoum*). Respondents were not familiar with the word, although they did know the concept. One of the few respondents who did understand it was an Arabic teacher. The word was replaced with one commonly used in spoken Arabic (transliteration: *karfasa*).

> **I**: And if I say *jouthoum*, is that a word you know?
>
> **R**: What is *jouthoum*? You don’t know it either.
>
> **I**: […] It’s like *karfasa*
>
> **R**: Oh, oh, yeah, yeah, *karfasa*, yeah.

The struggle to understand formal Arabic was not solely due to unfamiliarity with particular words. In fact, respondents faced difficulties with both commonly used words and phrases as well as those that are less well known. This may be explained by the fact that the majority of our sample had lower levels of education for whom formal Arabic is especially rare in daily interactions.

For example, some respondents did not understand “What is your year of birth?”, as it is more commonly asked in spoken Arabic in the form of “What year were you born?”. Interestingly, this question was not intended to be probed and was only asked to facilitate the flow of the interview, which shows that even the simplest of questions can have unforeseen problems.

Additionally, as with the questions themselves, the probes were asked in formal Arabic to be consistent with the language of the draft questionnaire. Thus respondents faced the same lack of understanding as they did with the questions, as illustrated by the respondent’s silence in the following example, after the interviewer had probed in formal Arabic about the time period the respondent was considering. In response to this evidence that the respondent did not understand the probe, the interviewer tried in multiple ways to reword it in spoken Arabic. Despite these efforts, the respondent continued to have difficulty answering the different versions of the probe:

> **I**: in formal Arabic: What time period were you thinking of when you were answering this question?
>
> **R**: **Silence**
>
> **I**: in spoken Arabic: How far back did you think?
>
> **R**: Now, you mean?
>
> **I**: in spoken Arabic: You thought about what period that passed?
>
> **R**: Oh no, I didn’t [need time to] think about this at all. There’s no time period.

Thus it seems many respondents had difficulty understanding specific terms and questions in formal Arabic, rather than the concepts behind the questions. In fact, out of the 101 questions, 70 of them presented problems of some kind and more than half of these, 38 questions, presented comprehension problems attributable to diglossia.

3.2. Ageing-related challenges

3.2.1. Sensory problems

Hearing problems created a significant challenge for our respondents. Two persons who had wanted to participate were unable to due to hearing impairments. For those who were able to complete the cognitive interview, hearing impairment made it difficult for some to communicate effectively with the interviewers:

I: What is your date of birth?

R: *Silence*

I: What is your date of birth? Should I use a louder voice?

R: Yes.

I: What is your date of birth?

R: Left his what?

I [louder]: Are you able to hear me well now?

R: Ahh left who?

I [even louder]: Are you able to hear me better now?

R: Yes.

I: What is your date of birth?

R: My date of birth? What?

I: What is your date of birth?

R: Date of birth? The year I was born?

I: Yes.

R: 1944.

3.2.2 Fatigue

Interviewers observed signs of fatigue in many respondents that included reduced concentration, engagement, and increased irritability when responding to probes towards the end of the interview. In the first example (Question #8), a respondent answered the early question energetically, providing unprompted elaboration, but much later (Question #37) provided considerably shorter, off-topic responses, and showed weariness and disengagement (in the analysts' judgment) when probed about "remittances." While lack of familiarity with the term "remittance" could have led to the respondent's terse answer to Q#37, the lack of clarification requests and disengaged response pattern (i.e., silence) suggest that fatigue rather than comprehension is the issue.

[Question #8]

I: How is your vision after wearing your eyeglasses?

R: I recently had surgery and feel a bit blurry. Before the surgery, my vision was poor, and now, thank God, it has improved, but I still experience some blurriness.

[Question #37]

I: How much does this household receive in total from remittances?

R: No, No.

[off-topic exchange]

I: What does the word “remittances” mean to you?

R: **Silence**

I: You don’t know what “remittances” is?

R: It doesn’t mean anything to me.

Another respondent showed signs of irritability with a probe at the end of the interview, which we attribute to fatigue.

I: When I asked you about the medical costs what came to your mind?

R: What “what came to my mind?”! Healthcare. Can you neglect your health? If you were sick, or had back pain? Can you avoid going to the doctor?

3.2.3 Cognitive Ability

Very long questions were especially problematic for older adults, given their reduced attentional ability (Harada et al., 2013). In the following example, a long question (even longer in Arabic) is asked to measure the respondent’s social network size, listing numerous family members. Much as the respondent in the example, eight respondents exhibited difficulties answering this question, asking for clarification, as well as who to include in the answer:

I: Now I’d like to ask you about your other relatives (besides your spouse and children), people that you are related to by blood or marriage. Please include grandchildren, brothers, sisters, sons-in-law and daughters-in-law, parents, aunts and uncles, and cousins. Give me your best guess of how many of these relatives you have. Would you say zero, between 1 and 5, between 6 and 10, between 11 and 20, or more than 20?”

R: What do you mean, I didn’t understand the question. *You* explain it to me.

In this example, the interviewer then repeated the question, only naming grandchildren, siblings, and aunts as examples, after which the respondent was able to answer the question easily.

The probes also demonstrated that some respondents retained only content from the beginning or the end of the list. In this example, respondents might focus primarily on grandchildren (as first item in the listing) or cousins (as last item in the listing). This pattern of responding is consistent with well-known response order effects, in particular primacy and recency effects (e.g., Krosnick, 1991):

R: [Daughter's name] has 4 kids, my other daughter has 4 children, that's 8. And my son has 2, that's 10. And my daughter has 2, that's 12. So 12 grandchildren.

I: Other than grandchildren, you have your siblings, your in-laws–

R [interrupting]: There is, there is. [...] 20-23 people.

Moreover, the interviewer had to repeat the questions multiple times in order to clarify the questions for some respondents:

I: Now, thinking again about the past 12 months, did you have to reduce, interrupt, or completely stop a regular treatment or medication you need?

R: **Silence**

I: For example, if you need medication...

R [interrupting]: Hypertension medication.

I: Did you interrupt or completely stop this medication because of any reason?

R: **Silence**

I: Did you have to reduce, interrupt, or completely stop a regular treatment or medication you need?

R: So, I did not get it you mean?

I: Yes.

R: No.

4. Discussion and conclusions

The cognitive interviews conducted for the main study revealed numerous respondent misunderstandings and difficulties carrying out response tasks, leading to revised questions that presumably elicited higher quality responses in the production interviews. But the experience also made evident that conducting cognitive interviews with older adults in an LMIC involves challenges that can potentially compromise the value of the pretest. We now appreciate how respondents' lack of survey literacy, their suspicion of institutions, and their difficulty understanding questions and probes in a formal version of Arabic can reduce the quality and usefulness of cognitive interview data from populations like this one. And, while hearing loss, fatigue, and cognitive decline with which these older respondents struggled were not surprising in retrospect, it underscores the need to design the pretest for the respondents and their capabilities.

Relatively few cognitive interview studies in LMICs have been published, and of those, most have been

conducted primarily in African countries, while no studies have focused on the Middle East or Arab populations. Previous studies often have had small sample sizes (less than 30), with primarily female participants, and have focused on younger adults rather than older populations. Additionally, they primarily examine one module such as social support (Martin et al., 2017) and nutrition assessments (Andrew et al., 2022; Namaste et al., 2024). The current study builds on this literature by including a substantially larger and more diverse sample (roughly equal numbers of males and females, educational attainment distributed roughly as in the population). Unlike previous studies, we test multiple cognitively demanding and sensitive questions, including modules assessing cognitive functions and exposure to political conflict and violence. Furthermore, our study is set in a complex sociopolitical context, introducing unique challenges related to what topics are sensitive to respondents. While existing literature highlights some difficulties in conducting cognitive interviews with older adults, our findings reveal additional context-specific challenges that further complicate the use of cognitive interviewing in this population.

The cognitive interviews in the current study reveal unique challenges for LMIC, especially when older adults are the target population. Based on these experiences, we have developed a set of best practices for researchers to follow when developing questionnaires for use in contexts and with respondents like those in the current study.

First, cognitive interviewing methods were initially developed in English-speaking countries where linguistic diversity is relatively low compared to multilingual societies such as Arab countries. Because there are usually multiple “low” versions of diglossic languages, it may be necessary to translate a questionnaire as well as the probes for cognitive interviews into multiple low versions of the target language. This undoubtedly involves more effort than translating the questionnaire into one formal, written version of the language but use of formal language was associated with numerous comprehension problems in the current study, mostly involving words that are not used in everyday conversation. These challenges are not unique to the Arabic, as other diglossic languages, such as Haitian Creole and Tamil, suggest that the problems arising from translation exposed by cognitive interviews may well extend to questionnaires written in these and other languages used in LMICs.

Second, because regional dialects can differ substantially, we recommend assigning cognitive interviewers who speak the same local dialect as the respondents. For example, researchers might recruit interviewers from the same area – even neighbourhood – as the respondents in the target population. Matching interviewers and respondents on the basis of the dialect they speak may increase respondents’ willingness to disclose that they do not understand a question or the cognitive interviewing enterprise more generally, much as matching production interviewers and respondents on the basis of race has been shown to elicit more honest answers (e.g., Schuman & Converse, 1971).

Third, educational attainment may accentuate the two groups of challenges we have identified, with those lower in education likely to be lower in survey literacy and ability to communicate in formal Arabic. Thus recruiting cognitive interview participants whose level of educational attainment roughly matches that of the target population is likely to make the pretest results more realistic. In the current study, it was difficult to find participants with high educational attainment, particularly women, a trend that may well be observed in other LMICs (Graetz et al. 2020). While it is often considered a best practice in cognitive interviewing to include some demographic variation (e.g., Beatty & Willis, 2007), in the type of population on which the current study is focused, recruiting participants with this in mind is particularly important.

Fourth, it is important to adequately train cognitive interviewers so they are prepared for the kind of

challenges we observed in this study. For example, it is essential to train them to clearly explain what kind of tasks respondents are being asked to carry out, and to call to their attention that in settings like these, respondents may not immediately understand the purpose of cognitive interviewing. Furthermore, clearly outlining how the information will be stored and used is necessary for reducing mistrust. This transparency, along with making it clear to respondents that they don't have to answer sensitive questions, should help encourage them to share relevant details.

Fifth, it is important to anticipate and accommodate ageing-related changes such as hearing problems, fatigue, and decreased cognitive ability when collecting data from older adults. These difficulties can mostly be remedied by interviewers speaking loudly, frequently asking respondents if they need a break, being patient and prepared to repeat questions, and clearly enunciating. Dividing the pre-test questions into more than one interview can also help mitigate fatigue and loss of concentration.

In conclusion, a systematic approach is needed for conducting cognitive interviews in LMICs, especially for older adults. More documentation as well as sharing experiences and best practices is needed both in the literature and within the community of practitioners who conduct cognitive interviews in similar LMIC settings and populations.

Appendix

Cognitive Interview Script– 2022

References

1. Abou-Mrad, F., Tarabey, L., Zamrini, E., Pasquier, F., Chelune, G., Fadel, P., & Hayek, M. (2015). Sociolinguistic reflection on neuropsychological assessment: An insight into selected culturally adapted battery of Lebanese Arabic cognitive testing. *Neurological Sciences*, 36, 1813-1822.
2. Alijla, A. (2016). Between inequality and sectarianism: Who destroys generalised trust? The case of Lebanon. *International Social Science Journal*, *66*(219–220), 177–195. https://doi.org/10.1111/issj.12122
3. Andrew, L., Lama, T. P., Heidkamp, R. A., Manandhar, P., Subedi, S., Munos, M. K., ... & Katz, J. (2022). Cognitive testing of questions about antenatal care and nutrition interventions in southern Nepal. *Social Science & Medicine*, 311, 115318.
4. Arab Center Washington DC (2023). *Arab Opinion Index 2022: Executive Summary*, https://arabcenterdc.org/resource/arab-opinion-index-2022-executive-summary/.
5. Ashok, S., Kim, S. S., Heidkamp, R. A., Munos, M. K., Menon, P., & Avula, R. (2022). Using cognitive interviewing to bridge the intent-interpretation gap for nutrition coverage survey questions in India. *Maternal & child nutrition*, 18(1).
6. Avlund, K. (2010). Fatigue in older adults: an early indicator of the aging process?. *Aging clinical and experimental research*, *22*, 100-115.
7. Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public opinion quarterly*, *71*(2), 287-311.
8. Behr, D. (2023). Translating questionnaires. In *International Handbook of Behavioral Health Assessment* (pp. 1-15). Cham: Springer International Publishing.

9. Blair, J., & Conrad, F. G. (2011). Sample size for cognitive interview pretesting. *Public opinion quarterly*, *75*(4), 636-658.
10. Brenner, P. S. (2017). How religious identity shapes survey responses. *Faithful measures: New methods in the measurement of religion*, 21-47.
11. Burke, D. M., & Shafto, M. A. (2004). Aging and Language Production. *Current directions in psychological science*, *13*(1), 21–24. https://doi.org/10.1111/j.0963-7214.2004.01301006.x
12. Caruso, A. J., Mueller, P. B., & Shadden, B. B. (1995). Effects of aging on speech and voice. *Physical & Occupational Therapy in Geriatrics*, *13*(1-2), 63-79.
13. Conrad, F. G., & Blair, J. (2009). Sources of error in cognitive interviews. *Public Opinion Quarterly*, *73*(1), 32-55.
14. Craik, F. I. (1999). Memory, aging, and survey measurement. In Schwarz, N., Park, D., Knauper, B., & Sudman, S. (Eds.). *Cognition, aging and self-reports*. Psychology Press.*Cognition,* (pp. 95-112). Psychology Press.
15. Darwish, N. (2019) "The Middle Eastern Societies:Institutional Trust in Political Turmoil and Stasis," *Silicon Valley Notebook*: Vol. 17 , Article 9. Available at: https://scholarcommons.scu.edu/svn/vol17/iss1/9
16. Fasfous, A. F., Al-Joudi, H. F., Puente, A. E., & Pérez-García, M. (2017). Neuropsychological measures in the Arab world: A systematic review. *Neuropsychology review*, *27*, 158-173.
17. Ferguson, C. A. (2003). Diglossia. In *The bilingualism reader* (pp. 71-86). Routledge.
18. Franzen, S., van den Berg, E., Goudsmit, M., Jurgens, C. K., Van De Wiel, L., Kalkisim, Y., ... & Papma, J. M. (2020). A systematic review of neuropsychological tests for the assessment of dementia in non-western, low-educated or illiterate populations. *Journal of the International Neuropsychological Society*, 26(3), 331-351.
19. Graetz, N., Woyczynski, L., Wilson, K. F., Hall, J. B., Abate, K. H., Abd-Allah, F., Adebayo, O. M., Adekanmbi, V., Afshari, M., Ajumobi, O., Akinyemiju, T., Alahdab, F., Al-Aly, Z., Rabanal, J. E. A., Alijanzadeh, M., Alipour, V., Altirkawi, K., Amiresmaili, M., Anber, N. H., ... Local Burden of Disease Educational Attainment Collaborators. (2020). Mapping disparities in education across low- and middle-income countries. *Nature*, 577(7789), 235–238. https://doi.org/10.1038/s41586-019-1872-1
20. Hall, S., & Beatty, S. (2014). Assessing spiritual well-being in residents of nursing homes for older people using the FACIT-Sp-12: a cognitive interviewing study. *Quality of Life Research*, 23, 1701-1711.
21. Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*, 29(4), 737–752. https://doi.org/10.1016/j.cger.2013.07.002
22. Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, 26(4), 433–443. https://doi.org/10.1177/0956797614567339
23. Howrey, B. T., & Hand, C. L. (2019). *Measuring Social Participation in the Health and Retirement Study*. https://dx.doi.org/10.1093/geront/gny094
24. *HRS Around the World.* (n.d.). Retrieved October 14, 2024 from https://event.roseliassociates.com/hrs-around-the-world/
25. Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging, 16*(4), 600-614. https://doi.org/10.1037/0882-7974.16.4.600
26. Kim, B. J., & Oh, S. H. (2013). Age-related changes in cognition and speech perception. *Korean journal of audiology*, *17*(2), 54–58. https://doi.org/10.7874/kja.2013.17.2.54
27. Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology*, *5*(3), 213-236.
28. Lay, K., Hutchinson, C., Song, J. *et al.* Cognitive interviewing for assessing the content validity of older-person specific outcome measures for quality assessment and economic evaluation: a scoping review. *Qual Life Res* (2024).
29. Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing loss prevalence and risk factors

among older adults in the United States. *The journals of gerontology. Series A, Biological sciences and medical sciences*, *66*(5), 582–590. https://doi.org/10.1093/gerona/glr002
30. Mallinson, S. (2002). Listening to respondents: A qualitative assessment of the Short-Form 36 Health Status Questionnaire. *Social Science and Medicine, 54*, 11–21.
31. Martin, S. L., Birhanu, Z., Omotayo, M. O., Kebede, Y., Pelto, G. H., Stoltzfus, R. J., & Dickin, K. L. (2017). "I Can't Answer What You're Asking Me. Let Me Go, Please." Cognitive Interviewing to Assess Social Support Measures in Ethiopia and Kenya. *Field Methods*, 29(4), 317-332.
32. Martínez-Nicolás, I., Llorente, T. E., Ivanova, O., Martínez-Sánchez, F., & Meilán, J. J. G. (2022). Many Changes in Speech through Aging Are Actually a Consequence of Cognitive Changes. *International journal of environmental research and public health*, *19*(4), 2137. https://doi.org/10.3390/ijerph19042137
33. Massey, M. (2025). Survey Practice in Non-Survey-Literate Populations: Lessons Learned from a Cognitive Interview Study in Brazil. *Survey Practice, 19 Special Issue.* https://doi.org/10.29115/SP-2024-0035
34. Melgar, M. I. E., Ramos Jr, M. A., Cuadro, C. K. A., Yusay, C. T. C., King, E. K. R., & Maramba, D. H. A. (2022). Translating COGNISTAT and the Use of the Cognitive Interview Approach: Observations and Challenges. *Madridge J Behav Soc Sci*, 5(1), 89-95.
35. Nadeau S.E. Aging-Related Alterations in Language (2019). In Heilman K.M., Nadeau S.E. (eds), 106-126. *Cognitive Changes and the Aging Brain*. Cambridge University Press.
36. Namaste, S. M., Bulungu, A. L., & Herforth, A. W. (2024). Cognitive Testing of Dietary Assessment and Receipt of Nutrition Services for Use in Population-Based Surveys: Results from a Demographic and Health Surveys Pilot in Uganda. *Current Developments in Nutrition*, 10451.
37. Park, H., & Goerman, P. L. (2019). Setting up the cognitive interview task for non-English-speaking participants in the United States. *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts (3MC)*, 227.
38. Park, H., Sha, M.M. & Olmsted, M. (2016). Research participant selection in non-English language questionnaire pretesting: findings from Chinese and Korean cognitive interviews. *Qual Quant 50*, 1385–1398.
39. Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging*, 34(1), 17–24. https://doi.org/10.1037/pag0000288
40. Sarkar, M., & de Jong, J. "Conducting Cognitive Interviews to Test Survey Items and Translation in the International Development Context." Presented at the World Association for Public Opinion Research Annual Meetings, September 19-22, 2023 Salzburg, Austria. Paper also presented at the Annual Cross-cultural Survey Design and Implementation Workshop, March 18 – 20, 2024.
41. Schaeffer, N. C. (2000). Asking questions about threatening topics: A selective overview. *The science of self-report: Implications for research and practice*, 105-121.
42. Schuman, H., & Converse, J. M. (1971). The effects of black and white interviewers on black responses in 1968. *Public Opinion Quarterly, 35*(1), 44-68.
43. Schwarz, N., Park, D., Knauper, B., & Sudman, S. (Eds.). (1998). *Cognition, aging and self-reports*. Psychology Press.
44. Sudharsanan, N., & Bloom, D. E. (2018). The demography of aging in low-and middle-income countries: chronological versus functional perspectives. In *Future directions for the demography of aging: Proceedings of a workshop* (pp. 309-338). Washington, DC: National Academies Press.
45. Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response.* Cambridge University Press.
46. Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public opinion quarterly*, *60*(2), 275-304.
47. Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, *133*(5), 859.
48. United Nations Department of Economic and Social Affairs (UNDESA). 2020. "*World Population Ageing 2020 Highlights*".

49. Varma, R., Vajaranant, T. S., Burkemper, B., Wu, S., Torres, M., Hsu, C., Choudhury, F., & McKean-Cowdin, R. (2016). Visual Impairment and Blindness in Adults in the United States: Demographic and Geographic Variations From 2015 to 2050. *JAMA ophthalmology*, *134*(7), 802–809. https://doi.org/10.1001/jamaophthalmol.2016.1284
50. Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. sage publications.
51. Willis, G. B. (2015). The practice of cross-cultural cognitive interviewing. *Public opinion quarterly*, *79*(S1), 359-395.
52. Wingfield, A. (1999) Speech Perception and the Comprehension of Spoken Language in Adult Aging. In Park, D. and Schwarz, N. (Eds.), *Cognitive aging: A primer*. Psychology Press. Psychology Press (pp. 175-195).
53. Yorkston, K. M., Bourgeois, M. S., & Baylor, C. R. (2010). Communication and aging. *Physical medicine and rehabilitation clinics of North America*, *21*(2), 309–319. https://doi.org/10.1016/j.pmr.2009.12.011
54. Zeinoun, P., Iliescu, D., & El Hakim, R. (2022). Psychological tests in Arabic: A review of methodological practices and recommendations for future use. *Neuropsychology Review*, 32(1), 1-19.