#AWORLD4ALLAGES

GLOBAL CAMPAIGN TO COMBAT AGEISM

WHO Ageism Scale

Manual and User Guide

March 2025

WHO Ageism Scale

Manual and User Guide

March 2025

WHO Ageism Scale: Manual and User Guide

Suggested citation. *WHO Ageism Scale: Manual and User Guide*. Geneva, Switzerland: World Health Organization; 2025. License: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at http://apps.who.int/iris.

Sales, rights and licensing. To purchase WHO publications, see https://www.who.int/publications/book-orders. To submit requests for commercial use and queries on rights and licensing, see https://www.who.int/copyright.

Update revisions. This manual will be updated and revised as needed, based on validation and emerging data on WHO Ageism Scale. For the latest version, please refer to https://www.aworld4allages.org.

Contact. For any enquiries regarding the scale, please contact hello@aworld4allages.org.

Contents

Acknowledgements

The WHO Ageism Scale was developed by the World Health Organization (WHO) and the University of Edinburgh, with leadership from Vânia de la Fuente-Núñez and Christopher Mikton; and Aja Louise Murray and Xuefei Li (University of Edinburgh, United Kingdom of Great Britain and Northern Ireland), under the supervision of Alana Officer, Unit Head, Demographic Change and Healthy Ageing, and Etienne Krug, Director, Social Determinants of Health.

Aja Louise Murray and Xuefei Li (University of Edinburgh, United Kingdom) led the drafting of the Manual.

We wish to thank the international panel of experts who generously shared their knowledge, guided the framework for the scale, assisted with developing and refining items, and provided invaluable feedback at each stage: Negin Masoudi Alavi (Kashan University of Medical Sciences, Iran), Jordi Alonso (IMIM-Hospital del Mar Medical Research Institute, Spain), Liat Ayalon (Bar Ilan University, Israel), Monika Bullinger-Naber (Universität Hamburg, Germany), Harsh Chauhan (Alt Unfold LLP, India), Carlo Chiorri (University of Genova, Italy), David Cicero (University of North Texas, United States of America), Brita Dorer (GESIS- Leibniz Institute for the Social Sciences, Germany), Natasha Duell (University of North Carolina at Chapel Hill, USA), Marlene Krasovitsky (EveryAGE Counts, Australia), Franziska Meinck (University of Edinburgh, United Kingdom), Julie Ober Allen (University of Oklahoma, USA), Elisha Sibale Mwamkinga (The Good Samaritan Social Service, Tanzania), Stephen Ojiambo Wandera (Makerere University, Uganda), Wiebke Weber (Pompeu Fabra University, Spain) and Xiangdong Yang (East China Normal University, China).

We are grateful for the generous financial support from the Ministry of Health, Labour and Welfare of the Government of Japan.

Abbreviations

CAPI	Computer-Assisted Personal Interviewing
CFA	Confirmatory factor analysis
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
FMIL	Full information maximum likelihood
IRT	Item response theory
OSF	Open science framework
PAPI	Paper-and-Pencil Interviewing
SD	Standard deviation
SDGs	Sustainable Development Goals
TRAPD	Translation, Review, Adjudication, Pretest, and Documentation
UN	United Nations
WHO	World Health Organization

1. Introduction

Summary

Ageism has a range of negative effects on individuals and societies. Prior to the development of the WHO Ageism Scale, no suitable measure existed to assess levels of ageism across the world in accordance with its currently accepted conceptualisation encompassing stereotypes, prejudices, and discrimination.

Ageism is defined by the World Health Organization (WHO) as stereotypes (how we think), prejudice (how we feel), and discrimination (how we act) based on age. It can manifest in various forms, including institutional, interpersonal, and self-directed ageism and can affect all age groups, though most of what we know about ageism pertains to ageism against older persons *(1)*.

Ageism is highly prevalent, with evidence suggesting that approximately 1 in 2 people are ageist towards older people *(1)*. It also has serious and far-reaching consequences for people's health and well-being *(1, 2)* and economic analyses show that ageism can be hugely costly to societies *(3)*.

Despite the scale and detrimental consequences of ageism, a comprehensive analysis of available ageism measures as part of the United Nations (UN) *Global report on ageism* found that there were no existing measures capable of measuring ageism as it is currently understood *(4)*. The UN *Global report on ageism (1)*, therefore, called for a new comprehensive, psychometrically robust scale that could be used across age groups and cultural contexts. Such a scale was needed to help illuminate the prevalence of ageism, assessing its risk/protective factors and to evaluate the effectiveness of interventions.

In response to this growing need, the WHO Ageism Scale was developed to measure the ageism *experiences* of individuals (in the *WHO Ageism Experiences Scale*) as well as people's ageism towards older persons (in the *WHO Ageism Towards Older Persons Scale*). By encompassing the cognitive (stereotypes), affective (prejudice), and behavioural (discrimination) components of ageism, the scale serves as a robust tool for researchers, policymakers, clinicians, educators, and community members seeking to assess age-based biases within and across diverse contexts.

This manual is aimed at researchers, policymakers and government officials, civil society organisations, health and care professionals, and anyone else who has interest in addressing ageism. It provides a practical

guide for introducing, administering, scoring, and interpreting the scale. By promoting the adoption of the WHO Ageism Scale, this manual aims to support more valid and reliable measurement of ageism, ultimately contributing to more effective strategies to combat ageism in diverse cultural and generational contexts.

2. The WHO Ageism Scale: origins and development process

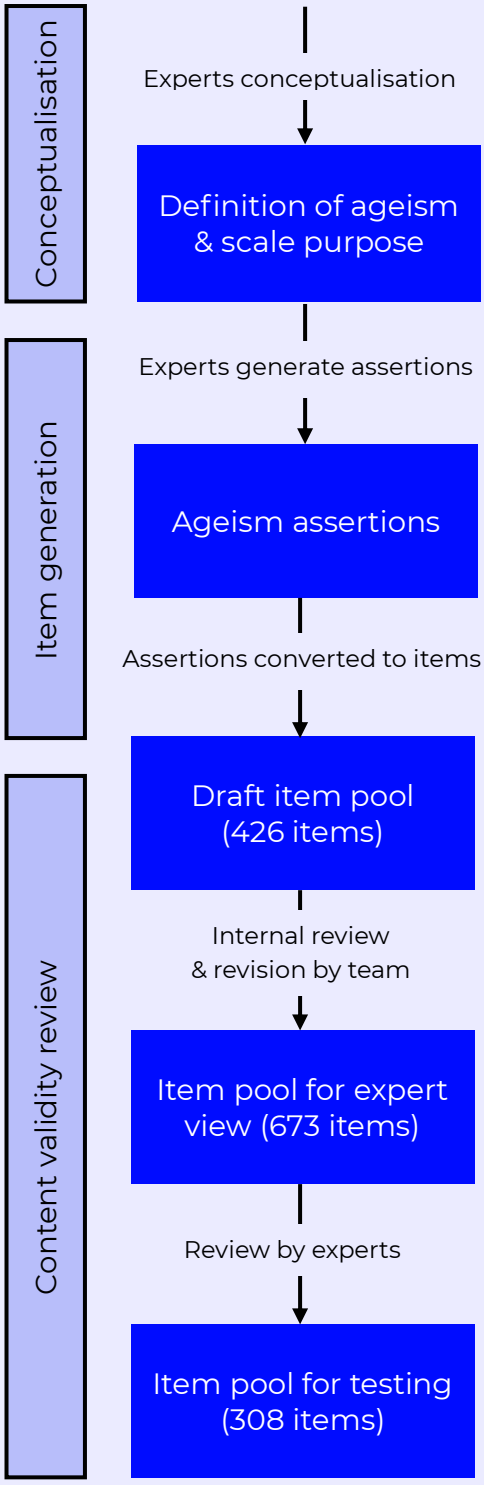
Summary

The WHO Ageism Scale was designed to measure ageism in the general population across the world. It uses a definition of ageism that recognises stereotypes (how people think about others due to their age), prejudice (how people feel about others due to their age) and discrimination (how people act towards others due to their age) as dimensions of ageism. It recognises that ageism can be self-directed, interpersonal, and institutional, affecting individuals across all age groups. It aims to capture both experiences of ageism and ageism that respondents direct towards others. The items of the WHO Ageism Scale were carefully developed in collaboration with experts from various disciplines and regions to ensure they accurately captured ageism as defined.

2.1 Creating the WHO Ageism Scale

The development of the WHO Ageism Scale was a rigorous multi-stage process involving the engagements of a large number of experts from across the world. A visual representation of the scale development process is provided in Figure 1.

Figure 1. Scale development process



At an expert meeting, the definition of ageism and the purpose of the scale was agreed. This was based on the expert input, a comprehensive review of ageism evidence, and collated items from existing ageism scales.

At the same meeting, experts generated 'assertions' based on the agreed conceptualisation of ageism. Assertions were short statements about ageism that aimed to represent the various dimensions of ageism. They were designed to form the basis of ageism items.

A psychometrics expert converted the assertions to items, in consultation with an ageism subject matter expert. This step followed best practice guidance in item construction.

Revisions were made by the internal team. Example revisions:

- Remove redundant items
- Edit items for clarity
- Remove items that did not fit the agreed conceptualisation of ageism
- Add items to address gaps
- Remove or edit items with clear content validity problems or that did not conform to best item construction practices
- Exclude experienced institutional stereotypes and prejudices items due to logical inconsistencies

Ageism and scale development experts reviewed and rated items using a structured form that assessed the scale as a whole and individual items. The internal team revised the scale and items in accordance with the feedback. Example revisions:

- Reverse word some items
- Further edits to items for clarity
- Remove or edit items that had clear content validity problems (including lack of cross-cultural relevance), that were not clear, or that were redundant
- Add a small number of new items
- Remove some of the most extremely worded items
- Simplify the language of some items
- Integrate item stems into items themselves rather than place them in pre-amble

Figure note. Adapted from *(5)*.

2.2 Design criteria

The initial phase of the development process of WHO Ageism Scale involved establishing design criteria to ensure a valid measure of ageism *(5)*. A critical aspect of this was agreeing on *conceptualisation*, i.e., the definition of ageism that would underpin the scale. This drew on the literature review evidence gathered as part of the UN *Global report on ageism*, the input of ageism experts, and a review of the content of existing ageism measures.

Six key criteria were established based on input from eight experts (three male, five female) in ageism and/or scale development, with representation from multiple WHO regions (the African Region, the European Region, the South-East Asia Region, and the Western Pacific Region) and age groups. Specifically, experts agreed that the new scale should:

- cover all three dimensions (stereotypes, prejudice, and discrimination) of ageism;
- consider three levels of manifestation (institutional, interpersonal, and self-directed);
- differentiate between experiences (i.e., as a target) and perpetration (i.e., as an actor);
- ensure all statement must be attributable solely to individual's *age;*
- suitable for all age groups; and
- maintain cross-cultural applicability.

2.3 Item generation

Following consensus on the design criteria, experts generated assertions: short statements about ageism that conformed to the agreed conceptualisation of ageism. These assertions were then converted to items and supplemented with other newly written items drawing on the ageism expertise of the team.

The item generation process aimed to minimise bias and ensure clarity and accessibility, especially for lower-literacy groups. English was used as the source language, with efforts made to avoid culturally specific expressions that would hinder translation. The draft items covered the following domains from an *experiences* perspective: self-directed stereotypes, self-directed prejudices, self-directed discrimination, interpersonal stereotypes, interpersonal prejudices, interpersonal discrimination, institutional discrimination. From the perspective of *ageism towards older persons* items covered: stereotypes, prejudices and discrimination against adolescents, young adults, middle-aged adults, and older adults.

2.4 Content validity review

Initial content validity review was conducted by experts, using a structured assessment form. A set of 673 items were evaluated for *accuracy* (alignment with the conceptualisation of ageism selected to underpin the scale), *clarity* (ease of understating/translation), *acceptability* (respondent appropriateness), and *lack of bias* (avoiding social desirability/reactivity). Items were rated as *low*, *moderate*, or *high* on these criteria, with open-ended feedback for improvements. Each item was reviewed by at least four experts (comprising a mix of scale development and ageism specialists), strategically assigned based on their expertise to ensure technical and contextual validity. Scale- and item-level feedback was obtained, and adjustments were made *(5)*. The output of this stage was the WHO Ageism item pool.

2.5 Item selection

Final items for the WHO Ageism Experiences Scale were selected from the WHO Ageism item pool and are in the process of being selected for the WHO Ageism Towards Older Persons Scale (the former was prioritised based on feedback from the Technical Advisory Group). To achieve a balanced measure of ageism whilst minimising participant burden and maintaining feasibility for large-scale studies, 15-item scales were considered an optimal length.

Item selection for the WHO Ageism Experiences Scale

In selecting the items content validity was prioritised over optimising statistical psychometric properties, such as reliability, to ensure all key domains of ageism were adequately represented. According to *(6)*, a selection strategy driven primarily by maximising reliability indices may lead to highly correlated items with similar content, thereby diminishing the conceptual breadth of the scale.

In addition, in the selection process considered two key features of content validity: *relevance* and *representativeness (7)*. Relevance refers to the appropriateness of an aspect of a scale to the target construct, while representativeness refers to the extent to which the facets of the target construct are proportionally represented by items. To achieve content validity, a scale should comprehensively cover all manifestations of the construct, accounting for both diversity and 'severity' in alignment with their occurrence within the population.

Based on these considerations, items were selected to ensure that core domains were well-represented in the scale, i.e., *interpersonal and self-directed stereotypes, prejudices, and discrimination* and *institutional discrimination* from an 'experiences' perspective *(1, 5)*. For the subscales

reflecting these focal domains, priority was given to more general items (e.g., items that were not specific to a particular context e.g., work, education, romantic relationships, finances) as more general items are likely to more effectively capture overall levels of exposure to ageism across different contexts. This is particularly important given one of the main goals of our scale that is to ensure its use across various age groups and country contexts. Further, to improve the acceptability of the scale, both positively and negatively worded items were included. This decision was guided by feedback from earlier content validity reviews, which highlighted the risk of an excessively negative scale. Items were also selected with the aim of capturing a range of 'severities' of ageism from mild or 'everyday' ageism to extreme ageism. This approach ensures a scale that can measure a wide range of ageism exposure levels with high reliability *(8, 9)*. Finally, items relating to discrimination experiences were given proportionally greater representation than stereotypes and prejudice, based on the assumption that discriminatory behaviours are the most direct and tangible forms of ageism and, therefore, may have the greatest impact on those experiencing them.

Item selection for the WHO Ageism Towards Older Persons Scale

Item selection is underway for the WHO Ageism Towards Older Persons Scale, following a similar process to the above-described. This manual will be updated to reflect future developments in this scale. Given that work remains underway for this scale, the remainder of this version of the manual focuses on the WHO Ageism Experiences Scale.

3. Scale Structure and Components of the WHO Ageism Experiences Scale

Summary

The WHO Ageism Experiences Scale is a 15-item measure encompassing all dimensions and levels of ageism experiences, including three positively worded items. It employs a five-point Likert-type (ranging from 1 = *strongly agree* to 5 = *strongly disagree*) with a '*don't know or not applicable*' option, which should be treated as missing. The final score should be reported as the mean of the sample, with guidance provided on managing missing values. Where space is constrained, a 5-item version is also available.

3.1 Scale structure

The WHO Ageism Experiences Scale contains 15 items, shown in Table 1. The items cover self-directed stereotypes (2 items), self-directed prejudice (1 item), self-directed discrimination (2 items), interpersonal stereotypes (2 items), interpersonal prejudices (1 item), interpersonal discrimination (3 items) and institutional discrimination (3 items). Of these 3 (20%) are positively worded.

Table 1. Item content and keying of WHO Ageism Scale

Item number	Content domain	Keying	Item in English
1	Self-stereotypes	+	At my age, my life has plenty of purpose
2	Self-stereotypes	-	I am a burden because of my age
3	Self-prejudice	-	I am embarrassed of my age
4*	Self-discrimination	-	Due to my age, I limit my participation in discussions even when they are about things that affect me
5	Self-discrimination	-	There are things I would like to do if I did not consider them inappropriate for my age group

6*	Interpersonal stereotypes	-	Others think that I have nothing valuable to contribute to society because of my age
7	Interpersonal stereotypes	+	Others think that at my age I am able to make decisions for myself
8*	Interpersonal prejudices	-	Others feel frustrated with me due to my age
9	Interpersonal prejudices	-	Other people feel uncomfortable around me because of my age
10	Interpersonal discrimination	-	Due to my age, other people talk to me as if I need things simplified
11*	Interpersonal discrimination	-	Others make decisions for me because of my age
12	Interpersonal discrimination	-	Due to my age, others make me feel excluded
13*	Institutional discrimination	-	Policies made by the government (e.g., on housing, social security, healthcare) do not meet the needs of people my age
14	Institutional discrimination	+	People my age are portrayed positively in the media
15	Institutional discrimination	-	I have been turned down for an opportunity (e.g., a job or volunteering opportunity) that I was qualified for because of my age

Note. Items with * sign are included in the five-item version.

3.2 Formative and reflective indicators

The interpersonal and institutional items in the WHO Ageism Scale have formative (or 'causal') rather than reflective (or 'effect') relations with the concepts they measure *(5)*. This perspective is based on the idea that for interpersonal and institutional ageism indicators, an individual's overall experiences of ageism do not arise from an internal latent (psychological) trait. Instead, these experiences are better conceptualised as a composite of interactions across various contexts and with multiple individuals and institutions, suggesting that each item contributes unique information rather than reflecting a single underlying dimension. In contrast, self-directed ageism, is best to be understood of as reflecting a psychological trait of internalised ageism.

3.3 Introductory statement

Before the questionnaire items, respondents should be presented with the following introductory statement: "*The following statements are designed to measure your experiences with different age groups. Use the response options below to tell us how much you agree with each statement. When answering, think about whether the statement applies in relation to the past 12 months*." Neutral language, such as 'experiences with other age groups', was used to avoid explicitly using the term 'ageism'. This can minimise the risk of priming socially desirable responses.

3.4 Reference period

A clearly defined reference period in psychometric instruments can reduce ambiguity; however, it is challenging to choose an appropriate timeframe. Too long a period can increase recall bias and participant burden because they have to think back a long way, whereas too short a period risks that important but infrequent manifestations of a construct are missed. Given that some items referenced infrequent events, such as being turned down for a job or voluntary opportunity, a one-year reference frame was selected. It should be noted that one potential drawback of this approach is that respondents' retrospective reports may be subject to forgetting and mood-contingent reconstruction (i.e., their responses are influenced by their current mood).

3.5 Response format and scoring

The scale employed a five-point Likert-type format, with response options: *Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly disagree*. A '*Don't know or not applicable*' option was provided to prevent participants from selecting the middle or a random response option when unable to complete the item, with such responses treated as missing values *(10)*.

For scoring responses, the following guidelines apply based on item wording direction (See Table 1 for the keying summary). For three positively worded items response options should be scored as: 1 = Strongly agree, 2 = Agree, 3 = Neither agree nor disagree, 4 = Disagree, 5 = Strongly disagree, Missing = Don't know or not applicable. For other negatively worded items, response options should be scored as: 5 = Strongly agree, 4 = Agree, 3 = Neither agree nor disagree, 2 = Disagree, 1 = Strongly disagree, Missing = Don't know or not applicable.

Next, to obtain the total or subscale scores, two approaches are recommended:

Average Composite Score: This involves summing the non-missing item scores and dividing by the number of non-missing items. This method assumes equal weighting across items and is appropriate when simplicity is prioritised.

Latent Variable Measurement Models: Confirmatory factor analysis (CFA) can generate factor scores for the self-directed ageism experiences items, which estimate the underlying construct scores by weighting items according to their factor loadings and accounting for measurement error.

Where there are missing item responses, these can be dealt with in several ways. For users confident in modern missing data methods such as multiple imputation methods or full information maximum likelihood (FIML) estimation, these methods can be used to impute missing values (multiple imputation) or estimate the relevant statistical models 'as if' the data were complete (FIML). Otherwise, we recommend taking the average of the non-missing responses as a person's overall score; however, please be aware that not all items measure an equally 'severe' level of ageism so this method can under- or over-estimate ageism levels. For 15-item version, if an individual has more than 50% of items missing, we recommend treating such scores with particular caution. Similarly, for the 5-item version, scores should be treated with caution if more than two items are missing.

3.6 Presenting and interpreting the results

The results of WHO Ageism Experiences Scale can be presented as the mean score of the sample. When presenting the descriptive statistics, it is recommended to include the number of non-missing responses, mean, standard deviation (SD), minimum and maximum values skewness and kurtosis.

There is no designated cut-off point for this scale. Following the appropriate scoring of negative and positive items, a higher mean score reflects greater exposure to ageism.

It is essential to recognise that this tool is not a diagnostic instrument, and its scores do not carry any clinical interpretation. Researchers are encouraged to contextualise mean scores based on different levels of manifestation and in relation to the sample's characteristics, study population, or comparable datasets.

4. Administration

Summary

Administration of the WHO Ageism Scale is ideally preceded by pilot testing and involves administration mode selection, and ethical compliance. Pilot testing with a representative sample of 5–10 individuals helps ensure clarity, appropriate length of the overall survey or interview, and identifies potential discomfort or technical issues early. Multiple administration modes, including CAPI (Computer-Assisted Personal Interviewing), PAPI (Paper-and-Pencil Interviewing), and online formats, have been used successfully. When interviewers are involved, proper training helps to minimise bias and inaccuracies. Training materials are available upon request for further guidance.

Pilot testing of the WHO Ageism Scale (especially the WHO Ageism Experiences Scale) prior to administration provides quality assurance and is strongly recommended. Selecting the appropriate administration mode selection, and ethical compliance are also important considerations. Below is a step-by-step guide to help with navigating these components and ensuring that the scale is implemented to a high standard.

4.1 Pilot testing

Pilot testing is recommended prior to administering the WHO Ageism Scale in the population of interest ('target population'), particularly when it is incorporated into a larger survey or research instrument (e.g., alongside health outcome measures or demographic variables). Piloting helps assess how well the scale integrates with other survey components, including logical consistency, survey flow, and respondent burden in a multi-measure context. Moreover, it can evaluate survey length and completion time, detect any mis-translated or mis-printed items and response options, and uncover technical or logistical issues.

Pilot testing is ideally done within a sample that closely matches the target population in terms of characteristics (e.g., age range, educational level) using a survey methodology that is close to the method that will be used in the target population (e.g., interviews, paper & pencil surveys). A sample size of 5–10 participants is typically sufficient for pilot testing.

4.2 Administration mode

The WHO Ageism Scale can be administered using various modes, including CAPI (Computer-Assisted Personal Interviewing), PAPI (Paper-

and-Pencil Interviewing), and online surveys. An English version of the scale, formatted for direct printing, can be found on the last page of this manual. There is no single *best* administration method; the choice should be guided by the specific context, available resources, and target population characteristics (e.g., literacy level). However, when using interviewers or fieldworkers, training should be provided to ensure consistency in administration and minimise interviewer bias.

Fieldworker training

A comprehensive fieldworker training programme (if data collection involves fieldworkers) should include, but is not limited to:

- project background and objectives;
- conceptual background, including ageism and other core study themes;
- survey structure and methodology, including an overview of questionnaire sections, data protection, and participant privacy considerations;
- key steps in data collection, with example scripts for various phases (e.g., initiating engagement, obtaining informed consent, survey closure);
- ethical principles in field research;
- safety and security measures for both researchers and participants;
- managing difficult situations, incorporating practical strategies and role-playing exercises;
- debriefing procedures and post-survey protocols; and
- Q&A opportunities.

A sample of the training materials, provided by a research team at the University of Edinburgh, can be accessed at https://blogs.ed.ac.uk/ageism-scale-validation/ but needs to be adapted to align with your research project.

4.3 Ethical considerations

Ethical standards in research dictate that informed consent must be obtained before administering the survey. Informed consent should be based on the provision of information that includes key details such as the nature of the data collection, any compensation provided, potential risks and benefits, confidentiality and data usage, data protection rights, voluntary participation, and the right to withdraw at any time.

5. Translation Guidance

Summary

WHO Ageism Scale was developed with the goal of being applicable in a wide range of languages. It was developed in English, with translatability considered during its development process. It has since been translated into multiple languages using gold-standard translation methods, including Arabic, Chinese, Italian, Portuguese, Romanian, Russian, Serbian, Spanish, Welsh (5-item version only). These translations are available on the WHO Ageism Scale webpage, with new versions added as they become available.

For new translations, the recommended approach is the TRAPD method (Translation, Review, Adjudication, Pretest, and Documentation). TRAPD ensures high-quality, culturally appropriate translations, with documentation maintained throughout the process. This guide provides instructions for adapting the scale to additional languages and contexts where a translation is not yet available.

The WHO Ageism Scale was developed in the source language of English, with the translatability of the items into other languages considered during scale development. A more detailed account of the item development process, including considerations of their translatability into other languages, is provided in the publication by *(5)*, which describes the process.

5.1 Translations of the WHO Ageism Scale

The WHO Ageism Experiences scale has been translated into a number of additional languages using gold standard translation methods. Available languages include Arabic, Chinese, Italian, Portuguese, Romanian, Russian, Serbian, Spanish, Welsh (5-item version only). The translations are provided on the WHO Ageism Scale webpage at https://www.aworld4allages.org/who-ageism-scale, with additional translations incorporated as they become available. This short guide provides guidance on translating the scale for use in new languages and contexts where a translation does not currently exist.

5.2 Recommended approach to translation – TRAPD method

Whilst translation methods continue to evolve, the current gold standard for producing translations of survey instruments is the 'TRAPD' method, which stands for **T**ranslation **R**eview **A**djudication **P**retest and

Documentation. This replaces older standards and differs from them primarily in involving two *forward* translations (from source to target language) and *no backward translations* (from the target language back to the source language). This is because it has been found that the backwards translation step was not effective in detecting translation problems. The TRAPD method has been adopted in large-scale cross-national survey efforts such as the *European Social Survey (11)*. The steps and their rationales are described in turn below. T, R, A, and P refer to distinct steps whereas D runs throughout the process and refers to the documentation of the process.

Steps in translation

Translation (T)

The first step is to obtain 2 or more forward translations conducted by translators with complementary expertise. The purpose of having two translations is to provide different translation options and stimulate discussions which can help ensure an overall more suitable translation. These translations should be conducted *independently.* Ideally, translator 1 is a professional translator (bringing linguistic expertise and if possible, survey translation expertise) and translator 2 is a subject matter expert familiar with survey methods (bringing expertise suitable for translating more technical terms and with an understanding of the construction of survey items). The aim for the two translators is not to provide a literal translation, but to provide a translation as conceptually similar to the source language as possible. The translators should be provided with background information about the concept of ageism and the purpose of the scale to ensure they have an understanding of the translation requirements (see 'Documentation'). The WHO Ageism Scale manual, available at the scale webpage, is a key source of this information.

Some adaptations to the translation stage may be justifiable in some contexts but should be carefully considered in terms of their potential impact on the translation and clearly documented. For example, one challenge often encountered relates to the fact that survey translation is a specialised form of translation, and it may be difficult to secure a professional translator with the right expertise. In these cases, a second social scientist may be a more appropriate translator, provided relevant translation expertise can be evidenced. Where resources are limited, it may also be difficult to secure two human translations. In these cases, a machine translation (i.e., artificial intelligence translation) with post-editing could be used to obtain a second translation. However, it is important to stress that machine translations, while becoming very accurate, will require careful review and post-editing by a human. You can read more about how machine translation is being used in survey translation in recent publications *(11)*.

Review and Adjudication (R&A)

In the review stage, the two original translators are joined by a broader review team who review and discuss the two translations. The translations should be reviewed and annotated ahead of the meeting and the meeting itself should be chaired by a third team member who was not involved in the translation stage but who has relevant (linguistic, subject matter, survey methods) expertise. Other team members with relevant expertise may also join the meeting to bring additional perspectives as relevant. The chair leads the discussions of the two translations with the aim of: i) understanding why translations might differ and identifying a best consensus translation; ii) identifying improvement; and iii) resolving challenges encountered (e.g., where there is no exact translation).

The review stage can be combined with the adjudication stage, which refers to making a decision on the 'best' translation considering the two independent translations and discussions around them. The chair of the review meeting would be well-placed to make the final decision. All areas of discussion, the final decisions, and their justifications should be noted. For example, if the team debated two alternative word choices for a translation the rationale for the final selection should be documented. Similarly, if there are any particularly challenging aspects to translate (e.g., no exact translation for a word is available in the target language), these should be documented along with the decision taken.

The original, annotated and pre-final translations (the final output of the review and adjudication stage) should also be retained as part of the documentation. If the meeting is recorded, a transcription of the discussions can also serve as helpful documentation.

Pre-testing (P)

In the pre-testing phase, the pre-final translation is tested with members of the target population (e.g., older persons living in the community). The purpose of this phase is to identify any issues with the comprehensibility and acceptability of the translation (e.g., ambiguous, offensive, or difficult-to-understand words or phrases that may cause issues). It is difficult to estimate how many target participants are needed in this phase as it is unclear when new difficulties may continue to be identified, therefore, an initial sample of 5-10 is recommended but if new issues continue to be uncovered, further participants may continue to be engaged until no new issues are being flagged. A common method of pre-testing is to use a cognitive interview *(12)*, in which the respondent is asked to complete the scale and is either asked to 'think-aloud' as they answer to reveal their comprehension and response process or is asked to comment on any

issues with the items after completing them. They might be asked the following types of questions using different types of 'probes':

- Comprehension/interpretation/meaning-oriented: 'What, to you, is X?';
- Paraphrase: 'Can you tell me in your own words what this question is asking?';
- Process-oriented: 'How did you arrive at that answer?';
- Confidence judgement: 'How sure are you that X?';
- Evaluative probe: 'Do you feel this question is easy or difficult to answer?';
- Elaborative probe: 'Why do you say that?' 'Tell me more';
- Hypothetical probe: 'How would you answer this if X'; and
- Sensitivity probe: 'Do you think that this question asks about things that are too private, or is it OK to ask this?'.

The selection of probes may depend on the questions that arise during the review and adjudication stages. For example, if it is unclear which word choice is more suitable for a given item, additional evidence can be gathered through pre-testing. Any issues arising at this stage and consequent changes to the translation should be documented.

The selection of the pre-testing sample of participants is important. If, for example, the participants are not well-representative of the target population then key issues may not be picked up with the translation. For example, if the scale is only pre-tested with very highly educated participants, difficulties with comprehending complicated or technical terms may not be picked up.

Documentation (D)

Documentation refers to ensuring that all relevant information about the translation input, process, and output are available. This is important for assuring users of the quality of translation and may also impact its interpretation in other ways. The documentation produced /retained should include the following information:

1. Any instructions given to translators
2. Information about the qualifications/experience of the translators and their selection
3. The material to be translated in the source language, the intermediate translations (including annotated translations from the review stage) and final translation
4. The changes made at each stage and the justification
5. Any further background information about the scale.

The main principle of the 'documentation' aspect of TRAPD is to ensure that anyone who was not involved in the translation should be able to

understand the process that was followed and why the final translation has the form that it does.

6. Validation Guidance

Summary

The validation of the WHO Ageism Scale is essential to ensure its reliability, robustness, and cross-cultural applicability, enabling insights into the prevalence, drivers, and consequences of ageism across the globe. Key psychometric evaluations can be guided by COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) criteria include factor analysis (e.g., CFA), reliability assessments (internal consistency, test-retest), validity tests (construct, convergent), and measurement invariance. Validation work should consider several recommendations, such as sample diversity, investigating intersectionality, and psychometric analyses.

Initial independent validation studies are available at https://blogs.ed.ac.uk/ageism-scale-validation/. For guidance or resources on validation studies, including psychometric testing, cross-cultural adaptation, or clinical validation, please contact the development team at the University of Edinburgh (aja.murray@ed.ac.uk and xuefei.li@ed.ac.uk) to ensure alignment, avoid duplication, and facilitate collaboration. Future studies can be shared with WHO at hello@aworld4allages.org.

6.1 Validation guidance

Validation of the WHO Ageism Scale is critical to ensure the robustness, reliability, and applicability of its scores across diverse populations and contexts. This process also helps researchers and policymakers capture the multifaceted experiences of ageism across different cultural, socioeconomic, and age-specific groups. By contributing to validating the scale, users can enhance comparability across populations, enabling meaningful cross-cultural and cross-context insights into ageism as a global issue. Users can contribute to validating the scale in two ways: i) by reporting the psychometric properties of the scale when it is used in new samples; and ii) by conducting dedicated validation studies aimed specifically at validating the scale.

Some psychometric tests that are recommended for the conducting on the scale scores, include the factor analysis (e.g., CFA), reliability (e.g., internal consistency, test-retest reliability), validity (e.g., construct validity, convergent validity, divergent/discriminant validity, concurrent/predictive

validity), and measurement invariance analyses. However, factor and internal consistency reliability analyses should only be conducted on the reflective items.

A one factor model is suggested for the self-directed ageism item when performing the CFA and assessing internal consistency.

6.2 Recommendations for validation work

Several validation studies of the WHO Ageism Scale have been conducted independently of WHO, led by researchers at the University of Edinburgh and other institutions. WHO has had no role in designing, conducting, or overseeing these studies.

For preliminary validation evidence:

Murray, A. L., Li, X., & Booth, T. (2025). Preliminary validation of the 15-item WHO experiences of ageism scales in a mixed-age United Kingdom sample. osf.io/preprints/psyarxiv/7jcqk_v2

Murray, A. L., Li, X., & Booth, T. (2025). Exploring the psychometric robustness of a 5-item version of the WHO ageism experiences scale in a United Kingdom adult sample. osf.io/preprints/psyarxiv/cbdh5_v1

Information on these independent validation efforts is being curated by the University of Edinburgh team at: https://blogs.ed.ac.uk/ageism-scale-validation/.

Building on insights from these validation studies, we recommend keeping the preamble and response options unchanged, as they have been carefully designed and tested. However, we understand that modifications may sometimes be needed. Future research can consider the following points to address gaps and further enhance the scale's application.

Expand population diversity

- Ageism affects all ages. While ageism is often discussed in the context of older persons, adolescents and younger adults also face age-related stereotypes, prejudice, and discrimination. Furthermore, studying ageism in adolescence—a formative period for identity development—can offer insights on how early exposure to ageism contribute to the internalisation of age-stereotypes over the lifespan.
- Validate the scale in high, medium and low-resource contexts. Analyses such as measurement invariance can help evaluate the equivalence across cultures.

- Ageism can intersect with other forms of “-isms”, amplifying marginalisation. It is important to test the scale among populations experiencing intersectional discrimination, such as older persons with disabilities.

Methodological enhancements

- Examine psychometric properties such as divergent validity, measurement invariance, as well as conduct item response theory (IRT)-based analyses which can complement the insights from factor analysis in reflective items and deepen our understanding of the interpretation of the scale scores.
- Adopt a mixed-methods approach by combining quantitative validation with qualitative interviews to explore nuances, such as how respondents interpret scale items.

If you plan to validate this scale in a particular context, we encourage you to review Section 8: Open Science and Pre-registration and contact hello@aworld4allages.org.

7. Recommended usage of the WHO Ageism Scale

Summary

The WHO Ageism Scale is adaptable for use from adolescence onwards, facilitating research on ageism across the lifespan. The WHO Ageism Experiences Scale and the WHO Ageism Towards Older Persons Scale are provided in printable English versions in Annex 1 and Annex 2, respectively. This scale is particularly valuable for intergenerational studies and can be applied in multidisciplinary academic research, healthcare contexts, community initiatives, and policy or programme evaluations.

7.1 Target population

The WHO Ageism Scale design allows for adaptation for use with adolescents through to older people, enabling researchers and practitioners to explore ageism across the lifespan. This flexibility supports its use in diverse contexts, including intergenerational studies that examine ageism experiences among younger, middle-aged, and older individuals. However, special considerations are necessary when administering it to participants under 18 or those with cognitive impairments, as they may have limited comprehension, require additional consent procedures, and need modifications to the administration methods.

It also can be used across different cultural contexts. When translated and culturally adapted using the TRAPD method, it becomes suitable for assessing ageism experiences in diverse global settings, ranging from high-income area to communities in low-resource regions. Furthermore, it is appropriate for both general population surveys and targeted studies of specific subgroups, including marginalised populations who may experience the intersectionality of ageism with other forms of "-isms" (e.g., ableism).

7.2 Potential settings for administration

Multidisciplinary academic research. Academic and research teams from different fields may find the WHO Ageism Scale useful for conducting systematic studies on ageism. This tool enables the collection of comparable data across various cultural contexts and demographic groups, contributing to the global understanding of ageism. Moreover, its standardised format and different length options make it ideal for large-

scale representative surveys and longitudinal studies to examine ageism change over time and its relations with hypothesises drivers and consequences. It can also be used as an outcome measure in interventions to see how a (policy, programmatic, or other type of intervention) impacts ageism.

Clinical and healthcare settings. In healthcare environments, this tool may illuminate how older persons perceive and experience ageism within care systems, offering critical insights to improve equity and patient-cantered practices. Clinics and hospitals might administer the tool to specific patient groups—such as older persons with depression, chronic illnesses, or cognitive impairments—to investigate how ageism experiences intersect with their health challenges. Geriatric care teams could use these findings to redesign care models that actively address gaps in dignity, communication, or resource allocation reported by patients. Rehabilitation programs might also apply the tool to track whether older persons perceive ageism as a barrier to accessing services or shaping providers' expectations of their recovery potential. Such applications align with broader efforts to reduce health disparities rooted in age-based discrimination.

Community-based settings. The WHO Ageism Scale could be a resource for capturing older persons' lived experiences of ageism in community contexts. For instance, community groups could use the tool to establish baseline before launching interventions such as intergenerational programmes, then re-test post-implementation to evaluate effectiveness. Public health departments could integrate the tool into population surveys to uncover how older persons perceive ageism as a barrier to healthcare access, social participation, or engagement in age-friendly city initiatives.

Policy evaluation. The multiple manifestation levels of the WHO Ageism Scale makes it possible for policymakers and governments auditing ageism at systemic levels. Moreover, national or regional agencies could deploy it to assess the prevalence of ageism before and after implement of anti-ageism policies, laws, or social security schemes and monitor progress over time. Global organisations like the United Nations (UN) or WHO might use cross-country data to benchmark ageism levels, aligning interventions with the UN Decade of Healthy Ageing (2021–2030), Madrid International Plan of Action on Ageing, the Sustainable Development Goals (SDGs) and other global goals.

Other intervention evaluation. Programmes targeting ageism across sectors can benefit from the tool's adaptability. For example, corporations might embed it into workplace diversity, equity, and inclusion (DEI) initiatives to assess employers' ageism experiences in hiring, promotion, or team dynamics, ensuring employees are not feel marginalised due to their age.

8. Open Science and Pre-registration

Summary

The WHO Ageism Scale supports open science to enhance research transparency, collaboration, and reproducibility. Researchers are advised to follow open science principles, including pre-registering study protocols, sharing anonymised data in repositories where permissible, and providing accessible materials (e.g., adapted scale versions) with proper attribution. For validation studies, such as psychometric or cross-cultural testing, contacting the development team before starting is recommended. This collaboration helps prevent duplication and aligns efforts with the scale's goals. These guidelines aim to ensure ethical use, robust evidence, and continual improvement of the WHO Ageism Scale.

The WHO Ageism Scale promotes open science practices to foster transparency, collaboration, and reproducibility in research. This section outlines guidelines for ethical use, licensing, and contributions to open science.

8.1 Permission and licensing

The WHO Ageism Scale is freely available under a CC BY-NC-SA 3.0 IGO license. This means that anyone can freely copy, reproduce, reprint, distribute, translate and adapt the work for non-commercial purposes, provided WHO is acknowledged as the source. In keeping with best practices for transparency and ethical scholarship, when using this scale:

Appropriate attribution. Always cite the original source of the WHO Ageism Scale in all publications, reports, or presentations where the scale is used. Proper attribution allows other researchers to trace back to the original materials for additional context.

Modifications and adaptations. If you adapt or modify the scale – such as translating it into another language, altering item wording, or changing the response format – clearly document all changes in supplementary materials or methodological sections of any related publications. By detailing your modifications, you help maintain transparency and enable other researchers to replicate or build upon your work accurately.

8.2 Open science practices (optional)

Researchers are welcome to adopt open science principles when using the scale. If you wish to contribute to transparency and broader scientific collaboration, consider the following steps:

Pre-registration. If applicable to your field, pre-register study protocols (e.g., hypotheses, methods, analysis plans) on platforms such as the Open Science Framework (OSF).

Data Sharing. Sharing de-identified datasets in public repositories (e.g., OSF) can allow other researchers to verify findings, conduct secondary analyses, and combine datasets to explore larger-scale trends. Always ensure you have the necessary permissions (e.g., from ethics committees and informed consent from participants) for sharing data.

Open Materials. Making your study materials publicly available (e.g., questionnaires, instructions, analysis coding scripts) enables others to replicate or adapt your work accurately.

8.3 Validation studies

If you plan to conduct validation studies (e.g., psychometric testing, cross-cultural adaptation, clinical validation), we encourage you to contact the development team at the University of Edinburgh: aja.murray@ed.ac.uk and xuefei.li@ed.ac.uk before you begin your work. This facilitates collaboration, avoids duplication of efforts, and ensures alignment with the scale's core objectives and global harmonisation. The team can also provide guidance or resources to support your work.

References

1. Global report on ageism. Geneva: World Health Organization; 2021 (https://www.who.int/publications/i/item/9789240016866).
2. Hu RX, Luo M, Zhang A, Li LW. Associations of ageism and health: a systematic review of quantitative observational studies. Research on Aging. 2021;43*(7-8)*:311-22 (https://doi.org/10.1177/0164027520980130).
3. Levy BR, Slade MD, Chang ES, Kannoth S, Wang S-Y. Ageism amplifies cost and prevalence of health conditions. The Gerontologist. 2020;60*(1)*:174-81 (https://doi.org/10.1093/geront/gny131).
4. Ayalon L, Dolberg P, Mikulionienė S, Perek-Białas J, Rapolienė G, Stypinska J et al. A systematic review of existing ageism scales. Ageing research reviews. 2019;54:100919 (https://doi.org/10.1016/j.arr.2019.100919).
5. Murray AL, Fuente-Núñez V. Development of the item pool for the 'WHO-ageism scale': conceptualisation, item generation and content validity assessment. Age and Ageing. 2023;52(Supplement_4):iv149-iv57 (https://doi.org/10.1093/ageing/afad105).
6. Meijer RR, Egberink IJL. Investigating Invariant Item Ordering in Personality and Clinical Scales: Some Empirical Findings and a Discussion. Educational and Psychological Measurement. 2012;72*(4)*:589-607 (https://doi.org/10.1177/0013164411429344).
7. Haynes SN, Richard D, Kubany ES. Content validity in psychological assessment: A functional approach to concepts and methods. Psychological assessment. 1995;7*(3)*:238 (https://doi.org/10.1037/1040-3590.7.3.238).
8. Murray AL, Eisner M, Ribeaud D. Can the Social Behavior Questionnaire help meet the need for dimensional, transdiagnostic measures of childhood and adolescent psychopathology? European Journal of Psychological Assessment. 2017 (https://doi.org/10.1027/1015-5759/a000442).
9. Van den Oord EJCG, Pickles A, Waldman ID. Normal variation and abnormality: an empirical study of the liability distributions underlying depression and delinquency. Journal of Child Psychology and Psychiatry. 2003;44*(2)*:180-92 (https://doi.org/10.1111/1469-7610.00112).
10. Murray AL, Booth T, Molenaar D. When middle really means "Top" or "Bottom": An analysis of the 16PF5 using Bock's nominal response model. Journal of personality assessment. 2016;98*(3)*:319-31 (https://doi.org/10.1080/00223891.2015.1095197).

11. Zavala-Rojas D, Behr D, Dorer B, Sorato D, Keck V. Using Machine Translation and Post-Editing in the TRAPD Approach: Effects on the Quality of Translated Survey Texts. Public Opinion Quarterly. 2024;88*(1)*:123-48 (https://doi.org/10.1093/poq/nfad060).

12. Willis GB. The practice of cross-cultural cognitive interviewing. Public opinion quarterly. 2015;79(S1):359-95 (https://doi.org/10.1093/poq/nfu092).

Annex 1. WHO Ageism Scale: experiences

The following statements are designed to measure your experiences with different age groups. Use the response options below to tell us how much you agree with each statement. When answering, think about whether the statement applies in relation to the past 12 months.	**Strongly agree**	**Agree**	**Neither agree nor disagree**	**Disagree**	**Strongly Disagree**	**Don't know or not applicable**
At my age, my life has plenty of purpose	○	○	○	○	○	○
I am a burden because of my age	○	○	○	○	○	○
I am embarrassed of my age	○	○	○	○	○	○
Due to my age, I limit my participation in discussions even when they are about things that affect me	○	○	○	○	○	○
There are things I would like to do if I did not consider them inappropriate for my age group	○	○	○	○	○	○
Others think that I have nothing valuable to contribute to society because of my age	○	○	○	○	○	○
Others think that at my age I am able to make decisions for myself	○	○	○	○	○	○
Others feel frustrated with me due to my age	○	○	○	○	○	○
Other people feel uncomfortable around me because of my age	○	○	○	○	○	○
Due to my age, other people talk to me as if I need things simplified	○	○	○	○	○	○
Others make decisions for me because of my age	○	○	○	○	○	○
Due to my age, others make me feel excluded	○	○	○	○	○	○
Policies made by the government (e.g., on housing, social security, healthcare) do not meet the needs of people my age	○	○	○	○	○	○
People my age are portrayed positively in the media	○	○	○	○	○	○
I have been turned down for an opportunity (e.g., a job or volunteering opportunity) that I was qualified for because of my age	○	○	○	○	○	○

Annex 2. WHO Ageism Scale: towards older persons

The following statements are designed to measure your thoughts, feelings, and behaviours toward older adults. Use the response options below to tell us how much you agree with each statement. When answering, think about whether the statement applies in relation to the past 12 months.	**Strongly agree**	**Agree**	**Neither agree nor disagree**	**Disagree**	**Strongly Disagree**	**Don't know or not applicable**
Older adults have a lot to contribute to society	O	O	O	O	O	O
Older adults should stick to being around people their own age	O	O	O	O	O	O
Older adults are too old for romance	O	O	O	O	O	O
Older adults are a burden	O	O	O	O	O	O
Older adults are too old to change	O	O	O	O	O	O
I feel frustrated with older adults	O	O	O	O	O	O
I feel bored listening to older adults	O	O	O	O	O	O
I feel pity for older adults	O	O	O	O	O	O
I enjoy being around older adults	O	O	O	O	O	O
I find older adults interesting	O	O	O	O	O	O
I talk to older adults in simplified language	O	O	O	O	O	O
I exclude older adults from certain conversations	O	O	O	O	O	O
I avoid spending time with older adults	O	O	O	O	O	O
I listen to older adults	O	O	O	O	O	O
I ask older adults for their view	O	O	O	O	O	O