

Exploring the use of self-report behavioural science questionnaires in sub-Saharan African countries

**Corina Weir** 

**University of Stirling** 

Candidate for Professional Doctorate in Health Psychology – Top Up

October 2023

#### **Thesis Abstract**

#### Background

The study of human behaviour is dominated by work published on Western, Educated, Industrialised, Rich, and Democratic (WEIRD) populations. Not only do samples in psychological studies need diversifying, but the research methods and tools used to assess human behaviour require further consideration for non-WEIRD populations. There is a demand to prioritise research efforts in sub-Saharan African (SSA) countries and ensure that research methods are culturally relevant.

#### Methods

Mixed methods were used to assess what researchers should consider when developing behavioral science measures to explore self-report measures in SSA countries. A qualitative study explored behavioural science researchers' experiences of the methodological challenges of using self-report questionnaires in SSA countries. A systematic review considered what selfreport measurement tools have been used to assess the behavioural influences on antimicrobial use for healthcare workers working in SSA countries, as an example area to explore potential problems with measurements. A secondary data analysis study was also conducted to assess whether response category descriptors of Likert scales were perceived differently by individuals in the UK and individuals in SSA countries.

#### Results

This thesis highlights key challenges which are particularly pertinent for researchers working in SSA countries, including differing interpretations of psychological constructs; complexities with adaptation and translation of question items and Likert response scales; and under reporting of pilot testing, validity, and reliability assessments.

#### Conclusions

Being aware of the potential limitations with quantitative self-report measures, reflecting on alternative methods for collecting data and ensuring adequate adaptation of measurement tools is crucial to developing questionnaires which are relevant in SSA countries. This may involve investing in research development practices to understand psychological constructs from the viewpoint of individuals in SSA countries, setting aside Western preconceptions on psychological constructs and how they should be measured.

# **Table of Contents**

Acknowledgements	5
List of tables	6
List of figures	7
List of appendices	8
Table of abbreviations	9
Chapter 1: Introduction	10
Chapter 1.1: WEIRD dominance in the literature	10
Chapter 1.2: Cultural differences in behavioural science	11
Chapter 1.3: Cultural differences influence self-reporting	12
Chapter 1.4 Research problem	13
Chapter 1.5 Aim and structure of thesis	14
Chapter 1.6 Changes to thesis	14
Chapter 1.7 Note to terminology	15
Chapter 2: Are self-report questionnaires in behavioural science too Western?	
Exploring the views of researchers working in sub-Saharan African	16
countries	
Chapter 2.1: Introduction	16
Chapter 2.2: Methods	20
Chapter 2.3: Results	27
Chapter 2.4: Discussion	38
Chapter 3: Measurement of behavioural influences of healthcare workers	
towards antimicrobial use in sub-Saharan African countries: A systematic	
review	42
Chapter 3.1: Introduction	42
Chapter 3.2: Methods	45
Chapter 3.3: Results	48
Chapter 3.4: Discussion	54

# **Table of Contents**

Chapter 4: Assessing response bias in agreement scale responses between			
individuals from the UK and sub-Saharan African countries			
Chapter 4.1: Introduction	93		
Chapter 4.2: Methods	97		
Chapter 4.3: Results	101		
Chapter 4.4: Discussion	110		

# Chapter 5: Discussion

Chapter 5.1: Overview	114
Chapter 5.2: Importance of findings for research practices	115
Chapter 5.3: Reflections	118
Chapter 5.4: Limitations	118
Chapter 5.5 Conclusions	119
References	132

# Acknowledgements

I would like to thank both my supervisors, Dr Lesley McGregor, and Prof Vivien Swanson, for their exceptional support and advice throughout this thesis. Without their encouragement, confidence and positivity, this thesis would not have been possible. Special thanks to Dr Lesley McGregor for her continual optimism and reassurance, especially in the final months.

Thank you to both of my advisors, Prof Lucie Byrne-Davis, and Dr David Musoke, for their guidance, support, and suggestions. Their input was extremely valuable and helped inform my thinking and approach throughout the entire research process. A special thanks to Dr Dan Powell, a good friend and colleague who shared his wisdom and offered a few laughs along the way. I would also like to thank Dr Neil Scott for his invaluable support with Chapter 4, Nicholus Nanyeenya for excellent advice and help with Chapter 2, and Muhammad Ahmed for his dedication with reviewing support in Chapter 3.

To the participants in all three studies included within this thesis, I thank you for your time and for sharing your experiences. The main goal of this thesis was to consider how we can improve research practices in relation to self-reporting, and without gathering the views from participants this would be impossible.

Lastly, I would like to extend my love and gratitude to my family, especially my incredible children, Mila, and Luca, who were both born during this research process. My mother taught me to be determined, believe in myself and to always persevere. I am extremely grateful for her support and encouragement.

This thesis would not have been possible without my wonderful husband, Agustin Gambaretto. Words can hardly describe my thanks and appreciation, and I dedicate this thesis to you.

# List of tables

Table 1 Participant information	23
Table 2 Results for quality assessment appraisal using AXIS	59
Table 3 Quality assessment using NHLBI	62
Table 4 Characteristics of included studies	63
Table 5 Publication information	77
Table 6 Questionnaire adaptation information	80
Table 7 Measurement information	
Table 8 Country frequencies	101
Table 9 Participant demographic information	
Table 10 Response category frequencies for equality scale items	103
Table 11 Response category frequencies for scepticism scale items	104
Table 12 Response category frequencies for males and females within each group	106
Table 13 Summary of individual ordinal logistic regressions predicting Likert response	category
for each question item	109

# List of figures

Figure 1. Map of the SSA countries included in the review	48
Figure 2. PRISMA flow diagram	49

# List of appendices

Appendix 1 Participant information sheet	121
Appendix 2 Consent form	124
Appendix 3 Interview topic guide	125
Appendix 4 Reflective notes and audit trail example	126
Appendix 5 Database search strategies	127
Appendix 6 Appraisal tool for Cross-Sectional studies (AXIS)	130
Appendix 7 NHBLI Quality Assessment Tool for Before-After (Pre-Post) Studies with No	
Control Group	131

# Table of abbreviations

Abbreviations	Meaning
HICs	High income countries
LMICs	Low- and middle-income countries
LS	Likert Scale
SSA	sub-Saharan Africa
WEIRD	Western, educated, industrialised, rich, democratic

#### **Chapter 1: Introduction**

#### 1.1 WEIRD dominance in the literature

The study of human behaviour has been dominated by work published on Western, Educated, Industrialised, Rich, and Democratic (WEIRD) populations (Henrich et al., 2010). The field of psychology has been driven by approaches and methodologies developed within Western ideology, and WEIRD based research has typically been viewed as universal and factual, regardless of the vast underrepresentation of the global population and lack of consideration for diverse cultures within Western countries (Beins, 2019; Ryba & Schinke, 2009). WEIRD populations represent a minority of the world, with high-income countries (HICs) representing approximately 6% of the population (World Bank, 2021), and it has been argued that incorporating research methods based on WEIRD populations in non-Western countries is potentially detrimental to understanding psychological constructs in different cultures (Hwang, 2023). Human behaviour is highly complex and influenced by multiple factors, requiring an understanding within the relevant cultural context to design and implement interventions for behaviour change (Beins, 2019). Not only do samples in psychological studies need diversifying, but the research methods and tools used to assess human behaviour require further consideration in non-WEIRD populations (Hruschka, 2020).

Four fundamental areas in psychology have been implicated by this WEIRD phenomenon; participant characteristics; theories and models; research methods; and educational and organisational structures (de Oliveira & Baggs, 2023). These aspects directly relate to a core bias in the lack of diversity within participant populations in research studies. One method which has been implemented to diversify samples within psychological research is the practice of Western researchers temporarily visiting countries, predominantly in Africa or Asia, to collect data (Olufadewa, Adesina & Ayorinde, 2020). This is known as "helicopter or parachute research", and this practice has received increasing criticism as potentially reinforcing ethnocentric views and perpetuating colonialism (Singh, 2022). Although this practice can help increase research output and add necessary resources, it largely fails to address the other aspects of the WEIRD problem, such as ensuring theories, models and measures are culturally relevant and tackling flaws within educational and organisational structures (de Oliveria & Baggs, 2023).

Between the years 2003-2007, 68% of participant samples in studies within six dominant American Psychological Association (APA) journals were American, with between 77%-88% of those individuals identifying as European American (Arnett, 2009). An update highlighted that journal articles appear to be becoming more diverse, with just over 60% of authors and samples being American based in the reviewed APA journals (Thalmayer et al., 2020). However, this is largely due to more European, English-speaking participation rather than an increase in global rates (Thalmayer et al., 2020). In both studies, individuals from African countries represented less than 1% of the participants in the dominant APA journals, even although Europe, North America, Australia, and New Zealand combined only represent approximately 14% of the global population (United Nations, 2022). The population of sub-Saharan African (SSA) countries also equates to 14% of the global population, and it is expected to double by 2050, contrasted by an expected decline in population rates in 2030 for Europe and North America (United Nations, 2022). These figures highlight that there is clearly a need to prioritise research efforts in SSA countries and ensure that research methods are developed to be culturally appropriate and representative (Adetula et al., 2022).

#### 1.2 Cultural differences in behavioural science

Many cultural differences exist across countries and regions which can influence behavioural science research, including perceptions of health and health related behaviours. Individualistic cultures tend to place emphasis on singular identities, personal goals, and achievements, whereas collectivist cultures prioritise identities within social groups (Pelham et al., 2022). This cultural difference is outlined in the research, highlighting that Western societies, particularly in Western Europe, are less collectivist than in SSA (Pelham et al., 2022). Research suggests these differences can influence emotional processing and social relationships, including social support seeking during stressful events (Kim, Sherman & Taylor, 2008; Mesquita & Walker, 2003). Societal norms and values also influence psychological constructs, such as well-being and mood (Wilson, Khumalo & Zulu, 2022). Within societies and cultures in SSA countries, health related beliefs can be influenced by religious and spiritual views (Marks, 2005). For example, questions on reproduction may be viewed as God's decision, as opposed to something an individual can influence (Batres, 2018). There are also cultural differences in how information is processed, how question items are perceived, response styles and aspects relating to social desirability, but most of the research on this fails to include SSA countries (Yang et al., 2010).

#### 1.3 Cultural differences influence self-reporting

These cultural differences can affect the validity of research, as most of the literature on the development and validation of self-report measures is derived from participants in WEIRD populations (Henrich et al., 2010). There is a growing number of studies which highlight key challenges relating to inadequate cultural adaptation of measures, limited translation efforts and reporting of validity and reliability studies only conducted in HICs. A large review exploring selfreported adherence and behavioral influences, conducted largely in Asian, Middle Eastern and South American low-middle income countries (LMICs), highlighted that most studies used measures in which there were no efforts to adapt the questionnaire to the local culture of interest. Specific examples of contextual factors which were omitted from the measures were discussed, including items on traditional medicine which were not incorporated in any measure developed in a Western country but was of significant cultural relevance to a number of LMICs. As such, there were several difficulties reported in using the self-report measures, and the overall use of these measures was assessed as inadequate, resulting in unreliable results and limitations on accurately ascertaining self-reported adherence to medication and the behavioral influences which contribute to adherence (Khoiry et al., 2023). Another study of over 50 LMICs around the globe assessed tools used to screen for self-reported behavioral problems and found similar issues with cultural adaptation and translation procedures (Maldonado et al., 2019).

Related to the adaptation of measures, literature has highlighted difficulties with the interpretation of response scales in various LMICs. There is evidence that participants in south American LMICs experience difficulties with Likert scales, including understanding and interpreting the scale points and anchor descriptions (Luetke Lanfer et al., 2021). Similar difficulties were reported in studies conducted in Guatemala, central America, which explored vaccine hesitancy using Likert scale response items (Domek et al., 2018). Many respondents omitted Likert scale questions and others would erase the response options and replace them with dichotomous answers (Domek et al., 2018). Some researchers have hypothesized that the variations in degree points within agreement Likert response scale items may not be interpretable or culturally relevant in some LMICs, due to tendencies of respondents to provide a dichotomous response when presented with Likert scale items (Flaskerud, 2012). The reported challenges of using Likert scales are further discussed in Chapters 2 and 4.

However, it is worth nothing that the limited body of research on questionnaire items and response scales has explored cultural differences between several geographic regions globally, including countries from every other region and African countries which are not in SSA, but SSA populations are largely omitted from the research (Van de Vijver et al., 2017, Yang et al., 2010). The reasons for SSA participants and researchers being underrepresented are multifaceted, but include limited resources, funds, capacity, and power imbalances with research partners from other countries (Mughogho, Adhiambo & Forscher, 2023). Young SSA researchers highlight a lack of support, funding, training, and motivation as factors which influence research activities (Kumwenda et al., 2017). Further research and exploration of these issues is clearly required in SSA countries to help improve research methodologies and tackle the WEIRD bias in behavioural science research.

In light of these difficulties, researchers have developed guidance on methods of adapting measurement tools to varying contexts and cultures. The Survey Research Center (discussed further in Chapter 4) developed guidelines based on research exploring the development of cross-cultural surveys (Survey Research Center, 2016). The Patient-Reported Outcomes Measurement Information System (PROMIS) includes more than 300 measurement tools which have been designed by a committee of international experts to support the translatability of the language, grammar and concepts included in the measures across varying countries. The PROMIS Health Organization have also developed guidelines and standards for the development and validation of measurement tools, including detailed language translation and cultural adaptation recommendations (PROMIS, 2013). A recent publication focused specifically on assessing psychological concepts using culturally relevant methods (Ambuehl & Inauen, 2022). Intended to complement existing translation guidelines, this paper outlines four key steps to adaptation and highlights the importance of initially exploring whether the psychological concept in question is applicable in the context in which it is being studied. These guidelines are invaluable resources for researchers working in LMICs, but more literature is required, specifically in SSA countries, to further support these initiatives and improve the quality of conducted research.

#### 1.4 Research problem

This thesis formed following the author's firsthand experiences of contributing to health partnerships between the UK and LMICs, with the aim of supporting healthcare professionals in behaviour change practices to improve health-related outcomes (Byrne-Davis et al., 2017; Bull et al., 2017). Throughout the author's time working in partnerships within SSA countries, challenges relating to methodological practices, in particular the use of commonly used psychological constructs in the UK and Likert scales, were highlighted as problematic and difficult to interpret. Recent research has highlighted that theoretical approaches and methods used in HICs to explore psychological constructs in behavioural science can be difficult to use in LMICs, and require adaptation (Turner et al., 2023). The author, in conjunction with others in the partnerships, identified that research methods used in the UK may not be directly applicable in SSA countries, and that this is an under-researched area that requires more exploration to support behavioural science researchers in developing culturally appropriate self-report measures to assess psychological constructs.

## **1.5 Aim and structure of thesis**

Overall, the main aim of this thesis is to assess what researchers should consider when developing behavioral science measures to explore self-reported psychological constructs in SSA countries. The following three chapters outline publishable papers with individual research questions that relate to the overall aim. These chapters are summarised as follows:

<u>Chapter 2</u>: a qualitative research study which aims to assess the methodological challenges of using self-report measures in SSA countries, as detailed by researchers working in the field of behavioural science in SSA countries.

<u>Chapter 3</u>: a systematic review looking at what self-report measurement tools have been used to assess the behavioural influences of healthcare workers working in SSA countries, with a particular focus on antimicrobial use as an example area to explore problems with measurements.

<u>Chapter 4</u>: a secondary data analysis study assessing whether response category descriptors of Likert scales are perceived differently by individuals in the UK and individuals in SSA countries.

## 1.6 Changes to thesis

The following studies outlined in Chapters 2-4 evolved drastically over the last four years, due to difficulties in recruiting for primary data collections studies in SSA following COVID-19 related restrictions. Originally, Chapter 3 was designed as a qualitative study involving a think aloud approach to assess healthcare practitioners' views on a questionnaire development in Uganda. Chapter 4 was also proposed as a primary data collection approach to disseminate the

questionnaire developed in Chapter 3 and assess the validity and reliability. However, these approaches were not possible, and the research methods were adapted to answer the main research aim.

#### 1.7 Note to terminology

Throughout this thesis, terminology relating to HICs, LMICs and SSA countries are used. However, the terminology on country grouping by income (HIC, LMICs) directly relates to Gross National Income (GNI) classifications assigned by the World Bank, and the threshold for groupings changes every year. GNI is not always the most appropriate terminology, and it can further instill divisiveness and perceptions relating to colonial and Eurocentric viewpoints (Lencucha & Neupane, 2022). Similarly, it is noteworthy that the vast majority of HICs are either dominated by English speakers, Anglosphere countries and/or are a European country (World Bank, 2021). Countries, in what some academics refer to as the Anglosphere (the US, UK, Australia, New Zealand and Canada), have many commonalities, such as language, historical connection and political systems (Gamble, 2021). WEIRD populations are largely represented within these countries, but this is not a well-defined cultural grouping. For this reason, within the subsequent chapters, terminology relating to HICs and LMICs are used as it relates to the literature in this area and emphasizes the issue of WEIRD populations dominating the literature base but does not necessarily represent the views of the author.

Related to this, SSA is a term used to describe the region of countries which lie south of the Sahara Desert, (full list of these countries is provided in Chapter 3.2). It has been argued that using SSA as a term to define these countries should be changed, as it is outdated, unrepresentative and has links to colonisation (Haldevang, 2016). This description is still used by major organisations (such as USAID and the United Nations), at times in combination with individual countries and/or groupings of clearer geographical regions, for example East Africa, Central Africa, etc. For this thesis, the term SSA is used as research has been undertaken in some of the African countries which are not represented by the SSA term, and this definition is used broadly to include a breadth of countries in this region. It was hypothesized at the start of the thesis, based on the authors knowledge, experiences and literature searching, that there would not be a vast amount of literature on the development and adaptation of self-report measures in SSA countries. Limiting the countries to defined geographical regions within Africa would have restricted the thesis considerably. The author has no intention of linking any terminology used in this thesis in a derogatory manner.

# Chapter 2: Are self-report questionnaires in behavioural science too Western? Exploring the views of researchers working in sub-Saharan African countries

# 2.1 Introduction

Cultural adaptation of measurement scales and their international use is one of the most important areas in psychological assessment (Lopez-Roig & Pastor, 2016). Testing equivalence of measures is important for psychological research and particularly for cross-cultural research, as a key consideration for adaptation procedures is to achieve the best possible level of linguistic, cultural, conceptual, and metric equivalence (Muniz et al., 2013). However, psychological research has largely been dominated by WEIRD populations, as outlined in Chapter 1, leading to biased samples in theories, methods, and research procedures (Henrich et al., 2010) and most of the research in psychological and behavioural science is still based on this population. Published papers largely fail to acknowledge biases in generalising findings (Apicella et al., 2020).

In behavioural science, the design and implementation of effective evidence-based behaviour change interventions are crucial to overcome health-related challenges which affect many populations (Michie et al., 2011). Several of these are related to individual behaviour, hence the requirement for behavioural change interventions. For example, in the management of chronic health conditions, medication adherence, attendance at physician appointments and lifestyle changes are key behaviours (Chauhan et al., 2017), as are infection control practices and appropriate medication prescriptions for antimicrobial resistance (Holmes et al., 2016). Behaviours are complex and challenging to change, and the identification of factors that influence these behaviours is necessary to develop effective interventions (Michie et al., 2008). Many influences on behaviour are self-reported by individuals and measured using quantitative tools, although qualitative methods can also be used to collect non-numerical information and gain further depth (Hall et al., 2016; Leggett et al., 2016). Self-reporting in a questionnaire or survey is the most common measurement method used in research, and particularly in behavioral science (Martinez et al., 2014).

Using self-report measures has several benefits, including practicality, being relatively low cost, and allowing researchers to ask respondents to directly report the factors that influence behaviour, including beliefs, attitudes, and intentions (Chambers & O'Carroll, 2016). However,

to avoid biases which can affect the measurement properties, the measures used need to be assessed for validity and acceptability (Chambers & O'Carroll, 2016; Martinez et al., 2014; Squires et al., 2011). The international initiative EQUATOR highlights the importance of using valid, reliable, and acceptable measures in research (Ogrinc et al., 2015; Tate et al., 2016). Quantitative self-report approaches typically incorporate rating scales to elicit responses from participants, in particular Likert scales, and several conditions are required for a rating scale to be used effectively. The question items used to measure the construct must be reliable and valid, each respondent should have a relatively precise and stable understanding of the meaning of each point on the scale, and respondents must agree in their interpretations (Krosnick & Presser, 2010). If some of these conditions are not met, data quality is likely to be affected (Krosnick & Presser, 2010).

In LMICs where questionnaires developed from Western based countries are utilised, differing cultural, methodological, and medical considerations can cause substantial bias if questions have not been appropriately evaluated or modified to the local context (Ares, 2018; Nielsen et al., 2017, Squires et al., 2011). This is also a challenge for Western educated researchers developing questionnaires to assess self-reported psychological constructs in LMICs, due to the bias in WEIRD samples discussed in Chapter 1. Consideration is required in relation to the pragmatics of language in specific cultural contexts to ensure the same meaning for psychological constructs, and other there are other concerns relating to comprehension and semantic equivalence (Lopez-Roig & Pastor, 2016). Research highlights that individuals from different cultures may have systematically different understandings of response scale terminology, for example response categories can have different meanings across societies (Beaton et al., 2000). Preliminary evidence suggests several possible explanations for this, including difficulties in directly translating adverbs used in the scales and semantic differences in the intensity associated with adverbs across cultures (Chung & Han, 2013; Harzing, 2006; Lopez-Roig & Pastor, 2016; Voss et al., 1996). Guidelines exist outlining procedures for crosscultural adaptation of existing self-report measures, including translation procedures (Beaton et al., 2000). However, there are difficulties with conceptual equivalence across cultures which can affect translation practices. The literature which these guidelines are based on is largely derived from Western and/or Eastern countries, failing to represent SSA countries, and are very general and focus on multinational, multicultural or multiregional studies (Harkness et al., 2016).

Likert scales appear to be particularly problematic, as briefly discussed in Chapter 1, and research demonstrates difficulties in the validity and reliability of these scales in LMICs (Borghi et al., 2018; Bull et al., 2017). A review of Likert scales used in 26 countries highlights disproportionate extreme response style biases in LMICs linked to understanding and acceptability (Harzing, 2006). Response styles refer to participant's tendency to systematically respond to questions in a specific way, regardless of the content (Baumgartner & Steenkamp, 2001). Consistent and systematic variances in response style among countries are evident in populations from Western, European countries, and populations from Eastern, Asian countries, with the latter emerging as reluctant to express negative opinions and thus more extreme positive responses are outlined in questionnaires (Harzing, 2006, Harzing et al., 2011). However, the African continent is largely omitted from research, and countries which are included from Africa are largely Northern, Arab countries, omitting SSA countries. This raises fundamental methodological and ethical concerns, as studies use unreliable approaches to reduce bias in SSA contexts, including the exclusion of participants who have difficulties in understanding questions response options and citations related to previous validity assessments from HICs, which also creates selection bias (Ares, 2018; Hall et al., 2016). To avoid these difficulties associated with Likert scales, reviews of studies utilising self-report questionnaires highlight inadequate attempts to reduce biases and provide minimal information on piloting, validation, and adaption of questionnaires in LMICs (Agampodi et al., 2015; Lotfi et al., 2016; Nielsen et al., 2017).

With little research exploring the use of self-report measures for psychological constructs in LMICs, it is difficult to determine if there is a problem with using and adapting behavioural science measures in SSA countries, due to the lack of work in this area. However, as most of the research on behavioural science and behaviour change interventions has been developed and conducted in HICs, and little methodological research considers SSA countries, there is an urgent need for more research investigating these methodological concerns in SSA. Currently, there are no accepted guidelines for designing and/or adapting self-report questionnaires specifically in SSA countries, as guidelines focus more generally on LMICs as whole (for example Carter et al., 2020; FAO & The World Bank, 2018). This is problematic, as there are a vast number of LMICs from various continents, and these guidelines does not consider cultural differences. Therefore, there is a need to establish robust questionnaires to produce valid research findings. To the authors' knowledge, no previous study has specifically explored

researchers' experiences of the methodological challenges in using self-report questionnaires in SSA countries.

The aim of this study is to explore behavioural science researchers' experiences of the methodological aspects of using self-report questionnaires in SSA countries, to understand: (1) the key methodological challenges of conducting self-report questionnaires in SSA countries and (2) the approaches which have previously been used to improve self-reporting and to understand researchers' perceptions of these.

#### 2.2 Methods

A qualitative approach was used to collect data from experienced researchers who had worked in SSA countries. As this is a relatively unexplored area of research with little background literature, a qualitative approach was determined as the most appropriate method to explore and discuss methodological considerations specifically relating to behavioural science research in SSA countries. To limit researcher bias and openly investigate methodological considerations, the focus of the interviews started with general questions on both quantitative and qualitative approaches used in behavioural science research, including challenges and solutions with these approaches. Discussion continued to the types of quantitative measures and response scales researchers had used, and their experiences of how well these approaches worked.

#### Theoretical perspective

A constructivist position was applied to this research, acknowledging that the world is constructed by individuals' thoughts and activities, and that creating meaning is viewed as a social process, shaped by social interactions, language, and culture. The aim of this approach is to understand different meanings by which people in different contexts make sense of the world, their lives, and social processes (Marks & Yardley, 2004). In terms of this study, the author produced themes through interpreting the meaning and importance of codes in relation to the research aims, and the level of importance expressed by participants, as opposed to solely considering the reoccurrence of themes throughout transcripts. This involved an interpretivist approach to understanding the viewpoints of participants and their views on how reality is socially constructed (Bogdan & Biklen, 1998). The initial coding and theme development followed an inductive process, and the data was open-coded, exploring individual interview excerpts and then reviewing similarities across the interview transcripts to understand the data from participants without preconceived notions in mind. Based on what was most relevant for the development of themes, the researcher pulled codes together into specific themes and ensured those themes were meaningful to the specific research aims.

#### Design

#### Semi-structured interviews

A cross-sectional qualitative approach was applied to answer the research aims and a semistructured interview guide was written by the author and reviewed by members of the supervisory team. This structure allowed flexibility to ensure that the topics and views which arose from each participant were explored in further detail. The interviews were conducted between October 2019 and January 2021 by the author, who is a Chartered Psychologist, has an MSc in Health Psychology and British Psychological Society Stage 2 qualification in Health Psychology, and experience in conducting behavioural science research in SSA countries and qualitative analysis.

#### Ethics

This project was reviewed and approved by the University of Aberdeen School of Medicine Medical Sciences and Nutrition Ethics Review Board (CERB/2019/7/1809; July 2019), the institution in which the author was working at the time of protocol development and study completion.

#### Procedure

#### Participants and recruitment

A purposive strategy of recruitment was employed, with invitation information posted on Twitter, Facebook and ResearchGate groups for behavioural scientists and psychologists working in LMICs. An invitation email was also sent to the European Health Psychology Society Equity, Global Health and Sustainability Special Interest Group. A snowball method was also used with researchers forwarding the study invitation to other relevant researchers to enhance reach. The invitation information highlighted the purpose of the study, inclusion criteria and included an email address to contact the author. Participants were informed in the invitation information that the interviewer (thesis author) had experience of conducting behavioural science research in SSA countries and was a health psychologist interested in exploring the experiences of other researchers, in relation to methodological considerations specifically. Although the interviewer had experienced specific methodological challenges in past research, participants were not informed of this and the interview schedule was designed to explore strengths and challenges, to limit bias. Once interested individuals contacted the author, the participant information sheet (PIS; Appendix 1) and consent form (Appendix 2) were sent by email for additional information. If at this point the individual wanted to proceed, a time and date was arranged to conduct the interview and the consent form was completed prior to the interview, either online or in person. Twenty-five people enquired about participating in the study and fourteen were interviewed; five researchers did not meet the inclusion criteria, three researchers had worked in other LMICs but not in SSA countries, and three individuals did not undertake research in health psychology or

behavioural science. No participants refused to participate or dropped out. Further details on the participants are outlined in Table 1.

# Table 1

# Participant information

Participant	Research focus	Region at time of	Countries research undertaken	Number of years
number		interview		worked in research
1	Health Systems Research	Europe	Malawi, Burkina Faso	8 years
2	Public Health & Environmental Heath	SSA	Malawi	14 years
3	Environmental Health	SSA	Uganda	10 years
4	Behavioural Scientist	SSA	Tanzania	15 years
5	Environmental Health	SSA	Uganda	2 years
6	Health Psychology	Europe	Democratic Republic of the Congo, Rwanda, Ghana	3 years
7	Environmental & Health Psychology	Europe	Ethiopia, Kenya	12 years
8	Counselling Psychology	SSA	Sudan, South Africa, various countries across SSA	21 years
9	Health Psychology	Europe	Mozambique, Tanzania, Ghana	10 years
10	Health Psychology	Europe	Ghana	11 years
11	Health Psychology	Europe	Uganda, Ethiopia	14 years

Participant	Research focus	Region at time of	Countries research undertaken	Number of years
number		interview		worked in research
12	Clinical Psychology	SSA	SSA	31 years
13	Health Psychology	Europe	Tanzania, Mauritius	13 years
14	Epidemiology & Health Psychology	Europe	Tanzania, Ethiopia, The Gambia	8 years

The inclusion criteria consisted of researchers who conducted research in health psychology or behavioural science (thus research involving the measurement of psychological constructs) in SSA countries. To ensure the inclusion of participants who were currently involved in collecting data in the field, obtaining a range of viewpoints from research leads to research assistants, no criteria was set in terms of length of experience. All participants were currently involved in research projects in SSA countries, six participants were nationals of an SSA country (those in Table 1 who were based in a SSA country at the time of the interview), and eight participants were nationals of a European country.

#### Data collection

The interview topic guide is provided in Appendix 3. Participants were offered a range of interview modalities to maximise recruitment and limit potential issues relating to internet connectivity in some SSA countries. Most interviews were conducted one to one online with video call through zoom or MS Teams (n=9), four interviews were conducted online with audio only through zoom and one participant was interviewed face to face. All interviews were audio recorded, with only the participant and interviewer present, and both field notes and debrief notes were written and recorded. The length of interviews ranged from 25-58 minutes (mean = 43 minutes). Interview recordings were transcribed verbatim and coded concurrently. Six rounds of recruitment drives were conducted during the recruitment phase to increase participation due to the specific inclusion criteria. All fourteen people who met the inclusion criteria engaged with the interviews, and initial coding of the final three interviews revealed no new themes.

#### Data analysis

The coding and analysis process was influenced by Braun and Clarke's initial guidelines for thematic analysis and work on reflexive thematic analysis (Braun & Clarke, 2006; Braun & Clarke, 2019). The author became familiar with the data by transcribing each interview individually and taking notes whilst reading the transcripts. The transcripts were checked against the recordings and re-read several times prior to initial coding. Thoughts and feelings in relation to the transcriptions were noted in a reflective log online. During the initial coding process, the author met with one of the supervisors to debrief on the initial coding process, review more complex codes and check the research process. The author (female) was aware that the second coder was crucial to mitigate Western and WEIRD viewpoints and include experience and understanding from a non-Western perspective. A second coder independently coded four transcripts (29%) and acted as a critical friend, questioning understanding and

discussing interpretations to support the reflexive process and enable a dialogue to review initial coding and theme development (Folger, 2010). Both researchers met several times and discussed the initial coding and generation of initial themes. The second coder (male) is an experienced qualitative researcher based in Uganda and supported collaborative reflection through providing alternative perspectives and thoughts on the data. Both coders met and discussed the two perspectives on initial coding, with comparable codes developed by both coders.

Following this, patterns were explored across the dataset and grouped into themes, through regularly checking theme development against the transcribed raw data. The process of developing themes involved taking an inductive and data-driven approach, ensuring that all data from the interviews and notes were kept in mind. Both themes and subthemes were stored and managed in NVIVO12, and both coders met again to review at this point and reached a team consensus on final themes. The author met with supervisors throughout the theme development process to discuss themes and debrief. Well organised files with a clear audit trail of code and theme development were maintained throughout this process, an example of which is included in Appendix 4.

#### Researcher reflexivity

The concept for this research developed from the experiences of methodological issues with self-report scales in several SSA countries by the author (detailed in Chapter 1). Anecdotally, other researchers had discussed similar methodological concerns, but there was scarce research on this topic. Throughout the process, the author was aware of predetermined assumptions regarding the research data, and although this was advantageous for data collection purposes, the author was aware of the potential issues this could pose. These assumptions included experiences with difficulties using response scales in SSA countries and challenges with Likert scales.

In terms of data collection, the prior experiences and understandings of the author allowed for an in-depth consideration of issues in the interviews, with important probes and follow up questions to explore the experiences of participants. The interview schedule was written with the author's experiences in mind, however the focus on reflexivity meant that the author reviewed and considered the questions in detail to write an open-ended and non-biased topic guide.

# 2.3 Results

Overall, nine themes were developed through the analytic process. These are outlined below and discussed in terms of the research aim which the themes relate to.

# (Aim 1) What are the key methodological challenges of conducting self-report questionnaires in SSA countries?

Two overarching, main themes reflect the underlying sub-themes relating to this research aim:

- Conceptual understanding of questions and scales
- Differences in research process in SSA countries

Sub-themes relating to the overarching theme of *conceptual understanding of questions and scales* are:

- Differences in conceptual interpretations compared to Western approaches
- Unfamiliarity with Likert scales
- Difficulties with translation and adaptation of measurement items and Likert response scale labels

Sub-themes which relate to an overarching theme of *differences in research processes in SSA countries:* 

- Social desirability bias
- Researcher administered questionnaire delivery can lead to interviewer bias
- Response fatigue in over researched/overburdened communities

## Conceptual understanding of questions and scales

Differences in conceptual interpretations compared to Western approaches

Several participants described difficulties with assessing specific psychological constructs when using self-report questionnaires. Examples of psychological constructs which raised difficulties included motivation, expectations, social influences, social norms, social support, and depression. Individuals referred to question items as often either not making sense to participants, participants not understanding what the question item was specifically asking, or that the items just did not work. This difficulty in assessing psychological constructs appeared to

occur at the researcher, interviewer, and participant level, as several individuals described differences in conceptual understanding of psychological constructs, which affected the measurement of these constructs:

"The term motivation in West Africa means financial there have been so many donor projects coming with financial incentives .... (researchers) conjured up the term motivation or you motivate me means I give you something basically. So something people have stumbled over a lot in sub-Saharan Africa, a lot because when you ask them what about motivation, they'll be like, yeah, I didn't get enough money." *(P1, Health Systems Researcher)* 

The conceptual differences in interpreting psychological constructs were reported as a common challenge when research teams used Western developed measures or developed their own measures using Western conceptual understandings of constructs:

"In other constructs such as stress and any psychological constructs really, it is important to consider constructs and the differences in understanding, for example we have had difficulties with the DSM descriptors of depression and how depression is perceived in our populations, and it is important to collect qualitative data from interviews on what depression means to people and their subjective experiences, rather than just the instruments." *(P8, Counselling Psychology Researcher)* 

Other participants stated that applying measures using Western developed constructs in SSA countries is not the correct approach, and instead research needs to focus on understanding the application and relevance of psychological constructs specifically in SSA countries, rather than adapting Western measures which may be ill-fitting in populations with varying interpretations:

"I think we should actually go back to basics .... we've developed this whole paradigm of health behavior and these questionnaires and, you know, based on Western thinking and, and actually, I would, I would try and go back to basics and say, okay, so if we develop as an African behavioral science .... what are the best ways of assessing particular constructs." (*P4, Counselling Psychology*)

Some participants highlighted that there are even difficulties with self-reporting of psychological constructs in Western countries, further emphasizing the need to thoroughly consider the relevance of Western developed self-report measurement tools in SSA countries.

### Unfamiliarity with Likert scales

Across certain participant populations, researchers outlined that some participants had not previously been asked to rate their responses in scale format, with most reporting a Likert scale format with verbal anchors as opposed to numerical values.

"People are not used to this, I don't know, (pause) I've grown up in a world where from as young as I can remember, I was asked to rate this is better than this, this is more than this, I prefer this to this and so on .... My world has always been a very rankable and quantitative for one, and I feel that that's not necessarily the case everywhere." (*P1, Health Systems Research*)

Participants' responses were contradictory in terms of the populations affected by unfamiliarity, with a couple of participants stating that this is less of a problem with student and educated populations, due to an increase in the use of rating scales in academic contexts, whilst other participants stated that this was an issue with all populations, including students and youth. Other participants discussed unfamiliarity as a general problem in non-WEIRD populations:

"Likert scales, okay, well, well, I have two words to say to you about that and the words are good luck (laughing) .... where does the idea of the Likert scale come from? where do they come from? what kind of assumptive world do they come from? .... [Likert scale] an ordinal measure, a ratio that you can assume, that the space between a three and a four is the same as the space between a two and a three. Which is, you know, you and I know, these are fictions. But these, but these fictions are culturally elaborated within the sort of culture that we come from, which is sort of the English-speaking, dominant culture, where people are familiar .... with these, these sorts of things." (*P12, Clinical Psychology*)

The utilisation of measures including rating scales with populations who have had no experience with this type of response option, in prior research contexts but also in cultural contexts, can affect the validity of tools and increase the likelihood of measurement error.

"You don't understand what it really means (agreement scale descriptors), you don't even understand what strong, like very strong or strong or medium mean, or all those kind of things....we find it a challenge with people who are educated....even with University students, this kind of questioning, I don't think that's what we are used to in this part of the world, so we just select anything (laughing)." (P2, Public Health & Environmental Health Researcher)

# Difficulties with translation and adaptation of measurement items and Likert response scale labels

Specific difficulties with interpreting the response scale labels used in Likert scales were highlighted by most participants. This largely referred to the adverbs used in the response scales to indicate the strength of the response highlighted by participants. This relates to the theme of unfamiliarity of Likert scales, but it was interpreted as a distinct theme due to the specific emphasis on the interpretation and translation of the chosen adverbs:

".... even with University students, this kind of questioning, I don't think that's what we are used to in this part of the world, so we just select anything (laughing). You don't understand what it really means, you don't even understand what strong like very strong or strong or medium mean or all those kind of things, so with a village is also quite a big challenge as well, if we are, we find it a challenge with people who are educated" (*P2, Public Health & Environmental Health Researcher*)

P2 interestingly uses the first-person term "we" several times when discussing the issues with interpreting the adverbs used in Likert scales, which implies that this researcher includes themself in the collective group of individuals who do not understand how to interpret the adverbs used. If the use of adverbs in rating scales is not interpreted in a meaningful way by researchers employing these tools, and the populations being assessed, then the validity and reliability of the study results is called into question, and again this can lead to measurement error.

Furthermore, there are additional issues with adapting tools to ensure the tool is translated adequately, not simply by direct translation but through using thorough translation guidelines to ensure that translated words make sense to the target population. In certain languages, adverbs

do not directly translate, and words used to replace adverbs may not make sense in a scale context:

"These ones [Likert scales], I would say in simple terms, currently they are not working. I don't know if we're gathering the required data but they are not working. Probably it's because sometimes we're using tools that are already developed, of course we say we have adapted them to our situation, but sometimes the adaption may not be adequate. So, we may need to find ways and means of developing our own tools and also finding ways and means of putting in the responses, because the response, for example, I strongly agree, I don't agree, if you use the local language, it doesn't make sense." (*P2, Public Health & Environmental Health Researcher*)

This quote above further highlights the need to develop tools using a bottom-up approach, specifically the response options, as using response options developed from WEIRD populations can lead to measurement error.

Researchers who had been educated and/or predominantly lived in Western countries suggested that this difficulty with translating and adapting Likert response options was an aspect of the measurement tool which they had not anticipated having problems with. This led to a couple of researchers expressing frustration at not being aware of this prior to the design of research studies and development of measurement tools:

"I had the impression that it was very much, uh, a language thing. So, that in they just didn't have these types of words in Swahili. You agree or you disagree, you don't somewhat disagree, what's somewhat disagree? So, um, yeah. So, I think that was an important, yeah, yeah. To be honest, I mean, it makes complete sense. So, it's annoying you know, when you start to ask these type of very basic questions about things that we just habitually do all the time." (*P4, Behavioural Science*)

#### Differences in research process in SSA countries

#### Social desirability bias

Several participants described concerns relating to social desirability bias in the use of selfreport measures, a form of response bias which can lead to ceiling effects. A few participants highlighted this as particularly problematic with healthcare worker populations, which some participants suggested could be related to specific expectations in healthcare contexts, and potentially expectations of the distribution of research output:

"It might just be because people want to be seen as good people. And it's really hard to get past that. There's a lot of like group compliance and really wanting to be seen good, at least in this cultural, in this specific cultural of the hospital as well, because obviously your expected to be good to your patients and everything. It's part of their identity. It's hard to delink these things. And yeah, these were the hardest things." *(P14, Epidemiology & Health Psychology)* 

Some participants discussed healthcare workers awareness of the behaviours which they should be adhering to, and that expectations in relation to who may view research outputs and the consequences of this can influence participants responses, leading to responses which are expected answers as opposed to the truth.

Participants may be reluctant to report negatively on self-report scales, as the social norm within specific cultures aligns with being positive and not discussing negative aspects or being seen to complain. Two participants raised response bias as an issue, highlighting that there is a need to consider the specific cultural context when considering the most appropriate measurement tool to use:

"The problem in Malawi, like in many countries in the world, is that people don't actually, if you give them a choice to respond fairly positively while still expressing sort of a bit of variation in their answer, they will use that fairly positive end of, of the scale. And people in Malawi specifically, there's no such thing as saying something negative .... So, one issue was that in Malawi, we had I think we used a four or five point scale, I can't remember and essentially everybody chose five or four maybe, there was no variation in the data." *(P1, Health Systems Research)* 

#### Researcher administered questionnaire delivery can lead to interviewer bias

All participants discussed the different data collection methods used to obtain self-report responses: participant administered, part-participant administered, and researcher administered, describing the data collection method as important in terms of methodological considerations. There were different reasons highlighted for using researcher administered methods, including to increase the response rate and to support populations who may experience literacy difficulties. Although these reasons also apply in other countries, increasing the response rate in SSA countries is particularly relevant as most of the research is conducted using paper methods, as opposed to online methods.

However, it is evident that there can be problems with interviewer bias and consistency when researcher-administered questionnaires are utilised:

"Normally a research assistant sometimes asks a question and then depending on how the respondent responds, the research assistant has to decide on or agree where they think the participant falls as opposed to somebody picking for their own. Sometimes it really depends on the research assistants' own judgement, that depending on how he has answered this question, I think the answer should be very good or good or not good. And that's, that again also brings in subjectivity because different research assistants could interpret different people's responses differently." (Participant 3, Environmental Health)

This is a unique approach to self-report measurements, which several participants described as an attempt to prevent social desirability bias, in addition to increasing response rates and for populations with literacy difficulties. Participants described researchers not reading the scale to participants, and instead asking respondents in a binary type format to respond to the questionnaire, and then determining the strength of that based on qualitative type data from a discussion. This raises issues with measurement errors and validity caused by potential interviewer bias.

#### Response fatigue in over researched/overburdened communities

Respondents expressed concerns with communities in SSA countries being targeted for research projects, and thus risking becoming overburdened with researchers and measurement tools, highlighting concerns with the responses provided if participants experience research fatigue. These over-researched populations may receive little benefit, from numerous research studies and questions were raised regarding the validity of responses from participants in these communities:

"I think you always need to consider do not harm because the more questions we will ask, the harder it will be for the next research team, because if people feel okay, I will not get any benefit from, from answering these questions, then they'll stop participating in studies. Because one thing someone told me, or someone told one of my colleagues was, um so, you will earn all your money by asking these questions and we will not get anything from it. And yet and then, it's true because the interviewers will earn money by asking questions, I will earn money by having research data, but the participants, in the end, if it's, for example, only for validating scales, they will not have any benefit from it." (*Participant 6, Health Psychology*)

This was highlighted as a challenge particularly in relation to research conducted by research groups and organisations from Western countries:

"I think the potential for exploitation is considerable when people collect data among poor populations and never be seen again. The researchers are often gone, you know, and leave the community without any benefit. I think it would be helpful if there were specific benefits to the research, not only to the researcher who is building a career and going off to fancy conferences and giving talks, etc. People need to benefit from that as well .... I am kind of against the parachuting model of research, where you parachute in and collect your data and go off again, you know, etc., etc." (*Participant 8, Counselling Psychology*)

## (2) Approaches which have previously been tested to improve self-reported measures

The themes relating to this research aim are as follows:

- Expertise in local context crucial in questionnaire development
- Investment in preparation essential for adaptation and development of measurement tools
- Using qualitative methods to improve understanding

## Expertise in local context crucial in questionnaire development

Although all participants who were Western based had discussed consulting and partnering with local researchers in the research process, most participants also highlighted the need for

Western based researchers to involve local researchers with particular expertise in the research topic area at the point of research design. Responses indicated that prior to determining the measures which will be adopted in studies, local expertise is needed to evaluate the constructs being measured, the questionnaire items and response scale options:

"I think we've got to be very careful about parachuting in. And I think all of these questionnaires need to be reviewed by people who live in the country and who work in the country and who are being brought up in that culture to have an input in how the wording is. And I'm not sure if it's necessarily about changing the scale, which has a four-point scale, but how you use the scale and the questions you would ask as part of that scale. I think perhaps somebody looking at it, asking the question in a way that makes it much more around the behaviour that people would normally be doing." *(P11, Health Psychology)* 

Furthermore, consideration of how to avoid gaining research data and then leaving populations with little benefit needs to be considered in research in SSA countries. This ongoing issue was conveyed by participants as not only relevant in the use of self-report measures, but for research in general.

Several participants highlighted that exploratory discussions are crucial with local researchers in the measurement development stage, rather than imposing Western developed tools and attempting to adapt the measures to meet the needs in the local context:

"I think again I would be looking at, asking people who are in living in that context, brought up within that context, part of that cultural group, thinking about what the language would be, what how would you ask that question? So rather than presenting that question, I'd say so what are some of the ways that people would define doing something, not at all or a great deal in this country." *(P7, Health Psychology)* 

This further supports earlier discussions in relation to taking a bottom-up approach to designing measurement tools in SSA countries; rather than trying to adapt a measure to make it work, start from the beginning again in terms of measurement development.

Investment in preparation essential for adaptation and development of measurement tools Linked to the theme above, participants described the need for resources to conduct initial exploratory work prior to designing and developing measurement tools. A range of ideas were provided, including conducting research specifically to assess measurement tool validity and reliability before focusing on the research topic, piloting tools in depth ensuring that response scales and question items were considered, and conducing feasibility studies to test the feasibility of certain measurement tools.

Some participants advocated that researchers need to budget for resources to enable this preparation work to be conducted thoroughly.

"Go more slowly. Do not think about these methods as things to overcome. Think about them as part of your research process and get funding and promise to create outputs, because that's what, what funders want, to create outputs about these kind of methodological issues, which is why your work is so incredibly important." (*P12, Clinical Psychology*)

"I think again I would be looking at, asking people who are in living in that context, brought up within that context, part of that cultural group, thinking about what the language would be, what how would you ask that question? So rather than presenting that question, I'd say so what are some of the ways that people would define doing something, not at all or a great deal in this country." *(P7, Health Psychology)* 

#### Using qualitative methods to improve understanding

Several researchers stated that they found it more beneficial to use qualitative research to explore and measure psychological constructs. These researchers, and others, highlighted that combining qualitative data collection methods with quantitative self-report measures can help to assesses any limitations in the questionnaire data:

"I don't think we had a great deal of reliability and validity [in the self-report questionnaire], but at least we could figure that out. And at least we had a variety of other measures. We had measures of actual behavior. We had qualitative interviews. We had our own structured observations. So, in a sense, we could deal with the
weaknesses of that data by triangulating from various other kinds of data." (*Participant* 13, *Health Psychology*)

Another participant outlined a specific example of times when it becomes evident through additional discussion during interviewer administered quantitative questionnaires that participants do not understand the question and/or response scale:

"I think sometimes, um, for example, they would realise afterwards that some of them that they sometimes didn't understand the question or when they were seeing the latrine or people were telling them like okay, I have a latrine, like a toilet, but they said it and that just meant it like okay, they had a, they had like a pit. They didn't have a toilet. So, they were just telling the interviewer, yeah I have a toilet, I constructed one. So, they would also notice by the talking between questions, this person didn't understand the scale now because the person told me I don't know. The person told me that they didn't like their toilet, so, they didn't want to use it. And then on the next question, they will ask, how much do you like the toilet? And then they will pick, very much so. Then they will notice, okay, this person seems to not understand the scale." (*P6, Health Psychology*)

This difficulty in understanding either the question and/or response scale would unlikely have been identified if the interviewer had not discussed this further with the participant, a discussion which was out with the actual questionnaire administration.

## 2.4 Discussion

The analysis of participants responses highlights several challenges in using self-report measures in SSA countries, and potential approaches to address some of these challenges. Several themes relate to conceptual understandings of both question items and response scales within self-report measures in general. Most of the sub-themes within this theme focus on challenges with the use of rating scales, with all participants highlighting issues relating to Likert scales. Translation, adaptation, and conceptual equivalence were also key challenges raised by participants. In terms of research processes, difficulties with social desirability, questionnaire administration and research fatigue were discussed by participants, some of which are unique to settings in SSA countries. Utilising the experiences of local researchers, investing in adaptation and development, and incorporating qualitative methods were discussed as potential approaches to address some of these challenges.

## Differences in conceptual interpretations

The challenges raised with conceptual interpretations and understandings highlight potential problems in terms of questionnaire validity and reliability. Participants discussed these difficulties largely referencing psychological constructs and measurement tools used in WEIRD populations, and how interpreting this into SSA countries is complex and problematic. But some highlighted that the attention should shift to understanding and exploring psychological constructs from SSA populations, rather than applying the understanding from WEIRD populations to these contexts. These differences in conceptual interpretations of psychological constructs are an interesting insight and have gained little attention in the published literature. This also relates to ensuring that the research techniques used to assess psychological constructs and develop questionnaires to explore this further in SSA populations should consider conceptual equivalence, and not just whether the question makes linguistic sense (Lopez-Roig & Pastor, 2016). Participants' views support the beliefs of other authors, who state that there is a need to review research development processes, "strip back to the foundations" and determine how things work in SSA countries (Lages et al., 2015).

The difficulties raised in general unfamiliarity with, and translation and adaptation of Likert scales, is evident in the literature on self-report measures in SSA countries (Borghi et al., 2018; Lotfi et al., 2016; Nielsen et al., 2017). Some participants indicated that these difficulties were unexpected, others argued that these difficulties did not apply to certain groups of people, such as students and educated populations. These potential challenges are complex and not well

documented or discussed in research. Only recently, as outlined in Chapter 1, has the WEIRD phenomena and potential implications of this gained attention (Hruschka et al., 2020). Although participants had mixed views on why and with whom these challenges occur, this appears to be a potentially significant limitation of Likert scales. This has implications for the development of new tools, and for the use of standardised measurement tools which incorporate Likert scales and requires further consideration.

Although there are guidelines for translating questionnaires in cross-cultural research projects, this study emphasizes that translation requires thorough consideration, arguably more than following guidelines. Response scales adverbs were discussed as particularly difficult to translate in SSA countries, not just in terms of words directly translating but also in terms of the adverbs used being unfamiliar to participants within this context. This relates to the discussions on unfamiliarity with Likert scales, and further emphasizes the need to explore the familiarity, acceptability, and relevance of response scale options with populations in SSA countries.

## Differences in research processes

The concept of social desirability was discussed by participants as an issue which was particularly evident in healthcare worker populations, likely due to concerns expectations from management and researchers, and social norms within cultures. This is problematic across research practices in various countries, not just SSA countries. Social desirability bias has been discussed in depth in the literature and is a large methodological issue in the use of self-report measures (Krumpal, 2013). However, participants highlight that it appears to be more problematic in SSA populations and should be considered by researchers working in these contexts to allow the incorporation of methods to potentially minimise this effect.

Question administration and response fatigue were also discussed as specific concerns in SSA populations. Certain participants discussed data collection techniques which may have been used to minimise problems with literacy issues and social desirability, but these methods increased the likelihood of researcher bias and could potentially affect significantly affect results. Clear considerations should also be given to issues relating to response fatigue and the amount of research which has been previously or is currently being undertaken in certain populations. This further highlights the issues briefly discussed in Chapter 1 with the "parachuting" approach and could potentially affect the validity of future research in overburdened communities.

## Translation

The frustrations highlighted by some researchers regarding difficulties with translating measurement items and response scale options indicate that many researchers, in particular those who are new to research in SSA countries, do not anticipate the level of difficulty which reportedly can arise when seeking conceptual and semantic equivalence. Research has explored difficulties in directly translating adverbs in scales and semantic differences in the intensity associated with adverbs across cultures (Chung & Han, 2013; Harzing, 2006; Lopez-Roig & Pastor, 2016; Voss et al., 1996), but this has largely omitted SSA countries and is not readily available in the evidence base. Without prior knowledge or awareness of these difficulties, how can researchers be expected to prepare, plan, and develop tools to explore and limit these issues affecting the validity of the data?

## Approaches to improve self-reported measures

Participants outlined several methods to improve the use of self-report measures, which could be utilised by researchers in the future as a means of addressing some of the issues relating to the WEIRD phenomena. The discussions on including local experts in questionnaire development may appear on obvious approach, but it is surprising the number of participants who stated that this was needed and not always widespread practice. This could strengthen the measurement validity and help to tackle the WEIRD phenomena (Henrich et al., 2010; Hruschka et al., 2020). Investing in time and resources to support the development and adaptation of measurement tools, incorporating key points outlined above in relation to the challenges of conducting research in SSA countries, were also stated as important. This links to the discussion on thoroughly exploring and understanding how psychological constructs may differ in SSA countries, and thus influencing the measurement practices associated with this. Qualitative methodologies were highlighted as a way to potentially explore this further, prior to undertaking research and/or in combination with quantitative methods to explore items in more detail. These approaches could be incorporated by researchers as methods to limit problems which may arise when collecting data through self-report measures in SSA countries.

### Limitations

There are several limitations to the research. The reliance on memory recall of participants, to consider methodological considerations of projects conducted in SSA countries, may have restricted the responses provided by participants. Most projects do not focus on methodological considerations, this is typically an afterthought which may be difficult for participants to fully

remember. Sub-Saharan Africa is a large geographical area, compromised of 1.14 billion people across 46 countries (The World Bank, 2021). There are vast cultural variations between these countries, which may further affect the validity and reliability of self-report measures which were not explored in depth in this research. There were themes developed from the data which are arguably relevant to other populations, not just SSA countries, and some of the challenges raised by participants are also representative of wider challenges with self-report research. For example, differences in interpretation of psychological constructs and the use of mixed methods have been discussed in the literature in Western countries. In addition, a combination of approaches was used to conduct the interviews, as both online and face to face approaches were utilised.

## Strengths

To the best of the author's knowledge, this study is the first to explore the views and experiences of researchers working in SSA countries, from a mix of different cultural backgrounds, focusing on behavioural science research and discussions of psychological constructs. These interviews raised challenges in terms of the interpretation and understanding of psychological constructs, and as behavioural science and health psychology are relatively new fields, these findings are imperative to ensuring that the continuing development of these fields consider the potential cultural differences in conceptualising theories and constructs. As the research base within these fields is predominantly developed from WEIRD populations, the overall suggestions by participants to consider testing and exploring the application of concepts and theories in SSA countries from the perspective of these countries is crucial and arguably methodologically and ethically vital.

## Conclusion

Although there are methodological issues with self-report measures, some of which are certainly not novel, the scale of the effect this can have on the validity and reliability of measures remains unknown. Undoubtably, without thorough consideration of the limitations highlighted in this study, future research may be more likely to have issues with validity and relevance. The findings from this research stress the need for consideration of the challenges discussed, and guidelines and/or recommendations for researchers assessing psychological constructs in SSA countries, particularly if the researchers intend to use concepts developed from Western based research or measurement tools.

# Chapter 3: Measurement of behavioural influences of healthcare workers towards antimicrobial use in sub-Saharan African countries: A systematic review

## 3.1 Introduction

Many measurement tools are adapted cross culturally, as a means of validating a measure for use in another context (Lopez-Roig & Pastor, 2016). Best practice guidelines for adaptation of measures outline the need to not only consider linguistic translation, but also the content, format, response scales and presentation to ensure the measure is appropriate for the culture (Beaton et al., 2000; Mohler et al., 2016). This includes reviewing the semantic and idiomatic equivalence, in terms of the meaning of words and any expressions and colloquialisms used. Developing culturally appropriate measures with established strategies for limiting method biases is one of the most important topics in behavioural science research, as stated in Chapter 2, and is crucial to assessing factors which influence healthcare workers behaviours (Muniz et al., 2013; Podsakoff et al., 2003). In order to explore factors which, influence behaviours, a common approach is the use of self-report questionnaires incorporating varying agreement question response options. This is discussed in depth in Chapter 2, and questionnaire development and validation are essential for research in behavioural science. Strategies for questionnaire development and evaluation are complex, requiring thorough consideration (Hansen et al., 2016; Boateng et al., 2018; Borghi et al., 2018; Lotfi et al., 2016; Nielsen et al., 2017). As most questionnaires measuring psychological constructs are developed on WEIRD populations, ensuring measures are adapted to new countries, cultures and languages is integral to validity and reliability (Beaton et al., 2000). However, researchers should consider whether adaptation is suitable, and some psychologists have argued that methods developed in WEIRD populations should be entirely reconsidered for appropriateness within Black, Indigenous and People of Colour (BIPOC) populations (Buchanan et al., 2021).

The literature and study outlined in Chapter 2 highlighted that Likert scales may be particularly problematic in certain populations, including SSA countries, yet Likert scales are the most widely used tool for assessing several constructs in behavioural science, including attitudes, beliefs, and preferences (Likert, 1932). Western dominated research practices, such as the use of Likert scales, are largely considered the first choice of measurement approach, see Chapter 1 for more details on this (Beins, 2019). As such, it is useful to understand and explore the type of measurement tools used in SSA countries to assess self-reported behavioural influences, to improve and build on non-Eurocentric behavioural science. This is currently highly relevant, as

the International Society for Behavioral Medicine, the Society of Behavioral Medicine and the European Health Psychology Society promote and advocate for the inclusion of behavioural science theories and approaches in global health initiatives, including the prevention of non-communicable diseases in LMICs (Dima et al., 2023; Odukoya et al., 2021).

A pertinent topic area in which self-report methods are popular is Antimicrobial Resistance (AMR), which is a worldwide public health concern, potentially threatening the lives of tens of millions of people (Braine, 2011). AMR, the process by which microbes develop a resistance to antimicrobials, is linked to behaviours relating to stewardship of antimicrobials, including inappropriate prescribing and overuse of medications (Jindal et al., 2015; Prestinaci et al., 2015). These behaviours are complex, multifaceted, and challenging to change, and the burden of AMR is significantly higher in SSA countries, due to the high prevalence of infectious diseases and reduced healthcare resources (Essack et al., 2017). The key to developing interventions to target these behaviours is understanding the problem and effectively identifying the specific factors that influence these behaviours (Calvo-Villamanan et al., 2022). However, the literature and knowledge base which informs intervention development is reliant on data from HICs (Wernli et al., 2020).

Behaviour change interventions are vital to combat AMR, but studies have shown that further research and intervention development is required in SSA countries (Iwu & Patrick, 2021). Assessing the effectiveness of interventions targeting antibiotic stewardship practices to measure and improve how antibiotics are prescribed and used in LMICs is problematic, due to poor study guality and vague intervention details (CDC, 2021; Van Dijck et al., 2018). In addition, studies do not always report behavioural influences, which are fundamental to stewardship interventions (Hulscher & Prins, 2017; Van Dijck et al., 2018). To improve the quality of intervention development to tackle AMR, it is necessary to determine which behavioural influences to assess and design measures to rigorously address the problem in each context (Rogers Van Katwyk et al., 2020). Knowledge, Attitude and Practice (KAP) surveys are a popular method used in health sciences research to gain information from a target population, with global health organisations advocating for this type of approach (Médicins du Monde, 2011). This health behaviour change model originated in the 1950s and assumes that knowledge and attitudes directly influence practice (Wan et al., 2016). Relatively cheap and easy to administer, this type of questionnaire can offer insights into behavioural influences to target interventions and provide useful baseline information (Andrade et al., 2020).

However, there have been many developments in understanding behaviour change over recent years, including the development of the Behaviour Change Wheel (BCW), established from a synthesis of 19 frameworks of behaviour change. Part of the BCW incorporates the COM-B model, developed to help identify potential factors influencing behaviour, stipulates that behaviour is influenced by Capability, Opportunity, and Motivation (Michie et al., 2011). This framework combines previous literature, the premise of which places varying degrees of importance on different influencing factors, at the expense of combining all important, potential drivers of behaviour (Michie et al., 2011).

In order to understand and intervene to reduce inappropriate antimicrobial prescribing and use measures to assess behavioural influences need to be relevant, comprehensive, valid, and reliable in the target population. The theories, models, and frameworks which the questionnaires are based on should also consider the unique challenges and barriers to AMR interventions, relating to health infrastructure, resources, access, and affordability (O'Neill, 2016). This review aims to identify what self-report measurement tools have been used to assess the behavioural influences of healthcare workers working in SSA countries towards antimicrobial use and determine how valid these measures are.

## Review question:

What self-report measurement tools have been used to assess the behavioural influences of healthcare workers working in SSA countries towards antimicrobial use?

Main objectives of this study:

- Identify and describe self-report measurement tools used to assess behavioural influences towards antimicrobial use in SSA countries among healthcare workers
- Assess the reported measurement properties of measures developed and validated for assessing factors which influence healthcare workers behaviour (including piloting, translating, adaptation, validity, and reliability)

## 3.2 Methods

This review was reported based on the Preferred Reporting Items for Systematic reviews and Meta-Analyses/PRISMA (Page et al., 2021). The protocol was registered with the Open Science Framework database in May 2023 (DOI: https://doi.org/10.17605/OSF.IO/ZSQF7).

## Inclusion criteria

Studies were selected that:

- (1) assessed self-reported behavioural influences towards antimicrobial prescribing behaviours or reported on the use of measures as a primary or secondary outcome which assessed behavioural influences towards antimicrobial use, including validation studies, new measurement tools or translated tools. Behavioural influences are those defined within the Theoretical Domains Framework (TDF; Atkins et al., 2017): social influences; environmental context and resources; social/professional role and identity; beliefs about capabilities; optimism; intentions; goals; beliefs about consequences; reinforcement; emotion; knowledge; cognitive and interpersonal skills; memory attention and decision processes; behavioural regulation and physical skills
- (2) included self-reported behavioural influences towards antimicrobial prescribing
- (3) were conducted in SSA countries, which is the region of Africa south of the Sahara. It consists of 46 out of the 54 countries in Africa including Angola, Benin, Botswana, Burkina Faso, Burundi, Cape Verde, Cameroon, Central African Republic, Chad, Comoros, Democratic Republic of the Congo, Republic of the Congo, Cote d'Ivoire, Equatorial Guinea, Eritrea, Eswatini (formerly Swaziland), Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, South Africa, South Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, and Zimbabwe (The World Bank, 2021)
- (4) included participants who were healthcare workers
- (5) used quantitative measurement tools (mixed methods included but only quantitative work reported in this study)
- (6) were published from 2001 onwards (WHO Global Strategy for AMR containment published this year; WHO, 2001)
- (7) were published in English

## Exclusion criteria

Studies were excluded if:

- (1) the work assessed antimicrobial prescribing in non-human populations
- (2) the study measured influences relating to infection, prevention, and control practices only
- (3) the measurement tool used was qualitative only

## Sources of data and search strategies

Relevant articles were identified in May 2023 through Medline, Scopus, African Index Medicus (AIM) PsycINFO and PsycARTICLES. All databases were searched from 2001 and limited to English language only. Search strategies for each database are outlined in Appendix 5. The reference lists of included studies and any systematic or scoping review studies were also screened against the inclusion criteria and relevant ones were included. Grey literature was sourced through searching Google Scholar and targeted websites, including NGO google custom search. Due to the limitations of Google Scholar, including but not limited to basic search capabilities, the search was restricted after 5 pages (100 results) and do not meet the inclusion criteria of this review following title review (Haddaway et al., 2015).

## Data extraction (selection and coding)

The titles and abstracts of studies were independently screened by the author, and a second reviewer (MSc in Public Health) independently checked 20% of the titles and abstracts at random to establish consistency. The full texts of eligible studies were then reviewed independently by the first reviewer, and the second reviewer independently checked 20% at random. At each stage, the studies were randomly selected with Microsoft excel using unique identifying numbers assigned to each database. Random numbers were then generated for each study in Excel. The dataset was then sorted in ascending order based on the generated random numbers. This ensured that the articles were randomised within their respective databases. The two lists of included articles following screening was then compared, and both reviewers met to discuss and resolve disagreements (of which there were only three). Articles within systematic reviews were assessed for duplicates at the full text stage, then their titles and abstracts were screened. Three studies which met the inclusion criteria had no information on the questionnaire items used in the research. Authors for all three studies were emailed and further details were requested to permit the inclusion of the article to this review. Only one author replied (27) with further information and this study was included in the review. The data collection form was designed based on the Cochrane data extraction form (Higgins & Green,

2011). Irrelevant items relating to trials were removed, and additional information relating to the measurement tools were included, such as reliability and validity assessments; piloting/pretesting; adaptation and response scales. This data was independently extracted by the first reviewer and all forms were checked by the second reviewer.

# Risk of bias (quality) assessment

Both reviewers discussed the quality assessment and independently assessed the quality of each study using the Appraisal tool for Cross-Sectional Studies (AXIS; Downes et al., 2016). This tool includes 20 items to appraise the reliability, relevance, and potential bias in the study, and each was assigned a score from 0 to 20, with higher scores demonstrating a lower risk of bias. The tool is outlined in Appendix 6. One study was a pre-post intervention design (15), and the NHLBI Quality Assessment Tool for Before-After (Pre-Post) Studies with No Control Group was used for this paper (NHLBI, 2013). This tool is outlined in Appendix 7.

# Ethical considerations

This review did not require ethical approval because it involved compiling together studies which had been previously published.

# 3.3 Results

The search yielded a total of 28 relevant articles from 13 SSA countries. Six studies were conducted in Ethiopia, four in Nigeria, four in South Africa, three in Zambia, two in Ghana, two in Uganda and one in Benin, Gambia, Sierra Leone, Cameron, Gabon, Lesotho and Tanzania. A visual representation of these countries in relation to the region of SSA is outlined in Figure 1. Most studies were published from 2020 onwards, 61%, and all studies were published after 2013. See Figure 2 for a Prisma flowchart (Page et al., 2020).

# Figure 1



Map of the SSA countries included in the review

*Note.* Image license: <u>https://commons.wikimedia.org/wiki/File:Africa-political-map.jpg</u>. Saudi Arabia, Yemen and Qatar are not part of SSA.

# Figure 2

PRISMA flow diagram



## Overview

Main study characteristics are provided in Table 4. Most studies were conducted in hospital settings (2,3,5,6,7,9,11-20,22,26) and the rest were in primary care or a combination of private and public healthcare facilities, with two national studies (4,13) and one study completed in a community setting (21). The healthcare workers included in the studies were predominantly pharmacists, nurses and doctors, and the number of healthcare workers included in each study varied and ranged from 51-2523 (Mean = 236.71). All studies used questionnaires to collect the relevant data on behavioural influences, with three studies employing interviewers to administer the questionnaires (2,7,24), two studies did not state modality (17,27) and the rest were selfadministered (Table 7). All studies were published in journals and no grey literature which met the inclusion and exclusion criteria was found. All eligible data was obtained from crosssectional research studies, apart from one study undertaken in Ghana as a pre and posttest intervention evaluation tool (24). The aim of this study was to evaluate the success of a training programme, and this differed from all other studies' objectives; to explore behavioural influences, with 16 studies stating that informing intervention development and/or implementation of AMR policies was either an aim of the study or highlighted as a potential application for the results of the study (4,5,8,9,11,13,14,16-18,20,22,24,26-28).

#### Quality assessment

An overview of the quality assessment information is outlined in Table 2 and Table 3. Only one study based the measurement tool on a previous scale which had been validated in a former study (21). Of the 28 studies, 21 scored 16 or higher out of 20 possible items, indicating high quality (1-5,7-9,11-14,16-18,20-22,24-26). Of the studies that missed items, most did not take measures to report, categorise or describe non-responders. Three studies scored 14 (6,10,19) and one study scored 11 (23), largely due to issues with reporting methods. This indicated moderate quality, but the remaining two studies were of low quality, with a score of eight and nine (27,28). Five studies did not report a pretest or pilot of the questionnaire prior to use (1,10,14,15,27). The pre-posttest study design received a score of 5 out of a possible 11, indicating moderate quality.

#### Measurement development

A total of seven studies did not report how the measurement tool was developed (15,16,19,22,23,27,28). Of the 21 studies that did provide details on how the measure was developed, ten used at least one previous study based in a SSA population (1,4,6-8,12-

14,20,25), one used WHO AMS toolkits developed with information specific to SSA countries (26) and ten studies did not report using studies undertaken in SSA countries to develop the measure (2,3,5,9,10,11,17,18,21,24; Table 6).

Twenty-three studies reported conducting some form of pre-testing or pilot testing, illustrated with other measurement considerations in Table 7. However, 30% of these studies provided no details on what the piloting involved or how it was conducted (4,5,12,18,19,22,23). Eight studies only outlined the population involved (3,6,7,9,11,13,16,20) and six studies stated whether the results from pilot/pretest were used in the analysis (2,8,24-26,28), with only one study incorporating this information in the results (28). In relation to changes based on pilot/pretest data, five studies stated that changes were made to the questionnaire based on this information (8,17,21,25,26). None of the included studies discussed piloting or adaptation specifically in relation to question stems or response scales. Although several studies used parts of questionnaires which had already been published or information from the literature (1,2,4-6,8-10,12,25,26), only one study detailed which questions from the studies were used and specifically how they had been adapted (20).

#### Response scales

Twenty-five of the 28 studies (89%) incorporated Likert scales as the response scales for some or all behavioural influences assessed (1,3-26 Table 4). Five studies used a dichotomous scale (2,4,25,27,28), one study used dichotomous with an option for "I do not know" (13), seven studies developed multiple choice questions to assess knowledge (3,4,13,16,20,26,27) and one study used a continuous scale from one to ten (9). Most Likert scales had five points, with one study using a three-point scale (24). and one other using a four-point scale (22). Agreement Likert scales (strongly agree, agree, neither agree nor disagree, disagree, strongly disagree) were the dominant label descriptors used to assess attitudes, perceptions, and beliefs, incorporated in questionnaires of 20 studies (1,3-5,7-10,12,14-19,21-24,26). Three studies developed a usefulness Likert scale (13,16,22) and a further 3 studies used a degree of influence/contribution Likert scale (3,11,14). Two studies used an importance Likert scale (3,20) or a confidence Likert scale (4,20). One study had a response scale indicating how often, using a Likert scale, and 1 study did not report what labels were used in the Likert scales incorporated in the questionnaire (6). All studies which reported Likert scale information (24 of the 25 studies) used verbal descriptors in the questionnaires for each point of the Likert response scale.

## Validity and reliability

Validity and reliability were reported for all included studies and assessed using standard definitions outlined in the literature (Dronavalli & Thompson, 2015; Taherdoost, 2016). Studies were deemed to have assessed content validity if a team of experts or judges had been used to assess specifically the relevance of the tool and whether the questionnaire is representation of the target construct (Taherdoost, 2016). Only four studies met this criterion for reporting content validity assessment (2,20,21,26). One study used a previously developed measure, of which the content validity had been assessed in the original development study and in the modified data collection tool by a team of experts (21). This study also assessed face validity, through determining where items linguistically and analytically represented the target construct (7,21,25,26). Six studies calculated the internal consistency of the questionnaire using Cronbach's alpha (1,5,7,13,21,26; Table 6).

## Translation

Only seven studies (25%) outlined the language(s) in which the questionnaire was delivered to healthcare workers, three of which were written in English (7,19,20), two were in Amharic, the official language of Ethiopia (18,21) and one questionnaire was in Swahili, an official language of Zambia (5). The questionnaire in one study conducted in Cameroon was produced for participants in both English and French (14). The author of one further study provided information on the language of the questionnaire after being contacted, and this questionnaire was written in French, an official language of Benin (20). Two studies highlighted that the questionnaires were prepared in English, translated into the target language, pre-tested and then back translated into English (5,21). The other studies did not report any information on translation practices.

### Authorship

Fifteen of the twenty-eight articles (54%) had contributing authors based or affiliated with a university or health facility in a Western country (2,3,5-9,12-15,18,20,25,28). Of these 15 articles, three studies did not include an author from the target population country (12,14,20) and two of these studies only included authors from European countries, Hungary, and Germany respectively (12,20). All other studies included authors based in the target population country (Table 5). Only three studies did not include an author from the target population country as a first or last author in the published paper (8,12,15).

## Behavioural influences

Findings demonstrate that knowledge, attitudes, and perceptions were the most reported behavioural influences assessed in the literature (Table 4). Twenty-one studies measured knowledge, 16 studies explored attitudes, 11 looked at perceptions and two assessed beliefs. The terminology for behavioural influences used in this study is the same as the authors originally proposed to identify the constructs measured. Most studies (27 out of 28) did not clearly describe the construct(s) they were assessing, behavioural influences, or the rationale, theory, or framework for the chosen construct(s), one study did (22). Seven studies referred to the KAP acronym, but no further information was reported on the related theory (2,4,5,12,14,20,26). In one study, the authors specifically discussed the construct of attitudes and the theoretical stance for the definition they used (5).

## 3.4 Discussion

This systematic review provides a summary of self-report measurement tools used in SSA countries to assess behavioural influences towards antimicrobial use in healthcare workers. Self-administered questionnaires, incorporating Likert scale response options to assess healthcare workers' knowledge, attitudes and perceptions toward AMR dominated the literature base. Overall, most studies were rated as high quality for cross-sectional studies and included a pilot or pre-testing phase, although the quality of reporting on these varied and very few studies conducted validity or reliability assessments. Seventeen studies either did not report how the measurement tool used was developed or did not disclose the use of studies undertaken in SSA countries to support the questionnaire development. All studies bar 3 used Likert response scales, however none of the studies stated that the Likert response scales, or any of the scales, were considered in the pilot/pre-testing phase. A high proportion of studies had authors based of affiliated in a Western country, and this proportion is similar to the findings of research studies discussed in Chapter 1 (Arnett, 2009; Thalmayer et al., 2020).

Most of the studies discovered in the literature base were conducted in hospital settings, in one or more specific sites within small geographical regions. As a potential result of this, all studies bar one developed a new measurement tool for the specific purposes of that research project. This is an interesting finding, as the topic area of AMR is currently very relevant, and there should be a greater focus on utilising existing measures as opposed to developing new measures exploring the same psychological constructs using different means (WHO, 2015). By using existing measures, either through adapting Western developed measures to ensure cultural relevance, or utilsing measures already developed in the culture of question, this could help eliminate potential concerns relating to biases in development processes associated with culture.

Although most studies reported conducting some form of pretesting or piloting of the questionnaire, and the quality of included studies was predominantly moderate-high, the reporting of the specific process adopted was omitted from most articles. As a result, the validity and reliability of most studies is questionable, particularly as all but one questionnaire was newly developed. Furthermore, authors referenced a large body of literature from Western countries as informing the development of questionnaires. This is understandable, due to the Western bias in publications and large body of research from high-income countries (Mulimani, 2019; also discussed in Chapter 1). However, this creates challenges in ensuring conceptual,

experiential, semantic and idiomatic differences across cultures are comprehensively considered in measurement design (Beaton et al., 2000; Lopez-Roig & Pastor, 2016; Hruschka et al., 2018). It is difficult to determine the impact of Western measurement approaches in the included papers, as no studies reported directly adapting a questionnaire developed in a Western country. This is likely due to the objective of most studies to report on local data at one hospital or several hospitals within a geographical region of a country. More than half of studies did include an author based or partially based in an institution in a Western country, with some studies published only by authors based in Western educational institutions. However, it should be highlighted that not all Western influences are potentially a drawback, but this should be balanced with considerations to the local context to ensure the tools developed are relevant to the culture in question (Hruschka, 2020).

It is surprising that very few studies stated the language of questionnaire delivery, especially considering English is not an official language in Tanzania, Ethiopia, and Gabon. This raises key questions regarding construct equivalence, as it is possible that a number of these measures were developed in one language and translated into official and/or local languages, and that translation practices include careful consideration of the intention and specific meanings of question items to avoid issues with interpretation (Hawkins et al., 2020). The few studies that did report translation highlighted this process and discussed back translation; once deemed a valuable part of the translation process but more recently criticisms of this approach have been raised, including the potential to miss important translation issues (Epstein et al., 2015; Behr, 2016). Research suggests that a panel or committee of experts to translate in a team approach is a more successful method (Epstein et al., 2015).

Considering the Western influences in the measurement development, highlighted above, key questions remain regarding the relevance of the response scales. With so many studies utilising Likert scales, particularly agreement scales, the question of reliability is an important one. None of the studies included reported pretesting or piloting the response scales with participants. Although this is not regular practice in cross sectional research studies, and there were no perceived issues reported in the papers in relation to the response scales, for the reasons outlined in Chapter 1 & 2, this should be considered further in populations that are not represented in WEIRD samples. Over and above measurement equivalence, which in this context is not so relevant due to lack of cross-cultural comparisons and no reporting of adaptation of measures, there is little evidence or understanding in relation to when Likert

scales will provide tangible, relevant data and when this may not be the case. Extreme responses and missing data have been reported in the literature, but this is a considerably complex and under researched area, and further work is required to explore this issue (Hruschka et al., 2018). Studies in general do not tend to highlight issues with response scales, unless this has been raised by several participants or probed further by interviewers, as seen in the findings outlined in Chapter 2.4, and potential issues can go relatively unnoticed in research studies. Therefore, like the findings outlined in Chapter 2.4, there is a need to thoroughly pilot/pre-test and utilise local expertise in the development and testing of questionnaires to ensure the measure is culturally appropriate.

Most studies were published in the last few years, but the KAP model of assessing influencing factors appeared to be the basis for most work, as authors targeted knowledge and attitudes (Andrade et al., 2020). Although very few studies referenced this model directly to support the choice of behavioural influences measured in the research, no other theories, models, or frameworks were cited as a rationale for the chosen influencing factors. The lack of reference to theories is not entirely surprising, as studies were based on practice and not in academic settings. Nevertheless, the use of behavioural approaches to determine which influencing factors to assess is important, especially considering that a large proportion of studies intended to potentially use the results to develop interventions and/or inform AMS action plans (Rogers Van Katwyk et al., 2020).

Caution must be warranted, as behavioural science approaches must consider the context in SSA countries, which often face unique resource and health constraints. Arguably, this further supports the need for incorporating comprehensive frameworks to consider a wider range of factors which may influence healthcare practitioners' behaviour, particularly in AMS intervention development. Behavioural science approaches have progressed in recent years, with a focus on developing frameworks which include potential influencing factors in a digestible form, and link to relevant theories (Michie et al., 2014). But reviews of AMS interventions in LMICs also found a lack of consideration for behavioural science theories in intervention development, despite recent evidence in support of this (Van Dijck et al., 2018; Craig et al., 2023). Further work is needed to understand, promote, and incorporate behavioural science theories and approaches in SSA countries.

#### Limitations

Whilst this review makes a unique contribution to the literature on measurement development in SSA countries within the field of behavioural science, it is important to view the findings in light of the study limitations. This review is limited due to a lack of information and reporting on measurement properties of developed questionnaires. Including English language studies only underestimates the contributions from non-English speaking countries and is complicit in publication bias towards Western countries. Although a google scholar search was used to identify grey literature, future attempts should include a more extensive and thorough grey literature search to minimize the effects of publication bias, including incorporating unpublished data through contacting researchers directly.

The interpretation of the results was also largely undertaken in a Western context, by a researcher based in the West. The topic area and population chosen for assessing self-report measurement resulted in participants from educated backgrounds. Further research is required to explore and understand the measurement tools and associated properties in community populations, as there are likely to be additional challenges in terms measurement equivalence and the use if Likert scales (Hruschka et al., 2018). This study focused on SSA countries, but it should be noted that only 14 out of 46 SSA countries were represented in this review. As discussed in Chapter 1, using the term SSA broadly is not representative of the diverse mix of included countries, and this should be considered when interpreting these findings. This also limits the generalizability of the results, as this study focused on research in a selective and relatively new area of interest. Further studies could explore behavioral influences across multiple healthcare worker behaviors to gain deeper insight into self-report measurement tools in SSA countries.

## Conclusion

Overall, this review found that there are several existing self-report measures available in countries across SSA to assess behavioural influences of healthcare workers towards antimicrobial use in humans. Measures are typically developed for the primary purpose of the study, and authors could potentially benefit from exploring the current literature base to support the development of measures, rather than using Western based studies to inform development, which could potentially be biased due to cultural differences. In general, studies did report pilot or pre-testing prior to rolling out the developed measure. However, this did not consider Likert response scales which could be a potential challenge for certain populations and should be

reviewed prior to implementation. Validity, reliability, translation and the use of behavioural theories and models to support the development of questionnaires are key areas which were under reported and should be considered further in measurement development studies to support the assessment of the cultural relevance of the measure.

	Results for AXIS appraisal																				
Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Author																					
1.Chetty et	1	1	0	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	16
al., 2023																					
2.Tembo et	1	1	1	1	1	1	0	1	1	1	0	1	0	0	1	0	1	1	1	1	15
al., 2022																					
3.Gulleen	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	15
et al., 2022																					
4.Balliram	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	16
et al., 2021																					
4.Mufwam	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	17
bi et al.,																					
2021																					
6.Nkinda et	1	1	0	1	1	1	0	0	1	1	1	0	0	0	1	1	1	1	1	1	14
al., 2022																					
7.Kimbowa	1	1	1	1	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	16
et al., 2021																					
8.Kabba et	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	18
al., 2020																					
9.Kalungia	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	17
et al., 2019																					

# Table 2Results for quality assessment appraisal using AXIS

	Results for AXIS appraisal   1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Total																				
Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Author																					
10.Farley	1	1	0	1	1	1	0	1	0	0	1	1	0	0	1	1	1	1	1	1	14
et al., 2018																					
11.Adorka	1	1	0	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	15
et al, 2013																					
12.Abubak	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	17
ar et al.,																					
2023																					
13.Adegbit	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	17
e et al.,																					
2022																					
14.Domche	1	1	0	1	1	1	0	1	0	1	1	1	0	0	1	1	1	1	1	1	15
Ngongang																					
et al., 2021																					
16.Gebrehi	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	16
wotet al.,																					
2020																					
17.Mersha,	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	18
2018																					
18.Gebrete	1	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	17
kle et al.,																					
2018																					

	Results for AXIS appraisal																				
Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Author																					
19.Tafa et	1	1	0	1	1	0	0	1	1	1	0	1	0	0	1	1	1	1	1	1	14
al., 2017																					
20.Chaw et	1	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	16
al., 2017																					
21.Erku,	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	18
2016																					
22.Abera	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	19
et al., 2014																					
23.Ajibade	1	1	0	1	1	0	0	1	0	1	0	0	0	0	1	1	1	1	0	1	11
et al., 2014																					
24.Asante	1	1	1	1	1	1	0	1	1	0	1	1	0	0	1	1	1	0	1	1	15
et al., 2017																					
25.Guma	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	17
et al., 2022																					
26.Ogoina	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	17
et al., 2021																					
27.Dougno	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	1	1	8
n et al.,																					
2020																					
28.Manga	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	9
et al., 2021																					

# Table 3

Quality assessment using NHLBI

Results for NHLBI appraisal													
Study Author	1	2	3	4	5	6	7	8	9	10	11		
15.Sneddon et al., 2020	1	1	1	1	0	CD <sup>a</sup>	0	NR⁵	NR⁵	1	1		

<sup>a</sup>Cannot determine

<sup>b</sup>Not reported

# Table 4

Characteristics of included studies

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
1.Chetty et	South	Public &	N = 55	Knowledge	LS: Strongly agree;	11 Questions	2 participants
al., 2023	Africa	private		Attitudes	agree; neutral;	assessing	missing data
		healthcare		Perceptions	disagree; strongly	knowledge,	
		sectors			disagree	attitudes, and	
		Pharmacists				perceptions	
						towards AMS	
						practices	
2.Tembo et	Zambia	Tertiary	N = 263	Knowledge	Dichotomous: Yes or	6 Questions	No missing
al., 2022		hospital		Attitudes	no	on knowledge	data
		Nurses &				about AMR 6	
		pharmacy				Questions on	
		workers				attitudes	
						towards AMR	

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
3.Gulleen et	Uganda	UCI Hospital	N = 61	Knowledge	Multiple choice	10 Questions	1 participant
al., 2022		staff		Perceptions	(knowledge) & LS:	on influencing	missing all
		Nurses,			Does not contribute;	factors	data
		pharmacists			usually does not	towards AMR	
		& physicians			contribute; neutral;	9 Questions	
					occasionally	on influencing	
					contributes; frequently	factors for	
					contributes	prescribing	
					LS: Strongly disagree;	3 Questions	
					somewhat disagree;	on knowledge	
					neither agree nor	questions	
					disagree; somewhat		
					agree; strongly agree;		
					don't know		
					LS: Very unimportant;		
					somewhat		
					unimportant; neutral;		
					somewhat important;		
					very important		

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
4.Balliram et	South	National	N = 2523	Knowledge	Dichotomous: Yes or	8 Questions	Only 2098
al., 2021	Africa	level		Attitudes	no	on knowledge	completed the
		Pharmacists,			Multiple choice	13 Questions	influencing
		Doctors &			questions (attitudes)	on confidence	factors
		Nurses			LS (knowledge &	7 Questions	sections (425
					attitudes): Strongly	on attitudes	missing)
					agree; agree; neutral;		
					disagree; strongly		
					disagree		
					LS (confidence): Very		
					confident; confident;		
					unsure; unconfident;		
					very unconfident; not		
					applicable		

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
5.Mufwambi	Zambia	Tertiary	N = 304	Knowledge	LS: Strongly agree;	11 Questions	22 missing
et al., 2021		Hospitals		Attitudes	agree; not sure;	on knowledge	
		Healthcare			disagree; strongly	of AMR	
		professionals			disagree	10 Questions	
						on attitudes	
						towards AMR	
						8 Questions	
						on knowledge	
						of prescribing	
6.Nkinda et	Tanzania	Hospital	N = 108	Knowledge	Two different 5-point	Knowledge	Not reported
al., 2022		Prescribers		Attitudes	LS (not reported)	on rational AB	
		& dispensers				use in	
						children &	
						how AR	
						occurs	
						Attitudes	
						towards AB	
						use in	
						children & AR	
						(no further	
						info)	

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
7.Kimbowa et	Uganda	Hospital	N = 582	Attitudes	LS: Strongly agree;	12 Questions	Missing 2
al., 2021		Healthcare			agree; neither agree	assessing	
		providers			nor disagree;	attitudes	
					disagree; strongly	relating to	
					disagree	AMS	
8.Kabba et	Sierra	Public sector	N = 119	Knowledge	LS: Strongly agree;	22 Questions	
al., 2020	Leone	hospitals		Attitudes	somewhat agree;	on knowledge	
		Doctors			neither agree nor	of AB and	
					disagree; somewhat	ABR	
					disagree; strongly	8 Questions	
					disagree	on attitudes	
						towards	
						prescribing	
						AB	
9.Kalungia et	Zambia	University	N = 198	Knowledge	Continuous scale from	3 Questions	No missing
al., 2019		teaching		Perceptions	1 of 10 (knowledge)	on knowledge	data
		hospitals			LS: Agreed,	on AMS	
		Physicians &			disagreed, unsure	8 Questions	
		pharmacists			(perceptions &	on	
					attitudes)	perceptions	
						and attitudes	
						about AMR	

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
10.Farley et	South	All GPs in	N = 264	Knowledge	LS: Agree, disagree;	12 Questions	Between 11-
al., 2018	Africa	South Africa		Attitudes	unsure (attitudes &	on attitudes	22 for
		Primary care		Perceptions	perceptions)	and	questions
		prescribers			Multiple choice	perceptions	
					questions (knowledge)	relating to AB	
					Case studies with	7 Knowledge	
					multiple choice	case studies	
					options (knowledge)	with multiple	
						choice	
11.Adorka et	Lesotho	Hospital	N = 51	Attitudes	LS of degree of	6 Questions	No response
al, 2013		Doctors &		Perceptions	influence: No	on influencing	between 5-0
		primary			response; not at all;	factors of	
		healthcare			minor degree; major	prescribing	
		nurses			degree	AB	

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
12.Abubakar	Nigeria	Hospitals	N = 313	Knowledge	LS: Strongly agree;	20 Questions	Not reported
et al., 2023		Health		Attitudes	agree; neutral,	on knowledge	
		workers			disagree; strongly	of AB and	
					disagree	ABR	
						15 Positive &	
						negative	
						questions	
						evaluating	
						attitude	
						towards	
						prescription of	
						ABs	
13.Adegbite	Gabon	Hospital	N = 47	Knowledge	Dichotomous: Yes; no;	8 Questions	Not reported
et al., 2022		Healthcare		Perceptions	l do not know	on knowledge	
		workers			(knowledge &	of ABR	
					attitudes)	9 Questions	
					LS: Very useful;	on attitudes	
					useful; neither useful	12 Questions	
					nor unnecessary;	on attitudes	
					useless; completely	towards	
					unnecessary	interventions	
					(attitudes)		

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
14.Domche	Cameroon	Tertiary	N = 98	Knowledge	LS: Strongly agree;	12 statements	Factors that
Ngongang et		hospitals		Perceptions	agree; neutral;	on	influence (25
al., 2021		Medical		Other factors	disagree; strongly	perceptions of	missing)
		doctors		(single items)	disagree (perceptions	AB use and	
					& attitudes)	ABR	
					LS: Strongly influence,	8 statements	
					influence, neutral, do	on	
					not influence, do not	perceptions of	
					influence at all	AB use and	
					Knowledge scale	ABR	
					unclear, scored as	Knowledge	
					correct or incorrect	questions	
						(various)	
15.Sneddon	Ghana	Hospitals	N = 60	Knowledge	Multiple choice	8 Questions	1 pre-test
et al., 2020		Healthcare		Attitudes	(knowledge)	on knowledge	missing
		professionals			LS: Strongly agree;	of AMS 10	responses
					agree; neither agree	Questions on	3 post-test
					nor disagree;	attitudes	missing
					disagree; strongly	towards AMR	responses
					disagree (attitudes)	& AMS	

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
16.Gebrehiwo	Ethiopia	University	N = 153	Knowledge	LS: Very useful;	7 Questions	Not reported
t et al., 2020		Hospital		Beliefs	useful; not useful; not	on beliefs	
		Healthcare			sure(beliefs)	about AMR	
		providers			LS: Strongly agree;	interventions	
					agree; disagree;	3 Questions	
					strongly disagree; do	on beliefs	
					not know (beliefs)	about the	
					Multiple choice	problem of	
					questions (beliefs)	AMR	
					Knowledge scale: not	7 Questions	
					reported	on knowledge	
						of cause of	
						AMR	
						2 Questions	
						on beliefs	
						about	
						unnecessary	
						AB	
						prescriptions	

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
17.Mersha,	Ethiopia	University	N = 270	Attitudes	LS: Strongly agree;	13 Questions	Not reported
2018		Teaching			agree; neutral;	on attitudes	
		Hospitals			disagree; strongly	towards AMS	
		Medical			disagree		
		Interns					
18.Gebretekle	Ethiopia	Hospital	N = 406	Perceptions	LS: Strongly agree;	19 Questions	No missing
et al., 2018		Pharmacists		(other data	agree; disagree;	on	data
		& Physicians		qualitative)	strongly disagree	perceptions of	
						AMR	
19.Tafa et al.,	Ethiopia	Hospitals &	N = 218	Knowledge	LS: Strongly agree;	14 Questions	No missing
2017		health		Attitudes	agree; disagree;	on AMR	data
		centers			strongly disagree	knowledge	
		Paramedical			(knowledge)	9 Questions	
		staff			Likert scale: Above &	on attitudes	
					don't know (attitudes)	towards AMR	
						4 attitudes on	
						unnecessary	
						prescriptions	
						3 attitudes on	
						importance of	
						AMR burden	
Study ID	Country	Setting	Sample size	Behavioural influences	Scale	Questions	Reporting of missing data
-------------------------	----------	---	-------------	--------------------------------------	---	--	--
20.Chaw et al., 2017	Gambia	Hospitals & health centers Health practitioners	N = 216	influences Knowledge Attitudes	Multiple choice (knowledge) LS: Very confident; confident; unconfident; very unconfident LS: Very important; important; moderately important; slightly important; not important; not	10 Questions on knowledge of inappropriate AB use 2 Questions on attitudes towards AB prescribing & important	missing data 14 questionnaires had some missing data
21.Erku, 2016	Ethiopia	Pharmacies Community pharmacists	N = 389	Perceptions	applicable (attitudes) LS: Strongly agree; agree; neutral; disagree; strongly disagree	factors influencing prescribing 8 Questions on perceptions towards AMS	No missing data

Study ID	Country	Setting	Sample size	Behavioural influences	Scale	Questions	Reporting of missing data
22.Abera et al., 2014	Ethiopia	Hospitals Physicians & nurses	N = 385	Knowledge Beliefs	LS: Strongly agree; agree; disagree; strongly disagree; don't know (knowledge) LS: Very useful, useful, not useful, not sure (beliefs) Additional knowledge questions scale not reported	13 Questions on knowledge of AMR 2 Questions on knowledge on scope of AMR problem 14 Questions on beliefs regarding AMR	No missing data
23.Ajibade et al., 2014	Nigeria	National level Nurse- Anesthetists	N = 67	Knowledge Attitudes	LS: Strongly agree; agree; unsure; disagree; strongly disagree	2 Questions on knowledge and 2 attitude questions on perioperative antibiotic prophylaxis prescribing	Varies between 7 and 0 for the 4 questions

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
24.Asante et	Ghana	Public &	N = 379	Knowledge	LS: Agree; disagree;	13 Questions	No missing
al., 2017		Private			don't know	on knowledge	data
		facilities				of ABR & use	
		Prescribers					
25.Guma et	South	Primary	N = 209	Knowledge	Dichotomous: Yes; no	10 Questions	34 incomplete
al., 2022	Africa	healthcare		Factors in	& multiple choice	on knowledge	questionnaires
		facilities		general	questions (knowledge)	of AB use	excluded
		General		Attitudes &	LS: Very often; often;	16 Questions	
		practitioners		perceptions	about half the time;	assessing	
				mentioned	rarely; almost never	environmental	
				as being in	(environmental	factors	
				questionnair	factors)	influencing	
				e, but not		AB	
				outlined in		prescribing	
				results			
26.Ogoina et	Nigeria	Tertiary	N = 1324	Knowledge	Multiple choice	28 Questions	Incomplete
al., 2021		hospitals		Attitudes	(knowledge)	on knowledge	questionnaires
		Physicians			LS: Strongly agree;	of AB & AMR	were counted
					agree; neutral;	15 Positive &	as non-
					disagree; strongly	negative	response and
					disagree (attitudes)	attitude	not included in
						questions	final numbers

Study ID	Country	Setting	Sample size	Behavioural	Scale	Questions	Reporting of
				influences			missing data
27.Dougnon	Benin	Healthcare	N = 330	Knowledge	Dichotomous: Yes; no	18 Questions	Not reported
et al., 2020		workers		Attitudes	(attitudes &	on knowledge	
		(nurses,			knowledge)	regarding AB	
		assistant			Multiple choice	prescribing	
		nurses,			questions (knowledge)	Number of	
		physicians,				attitude	
		midwives,				questions not	
		specialized				reported	
		physicians)					
28.Manga et	Nigeria	Primary	N = 442	Perceptions	Dichotomous: Yes: no	5 Questions	Not reported
al., 2021	gena	healthcare			,,	assessing	
, -		workers				perceptions	
						for not	
						adhering to	
						AB	
						prescribing	
						guidelines	

Publication information

Study ID	Countries where authors were	Lead author/final author	Journal published
	based at time of publication		
1. Chetty et al., 2023	South Africa	South Africa	Health SA Gesondheid - Journal of
			Interdisciplinary Health Sciences
2. Tembo et al., 2022	Zambia; UK	Zambia & UK	JAC-Antimicrobial Resistance
3. Gulleen et al., 2022	USA; Uganda	USA & Uganda	Antimicrobial Stewardship &
			Healthcare Epidemiology
4. Balliram et a., 2021	South Africa	South Africa	Southern African Journal of
			Infectious Diseases
5. Mufwambi et al., 2021	Germany, Sweden; UK; Zimbabwe; Zambia; South Africa	Zambia & UK	Frontiers in Pharmacology
6. Nkinda et al., 2022	Tanzania; Switzerland; USA; Rwanda	Tanzania & Rwanda	BMC Health Services Research
7. Kiba et al., 2021	Uganda; Sweden	Uganda	PLoS ONE
8. Kabba et al., 2020	China; Sierra Leone; Australia	China	Transactions of the Royal Society of Tropical Medicine and Hygiene

Study ID	Countries where authors were	Lead author/final author	Journal published
	based at time of publication		
9. Kalungia et al., 2019	Zambia; UK; Sweden; South Africa	Zambia & South	Journal of Chemotherapy
		Africa/UK/Sweden	
10. Farley et al., 2018	South Africa	South Africa	South African Medical Journal
11. Adorka et al, 2013	Nambia; Lesotho; Mexico	Nambia & Mexico	African Health Sciences
12. Abubakar et al., 2023	Hungary	Hungary	Journal of Infection Prevention
13. Adegbite et al., 2022	Gabon; Netherlands; Germany; Sierra Leone; South Africa	Gabon/Netherlands & Gabon/Netherlands/Germ any/Sierra Leone/South Africa	Antimicrobial Resistance & Infection Control
14. Domche Ngongang et al., 2021	South Africa; USA	South Africa/USA & South Africa	BMC Infectious Diseases
15. Sneddon et al., 2020	Ghana; Iraq; UK	UK & UK	JAC Antimicrobial Resistance
16. Gebrehiwot et al., 2020	Ethiopia	Ethiopia	Infection and Drug Resistance
17. Mersha, 2018	Ethiopia	Ethiopia	Antimicrobial Resistance and
			Infection Control
18. Gebretekle et al., 2018	Ethiopia; Canada	Ethiopia & Canada	PLoS ONE

Study ID	Countries where authors were	Lead author/final author	Journal published
	based at time of publication		
19. Taga et al., 2017	Ethiopia	Ethiopia	Annals of Clinical Microbiology
			and Antimicrobials
20. Chaw et al., 2017	Germany	Germany	Transactions of the Royal Society
			of Tropical Medicine and Hygiene
21. Erku, 2016	Ethiopia	Ethiopia	Interdisciplinary Perspectives on
			Infectious Diseases
22. Abera et al., 2014	Ethiopia	Ethiopia	BMC Pharmacology and
			Toxicology
23. Ajibade et al., 2014	Nigeria	Nigeria	Nigerian Journal of Medicine
24. Asante et al., 2017	Ghana	Ghana	BMC Health Services Research
25. Guma et al., 2022	South Africa; UAE; UK	South Africa & South	Antibiotics
		Africa	
26. Ogoina et al., 2021	Nigeria	Nigeria	Antimicrobial Resistance and
			Infection Control
27. Dougnon et al., 2020	Benin	Benin	International Journal of One
			Health
28. Manga et al., 2021	Nigeria; South Africa; UK	Nigeria & Nigeria	PAMJ One Health

# Questionnaire adaptation information

Study ID	Adaptation
1. Chetty et al., 2023	Incorporated questions from previous instrument designs in 4 similar studies:
	South Africa, questionnaire with student population
	<ul> <li>Pakistan, qualitative and quantitative study with students</li> </ul>
	Qatar, qualitative study, unclear
	Australia & France, survey with pharmacists
	No further details on adaptation process reported
2. Tembo et al., 2022	Based on 1 previous study:
	<ul> <li>Iraq, questionnaire with undergraduate pharmacy students</li> </ul>
	No further details on adaptation process reported
3. Gulleen et al., 2022	Based on 1 previous study:
	Peru, questionnaire with medical doctors
	No further details on adaptation process reported
4. Balliram et al., 2021	Adapted and piloted, based on questionnaires in 4 previous studies:
	France & Scotland, survey with junior doctors
	Ethiopia, questionnaire with physicians & nurses
	Brazil, survey with nurses & doctors & India, survey with paramedical staff
	No further details on adaptation process reported

Study ID	Adaptation
5. Mufwambi et al., 2021	Based on 2 previous studies:
	Europe, secondary data from prescribers
	<ul> <li>Southern Iran, pre/post intervention assessing antibiotic use</li> </ul>
	No further details on adaptation process reported
6. Nkinda et al., 2022	Based on 2 previous studies:
	DR Congo, survey with medical doctors & students
	Pakistan, survey with pharmacists
	No further details on adaptation process reported
7. Kimbowa et al., 2021	Based on 2 previous studies:
	Ethiopia, questionnaire with healthcare professionals
	Pakistan, questionnaire with pharmacists
	No further details on adaptation process reported
8. Kabba et al., 2020	Based on 5 previous studies:
	<ul> <li>Laos (Southeast Asia), questionnaire with doctors</li> </ul>
	Greece and Cyprus, survey with physicians
	<ul> <li>Portugal, questionnaire development study with physicians</li> </ul>
	Ethiopia, survey with paramedical staff
	China, questionnaire with students
	No further details on adaptation process reported

Study ID	Adaptation
9. Kalungia et al., 2019	Based on 2 previous studies:
	US survey with nurse practitioners
	US survey with medical students
	No further details on adaptation process reported
10. Farley et al., 2018	Based on 1 previous study:
	Systematic review on clinicians
	No further details on adaptation process reported
11. Adorka et al, 2013	Based on book guidelines:
	Book on practices for survey research
	No further details on adaptation process reported
12. Abubakar et al., 2023	Based on 3 previous studies:
	Sierra Leona, survey with medical doctors
	Nigeria, survey with healthcare workers
	Nigeria, survey with physicians
	No further details on adaptation process reported

Study ID	Adaptation				
13. Adegbite et al., 2022	Based on 2 previous studies:				
	France & Scotland, survey with junior doctors				
	Ghana, survey with prescribers				
	No further details on adaptation process reported				
14. Domche Ngongang et al., 2021	Adopted questions from a previous study:				
	South Africa, survey with medical students				
	No further details on adaptation process reported				
15. Sneddon et al, 2020	Not reported				
16. Gebrehiwot et al., 2020	Not reported				
17. Mersha, 2018	Based on a review of previous literature:				
	US survey with medical students				
	US survey with pharmacy students				
	South India, survey with medical students				
	Sweden, survey with the public				
	No further details on adaptation process reported				
18. Gebretekle et al., 2018	Adapted from a previous study:				
	<ul> <li>AMS toolkit best practices (includes pre/post assessment questionnaire)</li> </ul>				
	No further details on adaptation process reported				

Study ID	Adaptation
19. Tafa et al., 2017	Not reported
20. Chaw et al., 2017	Based on 3 previous studies:
	France & Scotland, survey with junior doctors
	Saudi Arabia, questionnaire with physicians
	DR Congo, survey with medical doctors & students
	Clearly defined process outlining which questions were from which study
21. Erku, 2016	Previously validated questionnaire used:
	Malaysia, questionnaire with community pharmacists
22. Abera et al., 2014	Not reported
23. Ajibade et al., 2014	Not reported
24. Asante et al., 2017	Questions modelled on WHO publication:
	WHO standard methodology for investigating drug use in health facilities
	No further details on adaptation process reported
25. Guma et al., 2022	Based on literature review of 3 previous studies:
	South Africa, survey with primary care prescribers
	South Africa, qualitative study on AB use by prescribers
	China, survey with doctors
	No further details on adaptation process reported

Study ID	Adaptation
26. Ogoina et al., 2021	Based on 4 previous studies:
	Systematic review on physicians AB prescribing
	Systematic review on qualitative studies to understand AB prescribing behaviour
	AMS toolkit for LMICs
	Multi country AMS survey
	No further details on adaptation process reported
27. Dougnon et al., 2020	Not reported
28. Manga et al., 2021	Not reported

Measurement Information

Stu	udy ID	Pretest or pilot	Validity assessment	Reliability assessment	Modality of data
					collection
1.	Chetty et al., 2023	Not reported	Not reported	Internal consistency –	Self-administered
				Cronbach's alpha	
				( $\alpha$ =0.9 for perceptions)	
2.	Tembo et al., 2022	Yes - Tested for	Content validity -	Not reported	Interviewer administered
		consistency, length and	Relevance of the		
		relevance of the	questions assessed by 5		
		questions, not used in the	physicians and 5 intern		
		final analysis, not reported	pharmacists		
		if any changes were made			
3.	Gulleen et al.,	Yes - Pilot tested on 10	Not clear - Experts used	Not reported	Self-administered
	2022	healthcare providers to	in development of		
		optimize clarity and	survey but no		
		readability, not reported if	information on changes		
		used in final analysis or if			
		any changes were made			
4.	Balliram et al.,	Yes - Process not reported	Not reported	Not reported	Self-administered
	2021				

Stu	udy ID	Pretest or pilot	Validity assessment	Reliability assessment	Modality of data
					collection
5.	Mufwambi et al.,	Yes - Process not reported	Not reported	Internal consistency –	Self-administered
	2021			Cronbach's alpha	
				$(\alpha = >0.7 \text{ or questions})$	
				removed)	
6.	Nkinda et al., 2022	Yes – Pilot tested with 10 prescribers & 10	Not reported	Not reported	Not reported
		dispensers, not reported if			
		used in final analysis or if			
_		any changes were made			
7.	Kimbowa et al.,	Yes - Pilot tested with 20	Face validity –	Internal consistency –	Interviewer administered
	2021	healthcare practitioners,	Specialists reviewed for	Cronbach's alpha	
		not reported if used in final	readability, clarity and	$(\alpha=0.9 \text{ for attitude})$	
		analysis or if any changes	comprehensiveness;		
		were made	changes made		
			accordingly		
8.	Kabba et al., 2020	Yes - Pretested with 10	Not clear – experts	Not reported	Self-administered
		doctors, not used in final	reviewed content and		
		analysis and changes	changes were made, but		
		made accordingly	no further information		
			reported		

Study ID	Pretest or pilot	Validity assessment	Reliability assessment	Modality of data
9. Kalungia et al., 2019	Yes – Piloted for consistency, length & relevance with 10 healthcare workers, not reported if used in final analysis, changes made accordingly	Not clear – relevance assessed in the pilot, but no further information reported	Not reported	Self-administered
10. Farley et al., 2018	Not reported	Not clear – discussion with experts but no further information reported	Not reported	Self-administered
11. Adorka et al, 2013	Yes - Piloted with 10 medical doctors and nurses, not reported if used in final analysis, changes made accordingly	Not reported	Not reported	Self-administered

Study ID	Pretest or pilot	Validity assessment	Reliability assessment	Modality of data collection
12. Abubakar et al., 2023	Yes - Process not reported	Not reported	Not reported	Self-administered
13. Adegbite et al., 2022	Yes - Piloted with 10 healthcare workers for comprehension & clarity, not reported if used in final analysis or if changes were made	Not reported	Internal consistency – Cronbach's alpha (α=0.78 for final questionnaire)	Self-administered
14. Domche Ngongang et al., 2021	Not reported	Not reported	Not reported	Self-administered
15. Sneddon et al, 2020	Not reported	Not reported	Not reported	Self-administered
16. Gebrehiwot et al., 2020	Yes - Pretested on 5 health professionals, not reported if used in final analysis or if changes were made	Not reported	Not reported	Self-administered

Study ID	Pretest or pilot	Validity assessment	Reliability assessment	Modality of data collection
17. Mersha, 2018	Yes - Pretested on 5% of the sample size, not reported if used in final analysis, changes made accordingly	Not reported	Not reported	Not clear
<ol> <li>Gebretekle et al.,</li> <li>2018</li> </ol>	Yes - Process not reported	Not reported	Not reported	Self-administered
19. Tafa et al., 2017	Yes - Process not reported	Not clear – experts reviewed for quality and applicability, no further information reported	Not reported	Self-administered
20. Chaw et al., 2017	Yes - Piloted with 7 healthcare workers, not reported if used in final analysis or if changes made accordingly	Content validity – Experts checked quality, completeness & applicability	Not reported	Self-administered

Study ID	Pretest or pilot	Validity assessment	Reliability assessment	Modality of data
				collection
21. Erku, 2016	Yes - Piloted with 20	Content and face validity	Internal consistency -	Self-administered
	community pharmacists,	<ul> <li>Experts and members</li> </ul>	Cronbach's alpha	
	not used in final analysis,	of target population	( $\alpha$ =0.72 for perceptions,	
	changes made accordingly	reviewed, and changes	reported in study	
		made accordingly	development article,	
			Khan et al., 2016)	
22. Abera et al., 2014	Yes - Piloted at local	Not reported	Not reported	Self-administered
	hospital, no further			
	information reported			
23. Ajibade et al.,	Yes - Process not reported	Not reported	Not reported	Self-administered
2014				
24. Asante et al.,	Yes - Pretested with 12	Not reported	Not reported	Interviewer administered
2017	healthcare workers, not			
	included in final analysis,			
	not reported if changes			
	made			
25. Guma et al., 2022	Yes - Pilot with 30 general	Face validity – during	Not reported	Self-administered
	practitioners, not included	pilot, clarity was		
	in final analysis, changes	assessed but not		
	made accordingly	content		

Study ID	Pretest or pilot	Validity assessment	Reliability assessment	Modality of data
				collection
26. Ogoina et al.,	Yes - Pretested by 30	Face and content	Internal consistency –	Self-administered
2021	physicians, not included in	validity – Experts in the	Cronbach's alpha	
	final analysis, changes	field of public health and	(α=>0.75 for all	
	made accordingly	adjustments made	questionnaire	
		accordingly	components)	
27. Dougnon et al., 2020	Not reported	Not reported	Not reported	Not clear (simply states administered)
28. Manga et al., 2021	Yes - Pilot survey with 20% of participants, included in final analysis, not reported if changes made	Not reported	Not reported	Self-administered

# Chapter 4: Assessing response bias in agreement scale responses between individuals from the UK and sub-Saharan African countries

## 4.1 Introduction

The development of culturally appropriate questionnaires has been highlighted in the previous chapters, and ensuring established strategies for limiting method biases is crucial in behavioural science research (Podsakoff et al., 2003; Muniz et al., 2013). The cultural adaptation of measurements and their international use is an important area in psychology assessment (Evers et al., 2013; Muñiz, Elosua, & Hambleton, 2013). Method bias relates to issues with comparability, and can be caused by the questionnaire design, including the response scales (He et al., 2017). Item bias infers that differences in response scores to items may not correspond to true differences in the psychological construct being assessed. If an item displays bias, there is a different probability of indicating an item for individuals from different cultures with the same trait level (He et al., 2017).

Cross-cultural adaptation and consideration of response scales incorporated in questionnaire items is under researched, and existing research tends to focus on translation difficulties and familiarity of scales (Lee et al., 2002). Item bias is another potential problem when using questionnaires cross-culturally, as some items in a questionnaire may not be appropriate for all cultures (López-Roig & Pastor, 2016). To ensure a measure is delivered correctly and adequately, it is crucial that all aspects of the questionnaire are adapted to be context specific to the population, language, and mode of delivery of the survey. This includes the content, format, response scales and the visual presentation of the items of the questionnaire (Survey Research Center, 2016). Given this complexity of using surveys cross-culturally, various methodological considerations are required in the development of new measures and adaptation of existing measurement tools cross-culturally. The limited studies available have explored cultural aspects related to language, and effective methods for translation of Western measures, but fail to include SSA countries (Johnson & Swedlow, 2019; Liu et al., 2018).

In order to measure psychological constructs which potentially influence behaviour, a common approach is the use of self-report questionnaires incorporating varying response scales. This is the predominant method used in research, but psychological constructs used in behavioural science research are considered to widely vary across cultures and countries, resulting in numerous challenges for research (Henrich et al., 2010; Johnson et al., 2005). However, the

literature on psychological constructs, including measurement, is largely dominated by WEIRD populations, as discussed in previous chapters (Henrich et al., 2010). More investigation is needed to consider how to design robust methods to explore and research these constructs effectively across differing cultures (Rad et al., 2018).

Self-report methods have several benefits, including practicality, low cost, and allow researchers to ask respondents to directly report the factors that influence behaviour, including beliefs, attitudes, and intentions. However, cultural differences in cognitive and communicative processes have the potential to affect individuals' responses, but little attention has been directed to examining variations in response styles, thus potentially inaccurately assuming differences are attributed to how the construct in question is perceived. Consistent and systematic variances in response style among countries is evident in the literature, as discussed in Chapter 2. Response biases are a concern in cross-cultural research, and there are several potential sources of this form of bias (Matsumoto & Hee Yoo, 2016). Two main types, acquiescence bias and reference group bias are discussed below.

Response categories can have different meanings in different countries (Johnson et al., 2005; Jurges, 2007) and research exploring cultural differences in response styles largely relate to the tendency to choose extreme positive responses (regardless of content), known as acquiescence bias (Baumgartner & Steenkamp, 2001; Johnson et al., 2005; Jurges, 2007). This is a form of response bias that has been linked to variations across countries in key cultural dimensions, including collectivism, masculinity, power distance, uncertainty avoidance, indulgence, and long-term orientation. However, SSA countries are largely omitted from research, and a large-scale review on response styles, including Likert scales, incorporated research from five of the six continents permanently inhabited, neglecting Africa (Johnson et al., 2011; Liu et al., 2018). This raises fundamental methodological and arguably ethical concerns, as studies use unreliable approaches to reduce bias in SSA contexts, including the exclusion of participants who have difficulties understanding questions response options and citations related to previous validity assessments from different cultures.

Research has highlighted the limitations of cross-cultural comparisons using Likert scales, due to potential differences in reference groups, as outlined in Chapter 1. This is another type of response bias and refers to individuals comparing themselves to people within the same culture, rendering comparisons between cultures problematic (Heine et al., 2002). Likert scales

dominate the field of behavioral science and survey research, which is likely due to the simplistic nature of this method in measuring psychological constructs that cannot be directly observed and require self-report from individuals (Likert, 1932). An example of this dominance is evident in Chapter 3, in which twenty-five of the twenty-eight studies included some form of Likert scale. Generally, research on Likert scales focuses on the limitations of Likert scales, the number of category responses to include, how to analyse the results and whether to treat the resulting data as ordinal or interval, and if categories in the scale should be numbered or labeled (Kusmaryono et al., 2022; Tanujaya et al., 2022; Willits et al., 2016). Cultural considerations regarding the use of Likert scales with populations other those defined as WEIRD, particularly in SSA countries, is limited.

In addition to problems with cross-cultural validity, researchers working in countries other than those in the West should consider the applicability and appropriateness of using questionnaires and/or response scales designed and developed on a population other than the one in question. Several studies have explored measurement equivalence across countries in large scale surveys, including the World Values Survey (Haerpfer et al., 2020), but this work has assessed construct comparability and largely omits specific consideration of item response scales (Davidov et al., 2008; Minkov & Hofstede, 2012). Given these potential biases and problems which can arise, in particular with Likert scales, it is imperative to assess what factors influence the validity of responses in cross cultural research. Furthermore, it is also vital for researchers to consider when and if Likert scales are appropriate for use in SSA countries.

Differential Item Functioning (DIF) methods are used to examine whether individuals' responses to particular items in questionnaires are systematically linked to a particular personal characteristic and are unrelated to the questionnaire's examining construct (Osterling & Everson, 2009). This refers to differences in the way an item functions across demographic groups, which are matched by a construct measured by the questionnaire (Zumbo, 1999). It occurs when individuals from diverse groups show differing probabilities of responding to an item because there are differences between the groups in the underlying construct being measured by the item. DIF displays whether a difference in results between two groups is due to a real difference or due to specific items causing bias in the measurement process (Osterlind & Everson, 2009). This analysis considers whether cultures differ in their individual responses to items, controlling for the underlying construct measured by the scale (Teresi & Fleishman, 2007).

Previous research using DIF methods has explored potential translation effects, it could be that although the cultures assessed use the same language, different cultures perceive the strength of the descriptors differently. A similar DIF study design to Scott et al. (2007) will be used to explore how individuals from SSA countries view the strength of Likert-scale responses compared to individuals from English-speaking Western countries. Although previous research highlights that there are differences in how individuals respond to Likert scales across cultures, little research has formerly assessed this with respect to SSA countries. It is acknowledged that cultures and countries are not mutually exclusive, and cultures can vary within a country and within countries grouped together. However, for the purposes of this study and as a starting point in the literature base, respondents from the UK will represent a culture from the West and the SSA countries included in the World Values Survey Wave 7 (2017-2022) will represent SSA. This survey was chosen as an example of a study conducted globally, including Likert scales and countries from the West and SSA. The aim of this study is to assess whether the response category descriptors of Likert scales are perceived differently by individuals in the UK and individuals in SSA countries.

#### 4.2 Methods

#### World Values Survey

This study is cross-sectional, utilising secondary data analysis from Wave 7 (2017-2022) of the World Values Survey (WVS7, REF). The WVS is a social survey which is conducted globally every 5 years and aims to explore societal values, beliefs and norms. This survey is the largest cross-national research project exploring people's beliefs and values throughout the world (Haerpfer et al., 2020). The data from each wave of the survey is freely available online through the WVS archive database (<u>https://www.worldvaluessurvey.org/wvs.jsp</u>) and it is conducted in more than 90 countries, including LMICs from all regions in the world. The inclusion criteria included individuals aged 18 and over, residing within private households in each country region, regardless of language, citizenship, or nationality. The WVS was chosen as it was an open access dataset with questions assessing psychological constructs using Likert scales and included an overall scaled score, which was difficult to find in surveys conducted in SSA countries. The survey also included several SSA countries with a large number of respondents to support a DIF analysis.

#### Items selected

The survey comprised of 290 items, with only one scale ('gender equality') utilising an agreement Likert response format, and one other scale ('scepticism') incorporating a confidence agreement Likert response format. Other scales in the questionnaire include multiple choice responses or a dichotomous choice and were not suitable for analysis as the aim of this project was to explore Likert response scales due to the lack of literature and dominance of the Likert scale format in research. This questionnaire is predominantly delivered face to face through interviews by trained interviewers. One country included in this chapter, the UK, also offered a postal (n=747, 31%) and web-based option (n=80, 3%).

These scales were chosen, as opposed to other scales, as the overall measure which these scales contributed to assessed additional constructs. The group with individuals from the UK were chosen as the reference category because research on the development of Likert scales was conducted on WEIRD populations. The factor structure of the gender equality and scepticism scales were assessed through a hierarchical factor analysis in a previous study, indicating acceptable loading scores (Welzel, 2013).

Wave 7 included 4 SSA countries, Nigeria (n = 1237), Ethiopia (n = 1230), Kenya (n = 1266) and Zimbabwe (n = 1215; Total = 4948), and the UK (n = 2375). The questionnaire is translated into various national languages, but for the purposes of this chapter, only English language responses were included in the analyses. This is to avoid the complexities of potential response differences relating to item translation. Wave 7 data collection ran between January 2017 and September 2022.

'Equality' assesses an individual's views on gender equality, relating to equality values in terms of support of women's equal access to education, jobs, and power (Welzel, 2013). This was assessed with three questionnaire items in which respondents were asked to indicate how strongly they disagreed with the following statements:

- "Education is more important for a boy than a girl" (Item 29)
- "When jobs are scarce, men should have priority over women to get a job" (Item 30)
- "Men make better political leaders than women" (Item 33)

Items 29 and 30 assessed participants' agreement to the values questions based on a 4-point Likert scale: 1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree. Item 33 incorporated a 5-point Likert scale, including the same agreement responses as items 29 and 30, with an additional category: 1 = strongly agree, 2 = agree, 3 = neither agree nor disagree, 4 = disagree, 5 = strongly disagree.

'Scepticism' explores individual's views on their confidence towards state institutions, including courts, police, and army (Welzel, 2013). This was assessed in three questionnaires items in which respondents were asked to look at a card and inform the interviewer (or indicate), for each item listed, how much confidence they have in the following:

- "Armed Forces" (Item 65)
- "Justice Systems/Courts" (Item 69)
- "The Police" (Item 70)

All items assessed participants' confidence based on a 4-point Likert scale: 1 = a great deal, 2 = quite a lot, 3 = not very much, 4 = none at all.

#### Statistical analysis

To assess whether individuals in SSA countries perceive response category descriptors differently compared to individuals in the UK, the answers for both groups were compared. Differential item functioning (DIF) analyses were conducted using cumulative odds ordinal logistic regression analyses. There are various available methods to assess DIF, including item response theory and structural equation modelling, but ordinal regression analyses have been used previously for Likert scale response categories (Zumbo, 1999; Scott et al., 2010). This approach is viewed as a robust, effective, and simple method to detect DIF (Scott et al., 2010). An ordinal regression analysis was conducted on each individual question item within a scale (of which there were three question items for each scale) using the Likert scale response option to that item as the dependent variable. It is essential to match the groups (i.e., participants from the UK and participants from SSA countries) on their underlying scores (through using the overall scale scores), to explore the effect that the country grouping has on the item Likert response (Zumbo, 1999). Therefore, the overall scale score (used as a "matching" variable) and a variable representing each cultural grouping (UK and SSA) were included as independent variables in the model. A dummy variable was created to group individuals into either the UK or SSA grouping, and the countries represented in each grouping are outlined in Table 8 below. A total of six ordinal regression models were conducted for six individual Likert scale items included in the questionnaire. For the three items in the gender equality scale, the regression analyses were adjusted for sex (male or female). There were no other options available for sex and this information was chosen by the survey interviewer based on appearance. SPSS statistics version 28 was used for the analysis. Currently, there is little guidance on the sample size needed for DIF analyses, but based on previous literature, a sample size of more than 200 individuals per group is required for DIF using logistic regression (Scott et al., 2007; Scott et al., 2010).

#### Interpretation

Due to the number of statistical tests being conducted, a conservative cut-off of p<.001 was used to indicate statistical significance. To consider DIF effects of a significant magnitude, a log odds ratio of >0.64 or <-0.64 is also required (Cameron et al., 2014; Scott et al., 2010). In terms of odds ratios, for a log odds ratio >0.64, the odds ratio would be >1.89, and for a log odds ratio <-0.64, the odds ratios, the 95% confidence intervals and odds ratios were derived from the regression analyses and are reported below. A log odds ratio greater than zero would highlight that individuals from the SSA grouping were more likely to

99

indicate a higher Likert scale score (indicating less agreement or less confidence) than individuals from the UK (the reference group). A log odds ratio less than zero would indicate that individuals from the SSA grouping were less likely to select a higher Likert scale score than individuals from the UK grouping.

The assumption of proportional odds for each regression model was met and checked through running individual binomial logistic regressions on each cumulative split of the Likert scale categories for each item and comparing the parameter estimates for all three items on each subscale.

## 4.3 Results

## Descriptive data

Between all five countries, there were 4357 participants in total. Respondents with missing data for any of the Likert items or scale scores relevant to the analysis were removed, to avoid potential complications with matching the country groupings to their underlying total scale score and affecting the analyses. Table 8 highlights the missing cases which were removed prior to analysis and provides an overview of the number of participants from each country. Respondents from Ethiopia, Kenya, Nigeria, and Zimbabwe were grouped together to form the SSA group (n = 1982).

## Table 8

## Country frequencies

	Total prior to	Not in	Missing	Total cases in	
	analysis (n) English (n)		responses	analysis (n/%)	
			(n)		
Ethiopia	1230	1220	0	10 (0.2%)	
Kenya	1265	175	44	1046 (24.0%)	
Nigeria	1237	562	22	653 (15.0%)	
Zimbabwe	1215	934	8	273 (6.3%)	
United Kingdom	2607	1	231	2375 (54.5%)	
Total	7554	2892	305	4357 (100%)	

Table 9 provides relevant participant demographic information. Age was missing for 155 participants and sex was missing for 37 participants.

Participant demographic information

	Total sample	UK group	SSA group
	M/n (SD or %)	M/n (SD or %)	M/n (SD or %)
Age (years)	41.87 (18.30)	51.64 (18.11)	30.74 (10.51)
Sex (male)	2064 (47.40%)	1022 (43.62%)	1042 (52.24%)

*Note*. M = mean; n = number; SD = standard deviation.

An independent samples t-test was performed between age and country grouping (UK or SSA), which revealed a statistically significant difference between the age of participants in the UK group compared to the SSA group. Individuals in the UK group were more likely to be older than individuals in the SSA group, t(3684.77) = 46.54, p<.001. Similarly, an independent samples t-test between sex and country grouping was statistically significant, t(4318) = 5.98, p<.001, highlighting that there were more females in the UK group than in the SSA group.

For the three agreement Likert scale items, the response category frequencies are outlined in Table 10 and for the three confidence Likert scale items, this information is provided in Table 11.

Response category frequencies for gender equality scale items

Item number	Country	Strongly	Agree	Disagree	Strongly	Neither agree
	grouping	Agree	n (%)	n (%)	Disagree	nor disagree
		n (%)			n (%)	n (%)
Item 29 "Men make better political	UK	44 (2%)	146 (6%)	1087 (46%)	1098 (46%)	N/A
leaders than women"						
	SSA	473 (24%)	528 (27%)	662 (33%)	319 (16%)	N/A
Item 30 "University is more	UK	18 (1%)	40 (2%)	789 (33%)	1528 (64%)	N/A
important for a girl than a boy"						
	SSA	188 (9%)	235 (12%)	756 (38%)	803 (41%)	N/A
Item 33 "Men should have more	UK	18 (1%)	91 (4%)	746 (31%)	1295 (55%)	225 (9%)
right to a job than women"						
	SSA	207 (10%)	661 (33%)	668 (34%)	233 (12%)	213 (11%)

Response category frequencies for scepticism scale items

Item number	Country	A great deal	Quite a lot	Not very	None at all
	grouping	n (%)	n (%)	much	n (%)
				n (%)	
Item 65 "For the following organizations, can you indicate	UK	774 (33%)	1163 (49%)	366 (15%)	72 (3%)
how confident you are in them (list options including):					
Armed Forces"	SSA	400 (20%)	634 (32%)	665 (34%)	283 (14%)
Item 69 "For the following organizations, can you indicate	UK	370 (16%)	1235 (52%)	647 (27%)	123 (5%)
how confident you are in them (list options including): The					
Police	SSA	213 (11%)	375 (19%)	710 (36%)	684 (34%)
Item 70 "For the following organizations, can you indicate	UK	362 (15%)	1235 (52%)	638 (27%)	140 (6%)
how confident you are in them (list options including):					
Justice Systems/Courts"	SSA	305 (15%)	581 (29%)	704 (36%)	391 (20%)

For items 29, 30 and 33, the regression analyses were adjusted to account for any potential influence of sex, as these items assessed values related to gender equality. Table 12 outlines the responses to each Likert scale response category for males and females from the UK and SSA countries.

Response category frequencies for males and females within each country grouping

Item	Sex	Country	Strongly	Agree	Disagree	Strongly	Neither	Total
		grouping	agree	n (%)	n (%)	disagree	agree nor	
			n (%)			n (%)	disagree	
							n (%)	
Item 29	Male	UK	26 (3%)	95 (9%)	547 (53%)	354 (35%)	N/A	1022
"Men make		SSA	299 (29%)	292 (28%)	307 (29%)	144 (14%)	N/A	1042
better		00,1	200 (2070)	202 (2070)		(		1012
political	Female	UK	15 (1%)	51 (4%)	524 (40%)	731 (55%)	N/A	1321
leaders than								
women"		SSA	171 (18%)	235 (25%)	354 (38%)	175 (19%)	N/A	935
	Male	UK	7 (1%)	25 (2%)	451 (44%)	539 (53%)	N/A	1022
Item 30 "University is more important for a girl than a boy"		SSA	120 (12%)	147 (14%)	397 (38%)	378 (36%)	N/A	1042
	Female	UK	10 (1%)	15 (1%)	331 (25%)	965 (73%)	N/A	1321
		SSA	68 (8%)	87 (9%)	357 (38%)	421 (45%)	N/A	935

Item	Sex	Country	Strongly	Agree	Disagree	Strongly	Neither	Total
		grouping	agree	n (%)	n (%)	disagree	agree nor	
			n (%)			n (%)	disagree	
							n (%)	
Item 33 "Men should have more right to a job than women"	Male	UK	11 (1%)	47 (5%)	372 (36%)	464 (45%)	128 (13%)	1022
		SSA	137 (13%)	407 (39%)	295 (28%)	89 (9%)	114 (11%)	1042
	Female	UK	7 (1%)	43 (3%)	364 (27%)	814 (62%)	93 (7%)	1321
		SSA	69 (7%)	253 (27%)	370 (40%)	144 (15%)	99 (11%)	935

## Regression analyses

For item 29, the final ordinal logistic regression model significantly predicted the item response category over and above the intercept-only model,  $\chi^2(2)=4471.17$ , *p*<.001. This was similar for item 30,  $\chi^2(2)=3099.75$ , *p*<.001, and item 33,  $\chi^2(2)=5541.73$ , *p*<.001. For item 65, the final ordinal logistic regression model significantly predicted the item response category over and above the intercept-only model,  $\chi^2(2)=3795.08$ , *p*<.001. This was similar for item 69,  $\chi^2(2)=5825.40$ , *p*<.001, and item 70,  $\chi^2(2)=4744.97$ , *p*<.001.

Tables 13 provides an overview of the ordinal logistic regression models for the gender equality scale items. None of these items met the statistical significance criteria. The regression models for the three items in the scepticism scale are also highlighted in Table 13. Compared to individuals in the UK, those from the SSA grouping with similar views on the scepticism scale were less likely to indicate low confidence with regards to the police (item 69), compared to the other items in this scale. Compared to individuals in the UK, those from the SSA grouping with similar views on the SSA grouping with similar views on the scepticism scale other items in this scale. Compared to individuals in the UK, those from the SSA grouping with similar views on the scepticism subscale were more likely to indicate low confidence with regards to justice/system courts (item 70), compared to the other items in this scale.
# Table 13

Summary of individual ordinal logistic regressions predicting Likert response category for each question item

Question item	Variable	B <sup>a</sup>	SE	Wald	р	Exp (B) 95% CI for Odds Ratio		
				statistic		Odds ratio	Lower	Upper
Item 29 (Model 1)	UK/SSA	0.17	0.08	4.85	0.03	1.19	1.02	1.39
	Sex	-0.12	0.07	3.00	0.08	0.89	0.78	1.02
Item 30 (Model 2)	UK/SSA	-0.88	0.08	104.26	<.001	0.42	0.35	0.49
	Sex	-0.05	0.07	0.52	0.47	0.95	0.82	1.09
Item 33 (Model 3)	UK/SSA	0.44	0.07	31.54	<.001	1.55	1.33	1.80
	Sex	0.01	0.07	0.02	0.89	1.01	0.88	1.16
Item 65 (Model 4)	UK/SSA	-0.03	0.07	0.26	0.61	0.97	0.85	1.10
Item 69* (Model 5)	UK/SSA	-0.69	0.08	84.80	<.001	0.50	0.43	0.58
Item 70* (Model 6)	UK/SSA	0.82	0.07	122.57	<.001	2.26	1.96	2.61

Note. Total scale score for the gender equality and scepticism scales were included in the relevant regression analysis

<sup>a</sup>B = log odds ratio

\*Meets criteria for important DIF (B< or > 0.64 and p<.001)

### 4.4 Discussion

Overall, this study used data from the WVS to assess whether the response category descriptors of Likert scales for six items were perceived differently by individuals in the UK and individuals in SSA countries. The results demonstrate that DIF appeared to be evident in two items out of the six, after controlling for potential effects of sex in the three items which assessed values in relation to gender equality. However, these significant results are in opposite directions, which could potentially indicate an issue with "pseudo-DIF". This can occur in scales with very few items, as is the case here, and means that one question item could be causing an apparent and opposite DIF effect in another item in the same scale, even although the other question item is not biased (Scott et al., 2010). The log odds ratios for the three items within each of the scales sums to roughly zero, meaning that the log odds ratio for a question item which does not demonstrate real DIF could be showing in the opposite direction as a way of compensating for the question item with true DIF (Scott et al., 2010). The results of these significant regression analyses are too complex to interpret without further information. Ideally, DIF should be undertaken in combination with qualitative research, as DIF is a statistical approach to test for bias and does not provide explanations for the bias (Teresi & Fleishman, 2007). Thus, interviews or open text boxes, encouraging participants to give comments on specific items would have given further information to support the interpretation (Scott et al., 2008). As this was a secondary analysis, this was not possible.

These results potentially highlight that the Likert response scale was perceived differently by individuals from both country groupings, but this is impossible to determine due to the issues outlined above. This was not supported by the regression analyses undertaken to assess DIF in the other question items. If one of the question items is demonstrating true DIF, this could be due to a number of differences between the SSA and UK populations, for example certain populations may be more willing to disclose their true opinions, interpretation of the Likert scale responses may differ culturally or interpretation of the question item itself could differ culturally (Breslau et al., 2008; Scott et al., 2010). It is also important to consider whether this result highlights benign DIF or adverse DIF. Benign DIF occurs when the two groups being assessed differ in their probabilities of endorsing a question item because the item is assessing an aspect of the constructs measured by the scale differently between the two groups. Whereas adverse DIF occurs when the two groups actually differ in the probabilities of endorsing a question item probabilities of endorsing a question item because the item is descented by the scale differently between the two groups. Whereas adverse DIF occurs when the two groups actually differ in the probabilities of endorsing a question item because of endorsing a question item

Likert scales response categories). These are important considerations, as benign DIF is not an indication of a measurement error and only adverse DIF highlights any potential biases in the question items (Breslau et al., 2008).

In this example, DIF refers to the probability of individuals who have similar views on the underlying construct (equality or scepticism) differing in endorsing a particular Likert response category when they belong to different groups (in this case UK or SSA). However, it should be noted that for items 29, 30 and 33 (Table 13), there were massive differences between the UK and SSA in terms of the numbers of individuals from each group who highlighted particular Likert response category options. This led to small cell counts in certain places, affecting the validity of the regression analyses which could have concealed any potential DIF effects. Similarly, there were also large differences between the males and females in the UK and SSA (Table 12) in terms of the numbers who choose certain Likert response options, again leading to small cell counts. This is likely to have dominated any potential DIF effect which may have existed in these three items.

The choice of psychological constructs to use as an example in this study was extremely limited, due to the lack of Likert response scales in global research in general, and specifically in the World Values Survey. Some researchers have hypothesized that this lack of use is due to challenges of using Likert scales, which are infrequently discussed in the literature. But in the case of the World Values Survey, most question items were assessed using a dichotomous scale, 0-10 scale or rank order scale, likely due to challenges with using Likert scales across cultures. Nevertheless, the constructs assessed in this study are of relevance in the field of behavioural science, as beliefs in relation to gender equality can have an impact on observed gender differences in self-reported well-being and psychological distress (Elwer et al., 2013; Tesch-Romer, Motel-Klingebiel & Tomasik, 2007). Scepticism towards state institutions can also influence health-related behaviours, and a prime example of this was evident during the COVID-19 outbreak, in which beliefs in relation to trust and confidence in governments and healthcare systems affected adherence to COVID-19 prevention behaviours (Latkin et al., 2021). Even although it is impossible to interpret the potential DIF results, future studies should explore this further with a larger number of participants in each grouping within the regression analyses to reduce the small cell counts which are potentially masking DIF effects in the question items.

### Limitations

A recommendation for future research is that all WVS items should be considered to get a more holistic view of the whole scale and its subscales in terms of item bias. The majority of the questions included in the WVS assessed psychological constructs through dichotomous or ranking questions, which limited the number of items available to answer the aim of this study. Only uniform DIF was assessed, exploring if items favour one group over the other for all items in the total score, therefore the DIF is in the same direction across the entire spectrum of the measured construct (Martinková et al., 2017). Future research should also study non-uniform DIF, where the regression includes an interaction term between total sub score scale and grouping factor to explore whether this has an effect on the item score (Zumbo, 1999). However, this was not required in this study as DIF was only detected in two items on the scale and likely related to "pseudo DIF". A drawback of using the logistic regression models to assess DIF is that the scale total score may not be an adequate matching variable, especially as this is a very short scale (Scott et al., 2010).

The SSA country grouping is only represented by four SSA countries, and problems with this grouping are highlighted in Chapter 1. However, the assumption was that in grouping countries together in this way, the SSA countries would broadly represent similar cultural effects when compared to the UK, but the choice of countries was based on the available respondents in the World Values Survey for Wave 7. This has implications for the generalizability of the results, across SSA countries and extending wider to other populations. In future, research exploring perceptions of Likert scales would ideally be collated from multiple SSA countries, allowing comparisons within the SSA country grouping which would give insights into similarities and/or differences across countries. Related to this, the number of participants excluded due to the responses not being in English is problematic, as this potentially limits and biases the data towards individuals who can read and write in English. This was necessary to avoid potentially confounding factors relating to difficulties in translating question items causing biases in the measure.

There is a diverse range of responses for each item within the scales, and this also differs within country groupings across the items (Tables 10 & 11). Thus, the overall scale score may not have been the most robust matching variable, as the scale only has 3 items and this may not be an accurate representation of the underlying constructs of equality and scepticism (Scott et al., 2010). Additionally, DIF analysis is not able to determine whether any detected differences are

due to differences in perceptions of the response categories, or differences in understandings of the psychological construct across country groupings. Therefore, further studies should include an additional component, such as qualitative follow up questions, to assess this further.

## Conclusion

Overall, only two items demonstrated a potential DIF effect, which was potentially related to complications with "pseudo DIF". Whether these differences were caused by true differences in interpretation of Likert response categories, or cultural differences on viewpoints in relation to scepticism is too complex to determine. The availability of datasets comparing SSA countries and Western countries was limited, and further reduced when considering datasets with Likert scale items. This could be due to issues with Likert scales, highlighted throughout this thesis, and should be discussed in the literature in more depth. Further studies should explore this further by including larger participant numbers and/or scales with several question items to increase the robustness of the DIF analyses.

### **Chapter 5: Discussion**

### 5.1 Overview

Overall, the main objective of this project was to assess what researchers should consider when developing behavioral science measures to explore self-report constructs in SSA countries. This thesis highlights three key studies, which iteratively developed throughout the project based on the results from the previous study. Chapter 2, the qualitative study assessing behavioural science researchers' experiences of the methodological aspects of using self-report questionnaires in SSA countries highlighted challenges researchers had experienced, which can be viewed as key considerations for researchers in the future. These include difficulties with conceptual understandings of both question items and response scales within self-report measures in general, challenges with the use of Likert rating scales and challenges relating to translation, adaptation, and conceptual equivalence. Some of the difficulties outlined are not unique to populations in SSA countries, but those relating to conceptual understandings of constructs, items and response scales are particularly relevant in considering how to use measures in SSA countries. Identifying the challenges and potential solutions can support research practices conducted in the future, to ensure measures are culturally appropriate and valid for the populations under consideration in SSA countries, and not just for individuals who are WEIRD (Hruschka, 2020).

The findings from Chapter 2 also provided potential solutions for researchers to tackle these methodological difficulties in using self-report measures in SSA countries, including investment in adaptation and development, utilising local research expertise, and incorporating qualitative methods to check potential issues with Likert response scales. These potential solutions outlined by participants in Chapter 2 lead to the development of a systematic review protocol, to understand if these methodological challenges were of relevance and to consider which of the solutions outlined were potentially part of practice already. Chapter 3 described this systematic review, looking specifically at what self-report measurement tools have been used to assess the behavioural influences of healthcare workers working in SSA countries (see Chapter 3). Antimicrobial use was chosen as an example topic to focus on in this review, as this is a worldwide public health concern which involves a range of complex, multifaceted behaviours linked to inappropriate prescribing and overuse of antimicrobials (WHO, 2011; WHO, 2015). Quantitative, self-report measures which were developed and disseminated on this topic area were identified and key aspects relating to measurement properties, and specifically the

development process for response scales, were reviewed. The importance of this study is evident from the literature outlined in Chapter 1, highlighting significant biases in the use of participants and authorship in journal articles from WEIRD participants (Arnett, 2009; Thalmayer et al., 2020). These two papers are cited throughout this thesis, as there is a limited body of research focusing on this issue, however it is imperative to consider which research methods are culturally relevant and appropriate in populations other than those containing WEIRD individuals (Hruschka et al., 2012).

Chapter 2 & 3 report interlinking findings in relation to measurement development in SSA countries, as the systematic review highlighted that studies were predominantly developed either using Western based studies, authors with affiliations to Western institutions and/or included Likert scale response categories with no reported information on validation or adaptation. The dominance of Likert scales in the published research was highlighted in Chapter 3, however, none of the included studies reported the piloting, or pre-testing of Likert scales in the measure development process. The author chose to explore the potential issues with Likert scales in detail in Chapter 4. This secondary data analysis study described assessing whether individuals from SSA countries and the UK perceived the Likert response scale categories differently. Although the findings from this study were too complex to interpret, largely due to potential limitations of the available data, two items out of six suggested a potential DIF effect for one item. There is a case to be made that given the findings in this thesis, validating, and assessing the reliability of Likert response scales should become part of the development process for behavioural science measures in SSA countries. Researchers should also consider alternative, more reliable approaches to Likert scales, where possible.

### 5.2 Importance of findings for research practice

This is a core theme within the three studies outlined in Chapters 2-4; the need to consider adaptation thoroughly and comprehensively in the development of measures in behavioural science. As evidenced and discussed in Chapter 1, the dominance of WEIRD populations in all facets of behavioural science and psychological research causes bias and raises questions regarding the validity and appropriateness of research theories, models and research methods, the latter of which was the focus of this thesis (Adetula et al., 2022; de Oliveria & Baggs, 2023; Henrich et al., 2010; Hruschka, 2020). The studies outlined above raise questions regarding the use of Likert response scales, the dominance of measurement practices tested on and developed for WEIRD populations, and the inclusion of local research expertise in measurement

development. Even in small scale and time limited projects, not allocating resources to these potential challenges could be detrimental to the relevance of the research output. Studies should use validation and reliability assessments to determine if a developed self-report measure is robust. The guidelines highlighted in Chapter 1, specifically the protocol developed by Ambuehl & Inauen (2022), relate to the challenges with differences in conceptual interpretations of psychological constructs identified by researchers in Chapter 2. This protocol provides researchers with guidance on how to approach the adaptation of psychological constructs to address the issues of cultural relevance and differing interpretations. Furthermore, the PROMIS standards (2013) and Survey Research Center guidelines (2016) are key resources which should be highlighted to researchers working in SSA countries.

Furthermore, there are inherent challenges with accurately and consistently defining psychological constructs across and even within differing disciplines in HICs. Recent research highlights limitations in construct validity studies and problems with definitions and clarity of psychological constructs (Flake et al., 2017; Peters & Crutzen, 2024). The measure's ability to predict the outcomes is also highly relevant, and in Chapter 3 this relates to the healthcare workers' antimicrobial resistance behaviors. Several of the included studies did not directly assess the outcome, behavior, and in future including methods to measure the behavior of interest could provide additional information on the predictive validity of the measures.

These issues are particularly pertinent for researchers working in SSA countries who are from WEIRD populations and/or were trained in Western institutions. Being aware of the potential limitations with quantitative self-report measures and reflecting on alternative methods for collecting data and ensuring adequate adaptation of all aspects of a measure, including understandings and interpretations of psychological constructs, and translation and adaptation of question items and response scales, is fundamental to ensure measures are appropriate in SSA countries. There are other methods of measuring behavioural influences which have recently gained more attention in the literature and could potentially address some of the key issues which have been outlined in this thesis. Visual Analogue Scales (VAS) permit an infinite number of endpoints, reduce the emphasis on psychological constructs which are prone to cultural variations, and these scales can be applied in a variety of formats (Klimek et al., 2017). Preliminary studies highlight the potential acceptability of VAS in SSA countries, however this method requires more research and consideration in relation to measurement adaptation across cultures (Finitsis et al., 2016).

Experimental methods have also been used in certain SSA countries, which do not require participants to directly self-report on psychological constructs. Vignettes, short narratives which outline a hypothetical person in a particular context, are used to assess psychological constructs indirectly through a range of methods, including visual, auditory, observational and illustrative. These narratives are used to elicit information on psychological constructs, such as attitudes, beliefs and behaviours, thus potentially limiting some of the issues raised in this thesis regarding self-reporting (Jenkins et al., 2010). However, vignettes may still include Likert scales to assess respondents' judgements regarding the vignettes. Factorial Survey Experiments, also known as vignette experiments, incorporate multiple vignettes which vary in terms of the attributes included in the narrative, allowing researchers to determine the relevance of the attribute on the respondents' judgements (Auspurg & Hinz, 2014). Research in SSA countries indicates this approach could be designed in a way which is culturally relevant and acceptable, potentially limiting some of the challenges with translation of complex psychological constructs (Liebe et al., 2020). Discrete Choice Experiments are also increasingly used by behavioural science researchers, which elicit preferences from individuals indirectly through presenting a number of alternative hypothetical scenarios which include several attributes. A systematic review found that DCEs were useful in assessing behavioral influences related to HIV prevention and intervention in SSA countries (McGrady et al., 2021).

Regardless of which measurement method is utilized, behavioural science approaches must strive to become more inclusive and less reliant on dominant, Eurocentric measurement approaches (Hruschka et al., 2018). Bottom-up approaches are required in research to decolonize psychology, to avoid preconceived notions and paternalistic, potentially patronising approaches, with a view instead to understand and question the theories, beliefs and questions which are based on a Eurocentric approach (Bulhan, 2015). This could involve investing in research development practices which aim to understand psychological constructs from the viewpoint of individuals in SSA countries, setting aside Western based preconceptions on what psychological constructs should be and how they should be measured. Researchers should strive to include expertise from relevant SSA countries in all aspects of research, but also support authorship contributions and potentially help to upskill research colleagues to support them with publishing.

#### 5.3 Reflections

Throughout this thesis, the author has continually reflected on their own bias as a Western based psychologist and researcher, who largely meets the WEIRD criteria. Collaborating with researchers who did not meet the WEIRD criteria, to have discussions and support the interpretation of results in Chapters 2 & 3, was crucial to try and limit some of the biases from the author. The intention of this work was to support the author, and other researchers, in developing research practices to use in SSA countries which were not based just on preconceived notions of how research is conducted in the West. This can lead to reinforcing inherent, colonial perspectives that Westerns know what is best for everyone and how things should be done. This is a very relevant issue in global health partnerships and cross-cultural projects, which may have time and financial restrictions limiting resources for the research development process. The author struggled to find information on this in the literature, and although discussions had occurred anecdotally with other researchers, there was little published research on this. However, since starting this thesis in 2019, more research has been published on WEIRD populations and the huge bias in psychological research (and research in general). This is a fundamental problem which will hopefully continue to gain increased attention.

### **5.4 Limitations**

It must be highlighted that only a proportion of SSA countries were represented in the work presented, and the author is aware that this is not representative of the variation in cultures across countries. This is similar to the number of Western and European countries outlined in the research, which again are not fully representative of the WEIRD populations discussed throughout. The restriction of including only English speakers and studies written in English further limits the generalizability of the findings. Although the intention of this thesis was to explore measurement development in SSA countries, an inherent complication of the limited literature resulted in an over representation of some SSA countries and others not receiving representation in this thesis. Focusing future research efforts on grey literature and countries underrepresented in published data is crucial. The author hypothesized in Chapter 1 that the literature conducted in SSA countries would be sparse appears to be supported, especially in relation to the limited number of articles found in the systematic review (Chapter 3) and the extensive difficulties in sourcing an existing dataset which included Western countries and SSA countries to answer the research question in Chapter 4. This further emphasizes the points raised in Chapter 1, relating to the published literature being dominated by WEIRD samples and authors. Clearly more work is needed to promote and encourage research conducted by and

within populations not represented by WEIRD samples, but this is out with the scope of this thesis to discuss.

Additionally, as mentioned in Chapters 2 & 3, behavioral science and health psychology are relatively new disciplines which are rapidly developing and broadening our understanding of how to change behaviour. There is a focus on improving and supporting other disciplines with incorporating psychological theories, models and frameworks to behaviour change interventions. Part of the difficulty in this work is the varying definitions and labels given to psychological constructs (De Boeck et al., 2023; Michie et al., 2005; Michie et al., 2011). The complexity of ensuring homogeneity in defining a psychological construct across numerous disciplines has implications for this thesis and the research body in general. Several researchers who participated in Chapter 2 were from environmental health, public health and health systems backgrounds. As such, these participants were less likely to have had a strong psychology foundation and this may have affected their views on interpretation of constructs. This was further evident and discussed in the systematic review outlined in Chapter 3, as most studies were conducted and authored by healthcare professionals with a medical background. This likely affected the self-report measurement tool development, as differing interpretations of psychological constructs raise concerns relating to content validity (Haynes et al., 1995). In future, studies focusing on researchers from behavioural science and/or health psychology disciplines could provide more discipline specific information on what needs to be considered when developing self-report measures in SSA countries, further highlighting the crucial need to clearly define psychological constructs.

### 5.5 Conclusion

Nevertheless, the work outlined highlights key challenges in conducting behavioural science research in SSA countries, which should be further discussed and explored, particularly as this relatively new discipline continues to develop, expand, and partner with international organisations to support health initiatives across the globe (Dima et al., 2023). Health psychologists can play a key role when working with international organisations and global health partnerships, to encourage research practices in this field which consider and reflect on the culturally appropriateness of WEIRD based norms in data collection. At the same time, highlighting the necessity to consider the conceptual interpretations of psychological constructs, Likert scale use and contributions from local research expertise in all aspects of the self-report

design process in populations within SSA countries, from tool development and piloting, right through the dissemination and publication of results.

Participant Information Sheet

## **Participant Information Sheet**

# Study Title: Methodological considerations in African countries: An exploratory study on the experiences and opinions of researchers

### Study Chief Investigator: Corina Weir

### **Study Introduction**

You are being invited as part of a research study to take part in an interview, to explore your experiences and opinions of implementing research in African countries. Before you decide if you want to take part, it is important for you to understand the background to this research and what taking part will involve.

Please take time to read the following information carefully and ask us if there is anything that is not clear or if you would like more information. Feel free to take time to decide whether

you wish to take part and thank you for reading this information sheet.

### What is the purpose of the study?

There is a need to establish accepted and valid methods of designing self-report questionnaires in African countries, to support the development of robust questionnaires to produce valid research findings.

The primary aim of this study is to explore researchers' experiences of the methodological aspects of conducting questionnaires in African countries.

### Why have I been chosen?

You are being invited to take part in this study as you are a researcher with experience of using self-report questionnaires in African countries. We are looking to interview between 20-24 researchers in total for this study.

### Do I have to take part?

No, it is entirely up to you to decide whether to take part. If you do decide to take part, you will be asked to electronically complete and sign the attached consent form, or you can print the attached consent form, sign your initials, scan and then email the consent form. You are still free to withdraw at any time and without giving a reason.

### What will happen to me if I take part?

Once you have completed the consent form and returned this to the research team, we will arrange to complete an interview either face to face, over the phone, through skype or using zoom technology. The interview will take place at a time convenient to you and the Chief Investigator (Corina Weir) will conduct all interviews.

The interview questions are designed to explore your experiences and opinions on the methodological challenges of conducting questionnaires in African countries, and the approaches which you have previously used. The interview will take between 45-60 minutes, the length of which will be determined by you. It will be audio recorded and transcribed verbatim immediately after. Transcription may be undertaken by an external company, authorised by the University of Aberdeen, but your personal details will not be given to the GDPR compliant company.

If you decide to withdraw following the interview, the interview data collected up until the point of withdrawal may still be used in analysis.

### What are the possible disadvantages of taking part?

It is unlikely that you will suffer any downside to taking part, as no sensitive information will be asked, and the questions only relate to your research experiences. You are free to stop the interview at any point, without giving any reason, and you can choose not to answer any of the questions.

## What will happen to my data?

The University of Aberdeen complies with the General Data Protection Regulation (2018) and all information will be treated with the strictest confidence. All interview recordings will be kept within a password-protected folder within a password-protected University of Aberdeen computer that only the Chief Investigator and the researcher will have access to. Once the study is complete, all data will be archived electronically in a password-protected file on the University system by the Chief Investigator for a period of 10 years.

### What are the possible benefits of taking part?

Whilst there is no direct benefit to you in taking part in this study, we intend to use the information gained to improve the design of future research projects in African countries. Your views and experiences will be highly valuable to ongoing projects.

# Will my taking part in this study be kept confidential?

The University of Aberdeen is the sponsor for this study based in Scotland, United Kingdom. We will be using information from your interview in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly. Anonymised data from this study may be looked at by individuals from the University of Aberdeen and/or the regulatory authorities, where it is relevant to your taking part in the research. The University of Aberdeen will keep identifiable information about you for 10 years after the study has finished.

Your rights to access, change or move your information are limited, as we need to manage your information in specific ways for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally-identifiable information possible. You can find out more about how we use your information <u>http://www.abdn.ac.uk/privacy</u>

### What will happen to the results of the research study?

The results from this study will be submitted for publication in an academic journal and presented at workshops/conferences. The results may also be written up as part of the Chief Investigator, Corina Weir's academic qualification and submitted to the University of Stirling. Your identity will be kept confidential and you will not be identified in any report/publication. If we use direct quotations from your interview, your name will be replaced with a pseudonym.

### Who is organising and funding the research?

This research is being conducted by Corina Weir, Health Psychologist. The University of Aberdeen is funding this research.

### Who has reviewed the study?

This study has been reviewed and approved by the College Ethical Review Board (CERB) at the University of Aberdeen.

### **Contact for Further Information**

For further information or to discuss any aspect of this project, please feel free to contact the Chief Investigator:

### Corina Weir (Health Psychologist), corina.weir@abdn.ac.uk, 01224 438108, 07791632065

If you wish to make a complaint about the study, please contact:

University of Aberdeen & NHS Grampian Clinical Research Governance, Foresterhill House Annexe, Foresterhill, Aberdeen, AB25 2ZB, researchgovernance@abdn.ac.uk

The above contact is not part of the research term.

# Many thanks for taking the time to read this information and for considering taking part in this project.

# If you would like a copy of the final results, please contact Corina Weir or let us know during your interview.

Consent Form



## Consent form

## Participant ID Number:

Name of CI: Corina Weir

Study Title: Methodological considerations in African countries: An exploratory study on the experiences and opinions of researchers

1. I confirm that I have read and understand the information sheet Version No: 2 Date: 26.08.2019 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

Please initial

- 2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. Data collected up until the point of withdrawal may still be used in analysis.
- 3. I understand that data collected during the study may be looked at by individuals from the University of Aberdeen and the regulatory authorities, if appropriate, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my data.
- 4. I agree to my interview being audio recorded. I understand that anonymised quotations from this interview may be used for presentations and publications.
- 5. I agree that my interview may be transcribed by an external company contracted by the University of Aberdeen.
- 6. I agree for my information to be stored on University of Aberdeen servers.
- 7. I agree to take part in the above study.

Name of participant	Date	Signature	
Name of researcher	Date	Signature	

Version 2 (26.08.2019)

Interview Topic Guide

# <u>Opening</u>

- Introductions and thank for participating
- Confirm consent for audio recording (start recorder)

# <u>Questions</u>

- Tell me about your research experience in African countries?
  - Can you specifically outline the methodological aspects of your research?
     (Questions, scales, psychometric properties, etc.)
- (If applicable) In terms of the methodological aspects, what similarities are there between conducting research in African countries and other countries?
- (If applicable) In terms of the methodological aspects, what differences have you found between conducting research in African countries and other countries?
- What are some the key challenges of conducting research in African countries?
  - Probe methodological challenges specifically
  - Probe question design and scales specifically
- What methods have you used to address these challenges?
  - Probe success of methods, what has worked or not worked?
- What methods would you like to use in an ideal world, with abundant resources available, to address these challenges?
- What crucial aspects of methodological considerations do you think all researchers working in African countries should be aware of/consider when designing research?

# <u>Closing</u>

- Anything you would like to add that has not been covered?
- Thank for taking part

Reflective notes and audit trail example

Example 1 (26/02/2022)

# P3 Coding

- The start of the interview involves discussions on adapting existing questionnaires to suit interests, and developing questionnaires based on existing literature (no mention of theory based questionnaire development however did not probe further) Consistent across interviews?
- Participant stated context informs questionnaire development (examples given how live in communities, access health services, cultural, social, economic, political factors) Is this the same for questionnaire development in the West? Differences between countries not as wide?
- This participant stated Likert scales can be useful when assessing attitudes (in particular), perceptions and behaviours however, later in the interview this relates to interviewer delivered questionnaires (in which interviewer asks questions qualitatively and assigns rating based on their opinion) 19.33
- Discussions of researcher administered questionnaires in 2 ways (give respondents choices in ratings of for example 1-5 low -high or ask qualitatively? then the researcher decides rating). Raised elsewhere and interesting point, as I was not aware of this from my own experiences. Are others?
- Alot of discussion regarding social desirability bias and using observational checklists to
   assess behaviour
- I am noticing a large number of nodes, which I am not too concerned about at this point as I want to be open in my analysis plan to discuss with critical friend on Tuesday

Example 2 (29/02/2022)

- Notes during second coding meeting
- Second coder spend the last week and a half second coding 4 interview transcripts
- Deliberately choose 2 in particular which I found arduous to code, with complications in points
- I spent this evening reviewing the second coding prior to moving to the next phase tomorrow
- Very interesting and there is a large cross over in coding, with some codes termed slightly differently, and additional interview parts linked to different codes, but plan to meet and discuss further later in the week
- Commented on individual codes in the shared word documents as discussion points for our meeting
- False responses (reworded from my coding), many self-report measures given as interviews (crucial point, reworded from my coding), clear statements required for questionnaire items (new code need to consider), Likert scales perceived as complicated/difficult (is it a perception? - worth discussing further), Likert scales not ethically ideal? (P13 - interesting code)
- This has helped with my confidence to move onto the next phase of analysis, and this process was very worthwhile
- Plan for tomorrow: Document and review second coding, move onto searching for themes to discuss at supervision meeting on Wednesday

Database search strategies

### Ovid MEDLINE(R) ALL <1946 to May 09, 2023>

PsycINFO and PsycARTICLES

1 (Antimicrobial\* or antibiotic\* or microbial\* or drug resistant or antibacterial\* or drug resistance).mp. [mp=title, book title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms, population supplementary concept word, anatomy supplementary concept word]

2 Community health worker\*or CHWs/ or Lay health worker\*.mp. or Health worker\*.mp. or healthcare worker\*.mp. or doctor\*.mp. or nurse\*.mp. or pharmacist\*.mp. or dentist\*.mp. or healthcare support worker\*.mp. or health practitioner.mp. or healthcare practitioner.mp. Or prescriber\* [mp=title, book title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms, population supplementary concept word, anatomy supplementary concept word]

3 (((Sub-Saharan Africa or East\* Africa or Central Africa or West\* Africa or South\* Africa or Angola or Benin or Botswana or Burkina Faso or Burundi or Cabo Verde or Cameroon or Central African Republic or CAR or Chad or Comoros or Democratic Republic of the Congo or DRC or Republic of the Congo or Cote d'Ivoire or Equatorial Guinea or Eritrea or Eswatini or Swaziland or Ethiopia or Gabon or Gambia or Ghana or Guinea or Guinea-Bissau or Kenya or Lesotho or Liberia or Madagascar or Malawi or Mali or Mauritania or Mauritius or Mozambique or Namibia or Niger or Nigeria or Rwanda or Sao Tome) and Principe) or Senegal or Seychelles or Sierra Leone or South Africa or South Sudan or Tanzania or Togo or Uganda or Zambia or Zimbabwe).mp. [mp=title, book title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms, population supplementary concept word, anatomy supplementary concept word]

## 4 1 and 2 and 3

## African Index Medicus (AIM)

(tw:(Antimicrobial\* OR antibiotic\* OR microbial\* OR drug resistant OR "antibacterial\* AND drug resistance")) AND (tw:(Health worker\* OR healthcare worker\* OR doctor\* OR nurse\* OR pharmacist\* OR dentist\* OR healthcare support worker\* OR health practitioner OR healthcare practitioner or prescriber\*))

## SCOPUS

1. Antimicrobial\* OR antibiotic\* OR microbial\* OR "drug resistant" OR antibacterial OR "drug resistance"

## AND

2. "Community health worker\*"OR CHWs OR "Lay health worker\*" OR "Health worker\*" OR "health care worker\*" OR doctor\* OR nurse\* OR pharmacist\* OR dentist\* OR "healthcare support worker\*" OR "health practitioner" OR "healthcare practitioner" or prescriber\* AND

3. "Sub-Saharan Africa" OR "East\* Africa" OR "Central Africa" OR "West\* Africa" OR "South\* Africa" OR Angola OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR "Cabo Verde" OR Cameroon OR "Central African Republic" OR CAR OR Chad OR Comoros OR "Democratic Republic of the Congo" OR DRC OR "Republic of the Congo" OR "Cote d'Ivoire" OR "Equatorial Guinea" OR Eritrea OR Eswatini OR Swaziland OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR Guinea-Bissau OR Kenya OR Lesotho OR Liberia OR Madagascar OR Malawi OR Mali OR Mauritania OR Mauritius OR Mozambique OR Namibia OR Niger OR Nigeria OR Rwanda OR "Sao Tome and Principe" OR Senegal OR Seychelles OR "Sierra Leone" OR "South Africa" OR "South Sudan" OR Tanzania OR Togo OR Uganda OR Zambia OR Zimbabwe

# Google Scholar

As Google Scholar has a 256 character or 150-word limit for searches, the search strategy was adjusted to the following:

(Antimicrobial OR antibiotic) AND (health worker OR doctor OR nurse OR pharmacist OR dentist OR health practitioner OR prescriber) AND (sub-Saharan AND Africa) AND (behaviour OR behavior) - Limited to 2001-2023 and English language only

After 411 hits, the next 100 hits did not meet the inclusion criteria assessed by the title. Therefore, no more hits were exported.

## Pubmed

(Antimicrobial\* OR antibiotic\* OR microbial\* OR "drug resistant" OR antibacterial OR "drug resistance") AND ("Community health worker\*" OR CHWs OR "Lay health worker\*" OR "Health worker\*" OR "healthcare worker\*" OR doctor\* OR nurse\* OR pharmacist\* OR dentist\* OR "healthcare support worker\*" OR "health practitioner" OR "healthcare practitioner" or prescriber\*) AND ("Sub-Saharan Africa" OR "East\* Africa" OR "Central Africa" OR "West\* Africa" OR "South\* Africa" OR Angola OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR "Cabo Verde" OR Cameroon OR "Central African Republic" OR CAR OR Chad OR Comoros OR "Democratic Republic of the Congo" OR DRC OR "Republic of the Congo" OR "Cote d'Ivoire" OR "Equatorial Guinea" OR Eritrea OR Eswatini OR Swaziland OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR Guinea-Bissau OR Kenya OR Lesotho OR Liberia OR Madagascar OR Malawi OR Mali OR Mauritania OR Mauritius OR Mozambique OR Namibia OR Niger OR Nigeria OR Rwanda OR "Sao Tome and Principe" OR Senegal OR Seychelles OR "Sierra Leone" OR "South Africa" OR "South Sudan" OR Tanzania OR Togo OR Uganda OR Zambia OR Zimbabwe)

Appraisal tool for Cross-Sectional studies (AXIS)

# For each question, yes = 1, no = 0 or don't know

Introduction	
1	Were the aims/objectives of the study clear?
Methods	
2	Was the study design appropriate for the stated aim(s)?
3	Was the sample size justified?
4	Was the target/reference population clearly defined? (Is it
	clear who the research was about?)
5	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?
6	Was the selection process likely to select subjects/participants
0	that were representative of the target/reference population under investigation?
7	Were measures undertaken to address and categorise non-responders?
8	Were the risk factor and outcome variables measured
	appropriate to the aims of the study?
9	Were the risk factor and outcome variables measured
	correctly using instruments/measurements that had been
	trialled, piloted or published previously?
10	Is it clear what was used to determine statistical significance
	and/or precision estimates? (e.g. p-values, confidence intervals)
11	Were the methods (including statistical methods) sufficiently
	described to enable them to be repeated?
Results	
12	Were the basic data adequately described?
13	Does the response rate raise concerns about non-response bias?
14	If appropriate, was information about non-responders described?
15	Were the results internally consistent?
16	Were the results presented for all the analyses described in
	the methods?
Discussion	
17	Were the authors' discussions and conclusions justified by the results?
18	Were the limitations of the study discussed?
Other	
19	Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?
20	Was ethical approval or consent of participants attained?

NHBLI Quality Assessment Tool for Before-After (Pre-Post) Studies With No Control Group

For each question, yes = 1, no = 0, CD = can't determine, NR = not reported

## Criteria

1. Was the study question or objective clearly stated?

2. Were eligibility/selection criteria for the study population prespecified and clearly described?3. Were the participants in the study representative of those who would be eligible for the

test/service/intervention in the general or clinical population of interest?

4. Were all eligible participants that met the prespecified entry criteria enrolled?

5. Was the sample size sufficiently large to provide confidence in the findings?

6. Was the test/service/intervention clearly described and delivered consistently across the study population?

7. Were the outcome measures prespecified, clearly defined, valid, reliable, and assessed consistently across all study participants?

8. Were the people assessing the outcomes blinded to the participants' exposures/interventions?

9. Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?

10. Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?

11. Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention (i.e., did they use an interrupted time-series design)?

12. If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis take into account the use of individual-level data to determine effects at the group level?

### References

- Abera, B., Kibret, M., & Mulu, W. (2014). Knowledge and beliefs on antimicrobial resistance among physicians and nurses in hospitals in Amhara Region, Ethiopia. *BMC Pharmacology & Toxicology*, *15*, 26. <u>https://doi.org/10.1186/2050-6511-15-26</u>
- Adegbite, B. R., Edoa, J. R., Schaumburg, F., Alabi, A. S., Adegnika, A. A., & Grobusch, M. P. (2022). Knowledge and perception on antimicrobial resistance and antibiotics prescribing attitude among physicians and nurses in Lambaréné region, Gabon: a call for setting-up an antimicrobial stewardship program. *Antimicrobial Resistance & Infection Control*, *11*(1), 44. <u>https://doi.org/10.1186/s13756-022-01079-x</u>
- Abubakar B, & Sárváry A. (2023). Knowledge, attitude, and practice on antibiotics use among healthcare workers: A cross-sectional study in Niger state, *Nigeria. J Infect Prev*, 24(5). doi: 10.1177/17571774231165407
- Adetula, A., Forscher, P. S., Basnight-Brown, D., Azouaghe, S., & IJzerman, H. (2022).
  Psychology should generalize from—not just to—Africa. *Nature Reviews Psychology*, 1(7), 370-371. <u>https://doi.org/10.1038/s44159-022-00070-y</u>
- Adorka, M., Dikokole, M., Mitonga, K. H., & Allen, K. (2013). Healthcare providers' attitudes and perceptions in infection diagnosis and antibiotic prescribing in public health institutions in Lesotho: a cross sectional survey. *African Health Sciences*, *13*(2), 344–350.
   <a href="https://doi.org/10.4314/ahs.v13i2.21">https://doi.org/10.4314/ahs.v13i2.21</a>
- Agampodi, T. C., Agampodi, S. B., Glozier, N., & Siribaddana, S. (2015). Measurement of social capital in relation to health in low and middle income countries (LMIC): A systematic review. Social Science & Medicine, 128, 95–104. <u>https://doi.org/10.1016/j.socscimed.2015.01.00</u>
- Ajibade, A., & Olaitan, P. B. (2014). Knowledge, attitude and practice of perioperative antibiotic prophylaxis among nurse-anaesthetists in Nigeria. *Nigerian Journal of Medicine: Journal of the National Association of Resident Doctors of Nigeria*, 23(2), 142–148.

- Ambuehl, B., & Inauen, J. (2022). Contextualized measurement scale adaptation: a 4-Step tutorial for health psychology research. International journal of environmental research and public health, 19(19), 12775.
- Andrade, C., Menon, V., Ameen, S., & Kumar Praharaj, S. (2020). Designing and Conducting Knowledge, Attitude, and Practice Surveys in Psychiatry: Practical Guidance. *Indian Journal of Psychological Medicine*, *42*(5), 478–481. https://doi.org/10.1177/0253717620946111
- Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, *41*(5), 319-329. <u>https://doi.org/10.1016/j.evolhumbehav.2020.07.015</u>
- Ares, G. (2018). Methodological issues in cross-cultural sensory and consumer research. *Food Quality and Preference, 64*, 253-263. <u>https://doi.org/10.1016/j.foodqual.2016.10.007</u>
- Arnett, J. J. (2009). The neglected 95%, a challenge to psychology's philosophy of science. *American Psychologist, 64*(6), 571–574. <u>https://doi.org/10.1037/a0016723</u>
- Asante, K. P., Boamah, E. A., Abdulai, M. A., Buabeng, K. O., Mahama, E., Dzabeng, F., Gavor, E., Annan, E. A., Owusu-Agyei, S., Gyansa-Lutterodt, M., & On behalf of the Ghana Antimicrobial Resistance Working Group. (2017). Knowledge of antibiotic resistance and antibiotic prescription practices among prescribers in the Brong Ahafo Region of Ghana; a cross-sectional study. *BMC Health Services Research*, *17*(1), 422. <u>https://doi.org/10.1186/s12913-017-2365-2</u>
- Atkins, L., Francis, J., Islam, R., O'Connor, D., Patey, A., Ivers, N., Foy, R., Duncan, E. M., Colquhoun, H., Grimshaw, J. M., Lawton, R., & Michie, S. (2017). A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implementation Science*, *12*(1), 77. <u>https://doi.org/10.1186/s13012-017-0605-</u> <u>9</u>

Auspurg, K., & Hinz, T. (2014). Factorial survey experiments (Vol. 175). Sage Publications.

- Balliram, R., Sibanda, W., & Essack, S. Y. (2021). The knowledge, attitudes and practices of doctors, pharmacists and nurses on antimicrobials, antimicrobial resistance and antimicrobial stewardship in South Africa. Southern African Journal of Infectious Diseases, 36(1), 262. <u>https://doi.org/10.4102/sajid.v36i1.262</u>
- Batres, C., Borras Guevara, M. L., & Perrett, D. I. (2018). Collecting data cross-culturally: methodological challenges that arise when testing non-WEIRD populations. SAGE Research Methods Cases. <u>https://doi.org/10.4135/9781526434081</u>
- Baumgartner, H., & Steenkamp, J.-B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38(2), 143–156. https://doi.org/10.1509/jmkr.38.2.143.18840
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186-3191.
- Behr, D. (2017). Assessing the use of back translation: The shortcomings of back translation as a quality testing method. *International Journal of Social Research Methodology*, 20(6), 573-584. <u>https://doi.org/10.1080/13645579.2016.1252188</u>
- Beins, B. C. (2019). Methodological and Conceptual Issues in Cross-Cultural Research. In K. D. Keith (Ed.) Cross-Cultural Psychology (2nd ed.) (pp. 59-81). Malden, MA: Wiley-Blackwell.
- Boateng, Godfred O., Torsten B. Neilands, Edward A. Frongillo, Hugo R. Melgar-Quiñonez, and Sera L. Young. "Best practices for developing and validating scales for health, social, and behavioral research: a primer." *Frontiers in public health* 6 (2018): 149. <u>https://doi.org/10.3389/fpubh.2018.00149</u>
- Bogdan, R. C., & Biklen, S. K. (1998). Foundations of qualitative research in education. *Qualitative Research in Education: An Introduction to Theory and Methods*, 1–48.
- Borghi, J., Lohmann, J., Dale, E., Meheus, F., Goudge, J., Oboirien, K., & Kuwawenaruwa, A. (2018). How to do (or not to do): Measuring health worker motivation in surveys in low-

and middle-income countries. *Health Policy and Planning*, *33*(2), 192–203. <u>https://doi.org/10.1093/heapol/czx153</u>

- Braine, T. (2011). Race against time to develop new antibiotics. *World Health Organization.* Bulletin of the World Health Organization, 89(2), 88. doi:10.2471/BLT.11.030211
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. <u>https://doi.org/10.1191/1478088706qp063oa</u>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597. <u>https://doi.org/10.1080/2159676X.2019.1628806</u>
- Breslau, J., Javaras, K. N., Blacker, D., Murphy, J. M., & Normand, S. L. T. (2008). Differential item functioning between ethnic groups in the epidemiological assessment of depression. The Journal of nervous and mental disease, 196(4), 297.
  DOI: <u>10.1097/NMD.0b013e31816a490e</u>
- Buchanan, N. T., Perez, M., Prinstein, M. J., & Thurston, I. B. (2021). Upending racism in psychological science: Strategies to change how science is conducted, reported, reviewed, and disseminated. *American Psychologist*, *76*(7), 1097. <u>https://psycnet.apa.org/doi/10.1037/amp0000905</u>
- Bulhan, H. A. (2015). Stages of Colonialism in Africa: From Occupation of Land to Occupation of Being. *Journal of Social and Political Psychology*, 3(1), 239-256. <u>https://doi.org/10.5964/jspp.v3i1.143</u>
- Bull, E. R., Mason, C., Junior, F. D., Santos, L. V., Scott, A., Ademokun, D., ... & Cavanagh, S. M. (2017). Developing nurse medication safety training in a health partnership in Mozambique using behavioural science. *Globalization and Health*, *13*(1), 1-10. DOI: <u>10.1186/s12992-017-0265-1</u>
- Byrne-Davis, L. M., Bull, E. R., Burton, A., Dharni, N., Gillison, F., Maltinsky, W., Mason, C.,Sharma, N., Armitage, C.J., Johnston, M., Byrne, G.J. & Hart, J. K. (2017). Howbehavioural science can contribute to health partnerships: the case of The Change

Exchange. *Globalization and Health*, *13*(1), 1-8. <u>https://doi.org/10.1186/s12992-017-0254-4</u>

- Calvo-Villamañán, A., San Millán, Á., & Carrilero, L. (2022). Tackling AMR from a multidisciplinary perspective: a primer from education and psychology. *International Microbiology*, *26*(1), 1–9. <u>https://doi.org/10.1007/s10123-022-00278-1</u>
- Cameron, I.M., Scott, N.W., Adler, M. *et al.* (2014). A comparison of three methods of assessing differential item functioning (DIF) in the Hospital Anxiety Depression Scale: ordinal logistic regression, Rasch analysis and the Mantel chi-square procedure. *Qual Life Res* 23, 2883–2888. <u>https://doi.org/10.1007/s11136-014-0719-3</u>
- Carter, S., Moncrieff, I. S., Bayugo, Y., Hickler, B., Gobat, N., Singh, N., Pham, P., Vinck, P., Poveda, J. D., Lissouba, P., Fricke, C., Shaw, A., Tam, W. J., Austin, C., Berman, G., Wood, G., Padoveze, M. C. & Niederberger, E. (2020). Guidance for health care worker surveys in humanitarian contexts in LMICs. Project Report. CASS/UNICEF. <u>https://reliefweb.int/report/world/guidance-health-care-worker-hcw-surveys-humanitariancontexts-lmics</u>
- CDC. (2021). The Core Elements of Human Antibiotic Stewardship Programs in Resource -Limited Settings: National and Hospital Levels. Atlanta, GA: US Department of Health and Human Services. Available at: <u>https://www.cdc.gov/antibiotic-</u> <u>use/healthcare/implementation.html</u>
- Chambers, J. A., & O'Carroll, R.E. (2016). Adherence to medical advice. In: Benyamini, Y., Johnston, M., & Karademas E.C. Assessment in Health Psychology. Hogrefe.
- Chauhan, B. F., Jeyaraman, M., Mann, A. S., Lys, J., Skidmore, B., Sibley, K. M., Abou-Setta,
   A., & Zarychanksi, R. (2017). Behavior change interventions and policies influencing
   primary healthcare professionals' practice—An overview of reviews. *Implementation Science*, 12(1), 3. <a href="https://doi.org/10.1186/s13012-016-0538-8">https://doi.org/10.1186/s13012-016-0538-8</a>
- Chaw, P. S., Schlinkmann, K. M., Raupach-Rosin, H., Karch, A., Pletz, M. W., Huebner, J., & Mikolajczyk, R. (2017). Knowledge, attitude and practice of Gambian health practitioners

towards antibiotic prescribing and microbiological testing: a cross-sectional survey. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *111*(3), 117–124. <u>https://doi.org/10.1093/trstmh/trx027</u>

- Chetty, D., & Rapper, S. L. (2023). Exploring the discord between pharmacy education and practice in antimicrobial stewardship. *Health SA Gesondheid*, *28*(1). <u>https://www.ajol.info/index.php/hsa/article/view/245361</u>
- Chung, S. J., & Han, J. Y. (2013). Utilizing psychological distances of hedonic phrases to develop a Korean hedonic scale. *Food Quality and Preference*, 28(1), 188–192. <u>https://doi.org/10.1016/j.foodqual.2012.07.011</u>
- Craig, J., Sadoff, R., Bennett, S., Bahati, F., & Beauvais, W. (2023). Behavior-change interventions to improve antimicrobial stewardship in human health, animal health, and livestock agriculture: A systematic review. *PLOS Global Public Health*, *3*(5), e0001526. <u>https://doi.org/10.1371/journal.pgph.0001526</u>
- Davidov, E., Schmidt, P., & Schwartz, S. H. (2008). Bringing values back in: The adequacy of the European Social Survey to measure values in 20 countries. *Public opinion quarterly*, *7*2(3), 420-445. <u>https://doi.org/10.1093/poq/nfn035</u>
- De Boeck, P., Pek, J., Walton, K., Wegener, D. T., Turner, B. M., Andersen, B. L., Beauchaine, T. P., Lecavalier, L., Myung, J. I. & Petty, R. E. (2023). Questioning psychological constructs: Current issues and proposed changes. Psychological Inquiry, 34(4), 239-257.
- de Oliveira, G. S., & Baggs, E. (2023). *Psychology's WEIRD Problems*. Cambridge University Press.
- Dima, A., Byrne-Davis, L., Baban, A., Hart, J., Johnston, M., & Araújo-Soares, V. (2023). EHPS engagement with international bodies on health policy: current contributions and future opportunities. *European Health Psychologist*, *23*(3).
- Domek, G. J., O'Leary, S. T., Bull, S., Bronsert, M., Contreras-Roldan, I. L., Ventura, G. A. B., Kempe, A., & Asturias, E. J. (2018). Measuring vaccine hesitancy: Field testing the WHO

SAGE Working Group on Vaccine Hesitancy survey tool in Guatemala. Vaccine, 36(35), 5273-5281. <u>https://doi.org/10.1016/j.vaccine.2018.07.046</u>

- Domche Ngongang, S. C., Basera, W., & Mendelson, M. (2021). Tertiary hospitals physician's knowledge and perceptions towards antibiotic use and antibiotic resistance in Cameroon. *BMC Infectious Diseases*, *21*(1), 1-11. <u>https://doi.org/10.1186/s12879-021-06792-3</u>
- Dougnon, V., Chab,i Y., Koudokpon, H., Agbankpe, J., Sefounon, R., Alle, D., Bankole, H., Baba-Moussa, L. (2020). Prescription of antibiotics as a source of emerging antibiotic resistance: Knowledge, attitudes, and practices of medical staff in the Dassa-Glazoué and Savalou-Bantè's health zones (Benin, West Africa), *Int. J. One Health*, 6(1): 34-40. www.doi.org/10.14202/IJOH.2020.34-40
- Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ open*, *6*(12), e011458. <u>http://dx.doi.org/10.1136/bmjopen-2016-011458</u>
- Elwér, S., Harryson, L., Bolin, M., & Hammarström, A. (2013). Patterns of gender equality at workplaces and psychological distress. PloS one, 8(1), e53246. <u>https://doi.org/10.1371/journal.pone.0053246</u>
- Epstein, J., Osborne, R. H., Elsworth, G. R., Beaton, D. E., & Guillemin, F. (2015). Crosscultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *Journal of clinical epidemiology*, *68*(4), 360-369. <u>https://doi.org/10.1016/j.jclinepi.2013.07.013</u>
- Erku, D. A. (2016). Antimicrobial Stewardship: A Cross-Sectional Survey Assessing the Perceptions and Practices of Community Pharmacists in Ethiopia. *Interdisciplinary Perspectives on Infectious Diseases*, 2016, e5686752. https://doi.org/10.1155/2016/5686752

- Essack S. Y., Desta A. T., Abotsi R. E. & Agoba E. E. (2017). Antimicrobial resistance in the WHO African region: current status and roadmap for action. J Public Health (Oxf), 1;39(1):8-13. <u>https://doi.org/10.1093%2Fpubmed%2Ffdw015</u>
- Evers, A., Muñiz, J., Hagemeister, C., Høstmælingen, A., Lindley, P., Sjöberg, A., & Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema*, *25*(3), 283-291. DOI: <u>10.7334/psicothema2013.97</u>
- Farley, E., Stewart, A., Davies, M.-A., Govind, M., van den Bergh, D., & Boyles, T. H. (2018). Antibiotic use and resistance: Knowledge, attitudes and perceptions among primary care prescribers in South Africa. SAMJ: South African Medical Journal, 108(9), 763–771. <u>https://doi.org/10.7196/samj.2018.v108i9.12933</u>
- Finitsis, D. J., Pellowski, J. A., Huedo-Medina, T. B., Fox, M. C., & Kalichman, S. C. (2016). Visual analogue scale (VAS) measurement of antiretroviral adherence in people living with HIV (PLWH): a meta-analysis. Journal of behavioral medicine, 39, 1043-1055. <u>https://doi.org/10.1007/s10865-016-9770-6</u>
- Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, 8(4), 370-378. <u>https://doi.org/10.1177/1948550617693063</u>
- Flaskerud, J. H. (2012). Cultural bias and Likert-type scales revisited. Issues in mental health nursing, 33(2), 130-132.
- Food and Agriculture Organization of the United Nations (FAO) and the World Bank. (2018). Food Data Collection in Household Consumption and Expenditure Surveys: Guidelines for Low-and Middle-Income Countries. Retrieved from: <u>https://www.worldbank.org/en/programs/lsms/publication/Food-data-collection-inhousehold-consumption-and-expenditure-surveys-guidelines-for-low-and-middleincome-countries</u>
- Gamble, A. (2021). The Brexit negotiations and the Anglosphere. *The Political Quarterly*, *92*(1), 108-112. <u>https://doi.org/10.1111/1467-923X.12957</u>

- Gebrehiwot, Z., & Tadiwos, Y. (2020). Knowledge and Beliefs of Health Care Professionals Towards Antimicrobial Resistance in Hiwot Fana Specialized University Hospital, in Harar, Ethiopia. *Infection and Drug Resistance*, *13*, 2027–2035. <u>https://doi.org/10.2147/IDR.S254237</u>
- Gebretekle, G. B., Haile Mariam, D., Abebe, W., Amogne, W., Tenna, A., Fenta, T. G., Libman, M., Yansouni, C. P., & Semret, M. (2018). Opportunities and barriers to implementing antibiotic stewardship in low and middle-income countries: Lessons from a mixed-methods study in a tertiary care hospital in Ethiopia. *PLoS ONE*, *13*(12), e0208447. <a href="https://doi.org/10.1371/journal.pone.0208447">https://doi.org/10.1371/journal.pone.0208447</a>
- Gulleen, E., Lubwama, M., Komakech, A., Krantz, E., Liu, C., & Phipps, W. (2022). Knowledge and perceptions of antimicrobial resistance and antimicrobial stewardship among staff at a national cancer referral center in Uganda. *Antimicrobial Stewardship & Healthcare Epidemiology*, 2(1), E54. <u>doi:10.1017/ash.2022.28</u>
- Guma, S. P., Godman, B., Campbell, S. M., & Mahomed, O. (2022). Determinants of the Empiric Use of Antibiotics by General Practitioners in South Africa: Observational, Analytic, Cross-Sectional Study. *Antibiotics*, *11*(10), 1423.
   <a href="https://doi.org/10.3390/antibiotics11101423">https://doi.org/10.3390/antibiotics11101423</a>
- Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PloS one*, *10*(9), e0138237. <u>https://doi.org/10.1371/journal.pone.0138237</u>
- Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano J., M. Lagos, P.
  Norris, E. Ponarin & B. Puranen et al. (eds.). 2020. World Values Survey: Round Seven
   Country-Pooled Datafile. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. <u>doi:10.14281/18241.20</u>
- Haldevang, M. de. (2016, September 1). *Why do we still use the term "sub-saharan africa"?* Quartz. <u>https://qz.com/africa/770350/why-do-we-still-say-subsaharan-africa</u>

- Hall, L. H., Johnson, J., Watt, I., Tsipa, A., & O'Connor, D. B. (2016). Healthcare staff wellbeing, burnout, and patient safety: a systematic review. *PloS one*, *11*(7), e0159015.
   <a href="https://doi.org/10.1371/journal.pone.0159015">https://doi.org/10.1371/journal.pone.0159015</a>
- Hansen, S. E., Benson, G., Bowers, A., Pennell, B. E., Lin, Y., cJ, D. B., & Hibben, K. C. (2016).
   Survey quality. guidelines for best practice in cross-cultural surveys. Ann Arbor, MI:
   Survey Research Center, Institute for Social Research, University of Michigan.

Harzing, A.-W. (2006). Response Styles in Cross-national Survey Research: A 26-country Study. International Journal of Cross Cultural Management, 6(2), 243–266. <u>https://doi.org/10.1177/1470595806066332</u>

- Harzing, A. W., Brown, M., Köster, K., & Zhao, S. (2011). Response Style Differences in Cross-National Research. <u>DOI:10.1007/s11575-011-0111-2</u>
- Hawkins, M., Cheng, C., Elsworth, G. R., & Osborne, R. H. (2020). Translation method is validity evidence for construct equivalence: analysis of secondary data routinely collected during translations of the Health Literacy Questionnaire (HLQ). BMC medical research methodology, 20(1), 1-13. <u>https://doi.org/10.1186/s12874-020-00962-8</u>
- Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. Psychological assessment, 7(3), 238. <u>https://doi.org/10.1037/1040-3590.7.3.238</u>
- He, J., Van de Vijver, F. J. R., Fetvadjiev, V. H., de Carmen Dominguez Espinosa, A., Adams, B., Alonso–Arbiol, I., Aydinli–Karakulak, A., Buzea, C., Dimitrova, R., Fortin, A., Hapunda, G., Ma, S., Sargautyte, R., Sim, S., Schachner, M. K., Suryani, A., Zeinoun, P., & Zhang, R. (2017). On Enhancing the Cross–Cultural Comparability of Likert–Scale Personality and Value Measures: A Comparison of Common Procedures. *European Journal of Personality*, *31*(6), 642-657. <u>https://doi.org/10.1002/per.2132</u>

- Heine, S. J., Lehman, D. R., Peng, K. & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *J Pers Soc Psychol*, 82(6):903-18. <u>PMID: 12051579.</u>
- Henrich, J., Heine, S. & Norenzayan, A. Most people are not WEIRD. *Nature* **466**, 29 (2010). <u>https://doi.org/10.1038/466029a</u>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? Behavioral and Brain Sciences, 33(2–3), 61–83. <u>https://doi.org/10.1017/S0140525X0999152X</u>
- Higgins J. P. T. & Green S. (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration. Available from <u>www.cochrane-handbook.org</u>.
- Holmes, A. H., Moore, L. S. P., Sundsfjord, A., Steinbakk, M., Regmi, S., Karkey, A., Guerin, P. J., & Piddock, L. J. V. (2016). Understanding the mechanisms and drivers of antimicrobial resistance. *Lancet (London, England)*, *387*(10014), 176–187. <u>https://doi.org/10.1016/S0140-6736(15)00473-0</u>
- Hulscher, M. E. J. L., & Prins, J. M. (2017). Antibiotic stewardship: does it work in hospital practice? A review of the evidence base. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 23(11), 799–805. DOI: <u>10.1016/j.cmi.2017.07.017</u>
- Hruschka, D. J., Munira, S., Jesmin, K., Hackman, J., & Tiokhin, L. (2018). Learning from failures of protocol in cross-cultural research. *Proceedings of the National Academy of Sciences*, *115*(45), 11428-11434. <u>https://doi.org/10.1073/pnas.1721166115</u>
- Hruschka, D. J. (2020). What we look with" is as important as "What we look at. *Evolution and Human Behavior*, *41*(5), 458-459. <u>10.1016/j.evolhumbehav.2020.07.011</u>
- Hwang, K. K. (2023). An Epistemological Strategy for Initiating Scientific Revolution against WEIRD Psychology. Integr. psych. behav. 57, 361–380. <u>https://doi.org/10.1007/s12124-022-09681-9</u>

- Iwu, C. D., & Patrick, S. M. (2021). An insight into the implementation of the global action plan on antimicrobial resistance in the WHO African region: A roadmap for action. *International Journal of Antimicrobial Agents*, *58*(4), 106411. <u>https://doi.org/10.1016/j.ijantimicag.2021.106411</u>
- Jenkins, N., Bloor, M., Fischer, J., Berney, L., & Neale, J. (2010). Putting it in context: the use of vignettes in qualitative interviewing. Qualitative research, 10(2), 175-198. <u>https://doi.org/10.1177/1468794109356737</u>
- Jindal, A. K., Pandya, K., & Khan, I. D. (2015). Antimicrobial resistance: A public health challenge. *Medical journal armed forces India*, 71(2), 178-181. <u>https://doi.org/10.1016/j.mjafi.2014.04.011</u>
- Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The Relation Between Culture and Response Styles: Evidence From 19 Countries. *Journal of Cross-Cultural Psychology*, 36(2), 264-277. <u>https://doi.org/10.1177/0022022104272905</u>
- Johnson, T., Shavitt, S., & Holbrook, A. (2010). Survey Response Styles Across Cultures. In D. Matsumoto & F. Van de Vijver (Eds.), *Cross-Cultural Research Methods in Psychology* (Culture and Psychology, pp. 130-176). Cambridge: Cambridge University Press. <u>doi:10.1017/CBO9780511779381.008</u>
- Johnson, B. B., & Swedlow, B. (2021). Cultural theory's contributions to risk analysis: A thematic review with directions and resources for further research. *Risk analysis*, *41*(3), 429-455. https://doi.org/10.1111/risa.13299
- Jürges, H. (2007). True health vs response styles: exploring cross-country differences in selfreported health. *Health Economics*, *16*(2), 163–178. <u>https://doi.org/10.1002/hec.1134</u>
- Kabba, J. A., Tadesse, N., James, P. B., Kallon, H., Kitchen, C., Atif, N., Jiang, M., Hayat, K.,
   Zhao, M., Yang, C., Chang, J., & Fang, Y. (2020). Knowledge, attitude and antibiotic prescribing patterns of medical doctors providing free healthcare in the outpatient departments of public hospitals in Sierra Leone: a national cross-sectional study.

*Transactions of the Royal Society of Tropical Medicine and Hygiene*, *114*(6), 448–458. <u>https://doi.org/10.1093/trstmh/trz137</u>

- Kalungia, A. C., Mwambula, H., Munkombwe, D., Marshall, S., Schellack, N., May, C., Jones, A. S. C., & Godman, B. (2019). Antimicrobial stewardship knowledge and perception among physicians and pharmacists at leading tertiary teaching hospitals in Zambia: implications for future policy and practice. *Journal of Chemotherapy (Florence, Italy)*, 31(7–8), 378–387. <u>https://doi.org/10.1080/1120009X.2019.1622293</u>
- Khan, M. U., Hassali, M. A. A., Ahmad, A., Elkalmi, R. M., Zaidi, S. T. R., & Dhingra, S. (2016).
   Perceptions and Practices of Community Pharmacists towards Antimicrobial
   Stewardship in the State of Selangor, Malaysia. *PLOS ONE*, *11*(2), e0149623.
   <a href="https://doi.org/10.1371/journal.pone.0149623">https://doi.org/10.1371/journal.pone.0149623</a>
- Khoiry, Q. A., Alfian, S. D., Van Boven, J. F., & Abdulah, R. (2023). Self-reported medication adherence instruments and their applicability in low-middle income countries: a scoping review. Frontiers in public health, 11, 1104510.
- Kim, H. S., Sherman, D. K., & Taylor, S. E. (2008). Culture and social support. American psychologist, 63(6), 518. DOI: <u>10.1037/0003-066X</u>
- Kimbowa, I. M., Eriksen, J., Nakafeero, M., Obua, C., Lundborg, C. S., Kalyango, J., & Ocan, M. (2022). Antimicrobial stewardship: Attitudes and practices of healthcare providers in selected health facilities in Uganda. *PLOS ONE*, *17*(2), e0262993. <u>https://doi.org/10.1371/journal.pone.0262993</u>
- Klimek, L., Bergmann, K. C., Biedermann, T., Bousquet, J., Hellings, P., Jung, K., Merk, H.,
  Schelnter, W., Stock, P., Ring, J., Wagenmann, M., Wehrmann, W., Mosges, R. & Pfaar,
  O. (2017). Visual analogue scales (VAS): Measuring instruments for the documentation of symptoms and therapy monitoring in cases of allergic rhinitis in everyday health care:
  Position Paper of the German Society of Allergology (AeDA) and the German Society of Allergy and Clinical Immunology (DGAKI), ENT Section, in collaboration with the working group on Clinical Immunology, Allergology and Environmental Medicine of the German
Society of Otorhinolaryngology, Head and Neck Surgery (DGHNOKHC). Allergo journal international, 26, 16-24. <u>https://doi.org/10.1007/s40629-016-0006-7</u>

- Krosnick, J. A., Presser, S., Fealing, K. H., Ruggles, S., & Vannette, D. L. (n.d.). *The National Science Foundation Advisory Committee for the Social, Behavioral and Economic Sciences Subcommittee on Advancing SBE Survey Research:* 163.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. Quality & Quantity, 47(4), 2025–2047. <u>https://doi.org/10.1007/s11135-011-9640-</u> <u>9</u>
- Kumwenda, S., El Hadji, A. N., Orondo, P. W., William, P., Oyinlola, L., Bongo, G. N., &Chiwona, B. (2017). Challenges facing young African scientists in their research careers:A qualitative exploratory study. *Malawi Medical Journal*, *29*(1), 1-4.
- Kusmaryono, I., Wijayanti, D., Maharani, H. R. (2022). Number of response options, reliability, validity, and potential bias in the use of the likert scale education and social science research: a literature review. *Int J Educ Methodol,* 8(4):625-637. <u>doi:</u> <u>10.12973/ijem.8.4.625</u>
- Lages, C. R., Pfajfar, G., & Shoham, A. (2015). Challenges in conducting and publishing research on the Middle East and Africa in leading journals. *International Marketing Review*, 32(1), 52–77. <u>https://doi.org/10.1108/IMR-12-2014-0374</u>
- Latkin, C. A., Dayton, L., Moran, M., Strickland, J. C., & Collins, K. (2021). Behavioral and psychosocial factors associated with COVID-19 scepticism in the United States. Current Psychology, 1-9. DOI: <u>10.1007/s12144-020-01211-3</u>
- Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, X. E. (2002). Cultural differences in responses to a likert scale. *Research in Nursing & Health*, *25*(4), 295–306. <u>https://doi.org/10.1002/nur.10041</u>
- Leggett, L. E., Khadaroo, R. G., Holroyd-Leduc, J., Lorenzetti, D. L., Hanson, H., Wagg, A., ... & Clement, F. (2016). Measuring resource utilization: a systematic review of validated self-

reported questionnaires. *Medicine*, *95*(10). https://doi.org/10.1097%2FMD.000000000002759

- Lencucha, R., & Neupane, S. (2022). The use, misuse and overuse of the 'low-income and middle-income countries' category. *BMJ global health*, *7*(6), e009067. <u>http://dx.doi.org/10.1136/bmigh-2022-009067</u>
- Liebe, U., Moumouni, I. M., Bigler, C., Ingabire, C., & Bieri, S. (2020). Using factorial survey experiments to measure attitudes, social norms, and fairness concerns in developing countries. Sociological methods & research, 49(1), 161-192. <u>https://doi.org/10.1177/0049124117729707</u>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22, 140*, 55–55.
- Liu, M., Suzer-Gurtekin, Z. T., Keusch, F., & Lee, S. (2018). Response Styles in Cross-cultural Surveys. In Advances in Comparative Survey Methods (pp. 477–499). John Wiley & Sons, Ltd. <u>https://doi.org/10.1002/9781118884997.ch22</u>
- Lopez-Roig, S., & Pastor, M.A. (2016). *Cultural adaptation of measures. In: Benyamini, Y.,* Johnston, M., & Karademas E.C. Assessment in Health Psychology. Hogrefe.
- Lotfi, T., Morsi, R. Z., Rajabbik, M. H., Alkhaled, L., Kahale, L., Nass, H., Brax, H., Fadlallah, R., & Akl, E. A. (2016). Knowledge, beliefs and attitudes of physicians in low and middle-income countries regarding interacting with pharmaceutical companies: A systematic review. *BMC Health Services Research*, *16*(1), 57. <u>https://doi.org/10.1186/s12913-016-1299-4</u>
- Löwe, B., & Kohlmann, S. (2023). International comparability of depression scores from selfreport scales: opportunities and challenges. *The Lancet Regional Health - Europe*, *31*, 100668. <u>https://doi.org/10.1016/j.lanepe.2023.100668</u>
- Luetke Lanfer, H., Reifegerste, D., & Kargbo, S. I. (2021). Collecting quantitative experimental data from a non-WEIRD population: challenges and practical recommendations from a

field experiment in rural Sierra Leone. *BMC research notes*, *14*(1), 1-6. <u>https://doi.org/10.1186/s13104-021-05828-w</u>

- Maldonado, B. N., Chandna, J., & Gladstone, M. (2019). A systematic review of tools used to screen and assess for externalising behaviour symptoms in low and middle income settings. Global Mental Health, 6, e13. doi:10.1017/gmh.2019.11
- Manga, M. M., Mohammed, Y., Suleiman, S., Fowotade, A., Yunusa-Kaltungo, Z., Usman, M., Abulfathi, A. A., & Saddiq, M. I. (2021). Antibiotic prescribing habits among primary healthcare workers in Northern Nigeria: a concern for patient safety in the era of global antimicrobial resistance. *PAMJ - One Health*, 5. <u>https://eprints.whiterose.ac.uk/186528/</u>

Marks, D. F. (2005). Health psychology: Theory, research and practice. Sage.

Marks, D. F., & Yardley, L. (2004). Research methods for clinical and health psychology. Sage.

- Martinez, R.G., Lewis, C.C. & Weiner, B.J. (2014). Instrumentation issues in implementation science. *Implementation Sci* **9**, 118. <u>https://doi.org/10.1186/s13012-014-0118-8</u>
- Martinková, P., Drabinová, A., Liaw, Y. L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking equity: Why differential item functioning analysis should be a routine part of developing conceptual assessments. CBE—Life Sciences Education, 16(2), rm2. doi:10.1187/cbe.16-10-0307
- Matsumoto, D., & Yoo, S. H. (2006). Toward a New Generation of Cross-Cultural Research. *Perspectives on Psychological Science*, *1*(3), 234-250. <u>https://doi.org/10.1111/j.1745-6916.2006.00014.x</u>
- McGrady, M. E., Pai, A. L., & Prosser, L. A. (2021). Using discrete choice experiments to develop and deliver patient-centered psychological interventions: a systematic review. Health psychology review, 15(2), 314-332.
- Médicins du Monde. (2011). The KAP Survey Model (Knowledge, Attitudes, and Practices). Retrieved from: <u>http://www.medecinsdumonde.org/Outils/Nous-contacter</u>

- Mersha, A. G. (2018). Attitude and perception of medical interns about antimicrobial resistance: a multi center cross-sectional study in Ethiopia. *Antimicrobial Resistance & Infection Control*, 7(1), 149. <u>https://doi.org/10.1186/s13756-018-0443-9</u>
- Mesquita, B., & Walker, R. (2003). Cultural differences in emotions: A context for interpreting emotional experiences. *Behaviour research and therapy*, *41*(7), 777-793. <u>https://doi.org/10.1016/S0005-7967(02)00189-4</u>
- Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D. & Walker, A. (2005)
  Psychological Theory Group. Making psychological theory useful for implementing
  evidence based practice: a consensus approach. Qual Saf Health Care. Feb;14(1):26-33. doi: 10.1136/qshc.2004.011155.
- Michie, S., Johnston, M., Francis, J., Hardeman, W., & Eccles, M. (2008). From Theory to Intervention: Mapping Theoretically Derived Behavioural Determinants to Behaviour Change Techniques. *Applied Psychology*, *57*(4), 660–680. <u>https://doi.org/10.1111/j.1464-0597.2008.00341.x</u>
- Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(1), 42. <u>https://doi.org/10.1186/1748-5908-6-42</u>
- Minkov, M., & Hofstede, G. (2012). Hofstede's fifth dimension: New evidence from the World Values Survey. *Journal of cross-cultural psychology*, *43*(1), 3-14. <u>https://doi.org/10.1177/0022022110388567</u>
- Mohler, P., Dorer, B., de Jong, J., & Hu, M. (2016). Translation. guidelines for best practice in cross-cultural surveys. *Ann Arbor, MI: survey research center. Institute for social research. University of Michigan.*
- Mughogho, W., Adhiambo, J., & Forscher, P. S. (2023). African researchers must be full participants in behavioural science research. *Nature Human Behaviour*, *7*(3), 297-299. https://doi.org/10.1038/s41562-023-01536-6

- Mufwambi, W., Stingl, J., Masimirembwa, C., Manasa, J., Nhachi, C., Stadler, N., ... & Munkombwe, D. (2021). Healthcare professionals' knowledge of pharmacogenetics and attitudes towards antimicrobial utilization in Zambia: implications for a precision medicine approach to reducing antimicrobial resistance. *Frontiers in Pharmacology*, *11*, 551522. <u>https://doi.org/10.3389/fphar.2020.551522</u>
- Muniz, J., Elosua, P., & Hambleton, R. K. (2013). International Test Commission Guidelines for test translation and adaptation. *Psicothema*, 25(2), 151-157. <u>https://doi.org/10.7334/psicothema2013.24</u>
- National Heart Lung and Blood Institute. (2014). Quality Assessment Tool for Before-After (Pre-Post) Studies With No Control Group [National Heart Lung and Blood Institute web site]. Retrieved from: http://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascularrisk-reduction/tools/before-after
- Nielsen, J. Ø., Shrestha, A. D., Neupane, D., & Kallestrup, P. (2017). Non-adherence to antihypertensive medication in low- and middle-income countries: A systematic review and meta-analysis of 92443 subjects. *Journal of Human Hypertension*, 31(1), 14–21. <u>https://doi.org/10.1038/jhh.2016.31</u>
- Nkinda, L., Kilonzi, M., Felix, F. F., Mutagonda, R., Myemba, D. T., Mwakawanga, D. L.,
  Kibwana, U., Njiro, B. J., Ndumwa, H. P., Mwakalukwa, R., Makuka, G., Kubigwa, S. W.,
  Marealle, A. I., Mikomangwa, W. P., Sambayi, G., Kunambi, P. P., Maganda, B. A., Sirili,
  N., Mfaume, R., Nshau, A.B., Bwire, G.M., Scherpbier, R., Nyankesha, E. &
  Ndayishimiye, P. (2022). Drivers of irrational use of antibiotics among children: a mixedmethod study among prescribers and dispensers in Tanzania. *BMC Health Services Research*, *22*(1), 961. https://doi.org/10.1186/s12913-022-08359-7
- O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, 19. <u>https://doi.org/10.1177/1609406919899220</u>

- Odukoya, O., Fox, R. S., Hayman, L. L., & Penedo, F. J. (2021). The International Society for Behavioral Medicine (ISBM) and the Society of Behavioral Medicine (SBM) advocate for the inclusion of behavioral scientists in the implementation of the Global Action Plan for the prevention of non-communicable diseases (NCDs) in low-and middle-income countries. *Translational behavioral medicine*, *11*(6), 1286-1288. <u>https://doi.org/10.1093/tbm/ibaa128</u>
- Ogoina, D., Iliyasu, G., Kwaghe, V., Otu, A., Akase, I. E., Adekanmbi, O., Mahmood, D., Iroezindu, M., Aliyu, S., Oyeyemi, A. S., Rotifa, S., Adeiza, M. A., Unigwe, U. S., Mmerem, J. I., Dayyab, F. M., Habib, Z. G., Otokpa, D., Effa, E., & Habib, A. G. (2021). Predictors of antibiotic prescriptions: a knowledge, attitude and practice survey among physicians in tertiary hospitals in Nigeria. *Antimicrobial Resistance & Infection Control*, *10*(1), 73. <u>https://doi.org/10.1186/s13756-021-00940-9</u>
- Olufadewa, I. I., Adesina, M. A., & Ayorinde, T. (2020). From Africa to the World: Reimagining Africa's research capacity and culture in the global knowledge economy. *Journal of Global Health*, *10*(1). <u>10.7189/jogh.10.010321</u>
- Ogrinc, G., Davies, L., Goodman, D., Batalden, P., Davidoff, F., & Stevens, D.. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. BMJ Qual Saf. 2016 Dec;25(12):986-992. <u>https://doi.org/10.1136/bmjqs-2015-004411</u>

O'Neill, J. (2016). Tackling drug-resistant infections globally: final report and recommendations.

Osterlind, S. J., & Everson, H., T. (2009). *Differential Item Functioning*. Sage.

- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery*, *88*, 105906. <u>https://doi.org/10.1016/j.ijsu.2021.105906</u>
- Pelham, B., Hardin, C., Murray, D., Shimizu, M., & Vandello, J. (2022). A truly global, non-WEIRD examination of collectivism: The global collectivism index (GCI). *Current*

Research in Ecological and Social Psychology, 3, 100030. https://doi.org/10.1016/j.cresp.2021.100030

- Peters, G. J., & Crutzen, R. (2024). Knowing what we're talking about: facilitating decentralized, unequivocal publication of and reference to psychological construct definitions and instructions. Meta-Psychology, 8. <u>https://doi.org/10.15626/MP.2022.3638</u>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, *88*(5), 879. <u>https://psycnet.apa.org/doi/10.1037/0021-9010.88.5.879</u>
- Prestinaci, F., Pezzotti, P., & Pantosti, A. (2015). Antimicrobial resistance: a global multifaceted phenomenon. *Pathogens and Global Health*, *109*(7), 309–318. <u>https://doi.org/10.1179/2047773215Y.000000030</u>
- Patient Reported Outcomes Measurement Information System (PROMIS). (May, 2013). Instrument Development and Validation Scientific Standards Version 2.0. <u>https://www.healthmeasures.net/images/PROMIS/PROMISStandards\_Vers2.0\_Final.pdf</u>
- Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences*, *115*(45), 11401-11405. <u>https://doi.org/10.1073/pnas.172116511</u>
- Rogers Van Katwyk, S., Hoffman, S. J., Mendelson, M., Taljaard, M., & Grimshaw, J. M. (2020).
   Strengthening the science of addressing antimicrobial resistance: a framework for planning, conducting and disseminating antimicrobial resistance intervention research.
   *Health Research Policy and Systems*, *18*(1), 60. <u>https://doi.org/10.1186/s12961-020-00549-1</u>
- Ryba, T. V., & Schinke, R. J. (2009). Methodology as a ritualized Eurocentrism: Introduction to the special issue. International Journal of Sport and Exercise Psychology, 7(3), 263– 274. <u>https://doi.org/10.1080/1612197X.2009.9671909</u>

- Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., De Graeff, A., Groenvold, M., Peterson, M. A. & EORTC and the Quality of Life Cross-Cultural Meta-Analysis Group. (2007). The use of differential item functioning analyses to identify cultural differences in responses to the EORTC QLQ-C30. *Quality of Life Research*, *16*, 115-129. <u>https://doi.org/10.1007/s11136-006-9120-1</u>
- Scott, N.W., Fayers, P.M., Aaronson, N.K. *et al.* (2010). Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. *Health Qual Life Outcomes* 8, 81. <u>https://doi.org/10.1186/1477-7525-8-81</u>
- Singh, L. (2022). Navigating equity and justice in international collaborations. *Nature Reviews Psychology*, *1*(7), 372-373. <u>https://doi.org/10.1038/s44159-022-00077-5</u>
- Sneddon, J., Cooper, L., Afriyie, D. K., Sefah, I. A., Cockburn, A., Kerr, F., Cameron, E., Goldthorpe, J., Kurdi, A., & Seaton, R. A. (2020). Supporting antimicrobial stewardship in Ghana: evaluation of the impact of training on knowledge and attitudes of healthcare professionals in two hospitals. *JAC-Antimicrobial Resistance*, 2(4), dlaa092. <u>https://doi.org/10.1093/jacamr/dlaa092</u>
- Squires, J. E., Estabrooks, C. A., O'Rourke, H. M., Gustavsson, P., Newburn-Cook, C. V., & Wallin, L. (2011). A systematic review of the psychometric properties of self-report research utilization measures used in healthcare. *Implementation Science*, 6(1), 83. <u>https://doi.org/10.1186/1748-5908-6-83</u>
- Survey Research Center. (2016). Guidelines for Best Practice in Cross-Cultural Surveys. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. <u>http://ccsg.isr.umich.edu/</u>
- Tafa, B., Endale, A., & Bekele, D. (2017). Paramedical staffs knowledge and attitudes towards antimicrobial resistance in Dire Dawa, Ethiopia: a cross sectional study. *Annals of Clinical Microbiology and Antimicrobials*, *16*(1), 64. <u>https://doi.org/10.1186/s12941-017-0241-x</u>

- Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., Horner, R., Kazdin, A., Kratochwill, T., McDonald, S., Sampson, M., Shamseer, L., Togher, L., Albin, R., Backman, C., Douglas, J., Evans, J. J., Gast, D., Manolov, R., Mitchell, G., Nickels, L., Ownsworth, T., Rose, M., Schmid, C. H. & Wilson, B. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *Physical Therapy*, *96*(7), e1–e10. <u>https://doi.org/10.2522/ptj.2016.96.7.e1</u>
- Tanujaya, B., Prahmana, R. C. I., & Mumu, J. (2022). Likert scale in social sciences research: Problems and difficulties.
- Tembo, N., Mudenda, S., Banda, M., Chileshe, M., & Matafwali, S. (2022). Knowledge, attitudes and practices on antimicrobial resistance among pharmacy personnel and nurses at a tertiary hospital in Ndola, Zambia: implications for antimicrobial stewardship programmes. JAC-Antimicrobial Resistance, 4(5), dlac107. <u>https://doi.org/10.1093/jacamr/dlac107</u>
- Teresi, J. A., & Fleishman, J. A. (2007). Differential item functioning and health assessment. *Quality of Life Research*, *16*, 33-42. <u>https://doi.org/10.1007/s11136-007-9184-6</u>
- Tesch-Römer, C., Motel-Klingebiel, A., & Tomasik, M. J. (2008). Gender differences in subjective well-being: Comparing societies with respect to gender equality. Social Indicators Research, 85, 329-349.
- Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American?. *American Psychologist*, *76*(1), 116. <u>https://doi.org/10.1037/amp0000622</u>
- Turner, R. R., Epuru Reddy, S., Byrne-Davis, L. M., Bull, E. R., & Hart, J. (2023). An interview study to explore applied psychologists' experiences of implementing health psychology in global health partnerships: The Change Exchange. *British Journal of Health Psychology*. DOI: <u>10.1111/bjhp.12669</u>

- United Nations. (2022). World Population Prospects 2022. https://desapublications.un.org/file/989/download
- Van de Vijver, F., He, J., & Kulikova, A. (2017, June). Response styles in cross-cultural perspective: TIMSS and PIRLS. In *Keynote presented at the 7th IEA International Research Conference 2017, Prague, Czech Republic.*
- Van Dijck, C., Vlieghe, E., & Cox, J. A. (2018). Antibiotic stewardship interventions in hospitals in low-and middle-income countries: a systematic review. *Bulletin of the World Health Organization*, 96(4), 266–280. <u>https://doi.org/10.2471/BLT.17.203448</u>
- Voss, K. E., Stem, D. E., Johnson, L. W., & Arce, C. (1996). An exploration of the comparability of semantic adjectives in three languages: A magnitude estimation approach. *International Marketing Review*, *13*(5), 44–58. <u>https://doi.org/10.1108/02651339610131388</u>
- Wan, T. T. H., Rav-Marathe, K., & Marathe, S. (2016). A Systematic Review of KAP-O Framework for Diabetes. *Medical Research Archives*, *3*(9). https://esmed.org/MRA/mra/article/view/483
- Wernli, D., Jørgensen, P. S., Parmley, E. J., Troell, M., Majowicz, S., Harbarth, S., Léger, A., Lambraki, I., Graells, T., Henriksson, P. J. G., Carson, C., Cousins, M., Skoog Ståhlgren, G., Mohan, C. V., Simpson, A. J. H., Wieland, B., Pedersen, K., Schneider, A., Chandy, S. J., Wijayathilaka, M. V. M, Delamare-Deboutteville, J., Vila, J., Lundborg, C. S., & Pittet, D. (2020). Evidence for action: a One Health learning platform on interventions to tackle antimicrobial resistance. *The Lancet. Infectious Diseases*, *20*(12), e307–e311. <a href="https://doi.org/10.1016/S1473-3099(20)30392-3">https://doi.org/10.1016/S1473-3099(20)30392-3</a>
- Willits, F. K., Theodori, G. L., & Luloff, A. E. (2016). Another look at Likert scales. *Journal of Rural Social Sciences*, *31*(3), 6.
- Wilson Fadiji, A., Khumalo, I. P., & Zulu, N. T. (2022). Well-Being in Africa: Towards an Africa (n) Centred Positive Psychology. In *Embracing Well-Being in Diverse African Contexts:*

Research Perspectives (pp. 31-49). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-85924-4\_2

- World Bank. (2021). World Development Indicators 2021. World Bank. <u>https://databank.worldbank.org/source/world-development-indicators</u>
- World Health Organisation (WHO). (2011). WHO Global Strategy for Containment of Antimicrobial Resistance. Retrieved from: <u>https://www.who.int/publications/i/item/who-global-strategy-for-containment-of-antimicrobial-resistance</u>
- Yang, Y., Harkness, J. A., Chin, T. Y., & Villar, A. (2010). Response styles and culture. Survey methods in multinational, multiregional, and multicultural contexts, 203-223.
- Zumbo, B. D. (1999). A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.