Identifying the core components and items to measure health workers' cultural competence in the Ethiopian context

Belay Moges¹, Tesfaye Semela²

¹School of Teacher Education, College of Education, Hawassa University, Hawassa, Sidama, 05, Ethiopia
²School of Teacher Education, College of Education, Institute of Policy and Development Research, Hawassa University, Hawassa, Sidama, 05, Ethiopia

Abstract

**Background:** Cultural competence (CC) is a crucial attribute in attaining quality healthcare service outcomes, mitigating malfunctioning practices, and improving patient satisfaction. Studies suggested comprehensive CC assessment requires appraising existing CC tools to measure health workers' CC in the Ethiopian context. Selecting existing CC tools, identifying sub-constructs, pinpointing demographic characteristics, and evaluating items are the study objectives.

**Methods:** 20 cultural competence tools of 20 to 83 items, 1-5 sub-constructs, and 4-10 Likert-Type rating options were identified, to be rated by eight experts in three groups. Consensus based Standards for the selection of health Measurement INstruments (COSMIN) and test construction literature are used to develop raters rating codes to compute Inter-Raters Reliability (IRR). The first group of three experts was rated to decide the inclusion of CC tools, factors, and demographic information. The second group of three experts selected six CC tools and 65 items. Two experts in the third group further evaluated the selected items.

**Results:** Reliability for the inclusion of CC tools, factors, and demographic variables were found to be 75%–87%, 50%–93%, and 50%–86% respectively. 13 items that violate test construction principles like absoluteness, endorsement, proneness to multiple interpretation, ambiguity and double barring were excluded. Cultural skill, cultural knowledge, and cultural awareness are the three most common sub-constructs. Cultural sensitivity, cultural desire, and cultural encounters are the least incorporated factors. Gender, age, language, department, patients encounter, education, and cultural diversity training are demographic variables to assess health workers’
Conclusions: Items of CC instruments cover cognitive, affective, and skill domains as prominent dimensions of CC tools. The reviewed tools consist of items that violate test construction principles. To conduct CC assessment in the Ethiopian context, apart from refining and validating selected items, generating additional items is recommended.

Keywords
Cultural Competence, Cultural Skill, Cultural Knowledge, Cultural Awareness, Cultural Sensitivity, Cultural Desire, Cultural Encounter, Inter-Raters Reliability

This article is included in the Sociology of Health gateway.

This article is included in the Health Services gateway.
Introduction

In recent decades, cultural competence (CC) is increasingly becoming a subject of scholarly attention globally with growing commitments to address the needs of culturally diverse clients. Being a key service sector, the provision of quality health care, by and large depends on the ability to effectively handle cultural diversity that encompasses patients’ variations of gender, ethnicity, education, geography, language, color, etc. Health is a key service sector that demands culturally competent health providers and health workers in order to guarantee better patient satisfaction. Countries should work to equip health professionals with the required CC knowledge, skills, and attitudes to address the needs of culturally diverse patients. Explorations of the current cultural situation and assessments of CC can help health researchers and policymakers to understand major constrictions and reduce health workers’ bias. CC is critical for providing appropriate health services for countries with culturally diverse populations like Ethiopia. However, understanding health workers’ and providers’ CC status is hampered by lack of conceptualizations and consistent measuring tools.

CC conceptualization gaps are evidenced in using different terms to detonate it and lack of a single agreed-upon definition. The concept transcultural competence or cultural competence were first coined by Pratt in 1957, interchangeably used terms such as cultural sensitivity, cultural-bound health care needs, cultural-specific care needs, transcultural, multiculturalism, and cross-cultural competence. CC refers to a trait where a person can coordinate, work, or interact with other people who are of different cultures and social backgrounds. CC has become an unquestionable and ubiquitous aspect of health professional development and linked to efforts to eliminate culture-based health disparities through sympathetic health workers training.

A growing body of literature recognizes that the health workforce or providers’ CC assessment become a priority of health research, education, and practice in countries where health workers encounter culturally diverse patients to improve healthcare outcome quality. To realize the benefit of CC assessment, scholars develop, adapt, and validate CC measuring tools for nursing students; nurses active at work; rehabilitation practitioners; professionals living and working in multicultural settings, and healthcare providers. Scholars rationales for the need of CC tools are increasing trend of foreign patients visiting homogeneous society of South Korean, increasing trend of nurse expatriates in Iran, and proven existence of ethnic diversity in Vietnam.

However, common limitations of health professionals’ CC research and assessment worldwide include: lack of clarity on CC concepts; scarcity of outcome-based research on effectiveness of CC strategies; scant numbers of standardized CC instruments in health cultural context of countries; and violation of one or more test construction principles. Beyond aforementioned gaps, Ethiopian health policy gives limited attention to CC, and health providers overlooked the role of CC in health practices were evidenced as additional weakness. Therefore, CC assessments necessitate an urgent need of either developing or adapting valid and reliable CC tool for use in Ethiopian health services. Beforehand, it is imperative to conduct such a type of study targeted to identify major factors to be considered in CC tools. This study was conducted as the first phase of developing the Ethiopian CC instrument with objectives, including: (a) to identify CC measuring tools and demographic variables; (b) to detect sub-constructs of CC tools; (c) to select items to be included in CC tools; and (d) to evaluate items’ adherence to test construction principles.

The remaining part of this study is presented in seven sections. The context of health service provision in Ethiopia is presented in section two. Conceptualization and theoretical frameworks are presented in section three. Section four, five, six, and seven will present methodology, findings, discussion, and conclusion respectively.

Health service provision in the Ethiopian context

Ethiopia has a short history of modern health services and upholds its roots during the Emperor Regime simultaneously with the 1948 Ministry of Health establishment. The Emperor Regime and Ethiopian subsequent governments commonly have preventative, curative, and rehabilitative health goals. Ethiopia evidenced high population growth rates with 120 million people in 2022, estimated 250 million in 2050, urban (21%) - rural (79%) population variations, is the home of more than 80 ethnic groups, and people communicate using 90 languages of Semitic, Cushitic, Omotic, and Nilo-Saharan families. These cultural factors conglomerated with others have made Ethiopia a home of paramount cultural diversity that will call for the need for CC assessment in crucial service sectors.

So far, heart-breaking conflict, violence, displacement, and war have continued as curses in recent and past Ethiopia. This might be due to inadequacy to address CC in the policies and practices. These devolved situations have supported Huntington’s 1993 assertion; human conflict and violence in the new millennium will not be primarily generated on economic or ideological grounds but rather from the limitations of designing culturally responsive systems in crucial service sectors.
The health sector, as a key service sector, is not only limited in incorporating CC but the health policies and practices overlooked it.\(^6\) Or else, past and current governments of Ethiopia commonly worked massive health facility expansion and increasing number of health workers to meet the increasing health demands of rapidly growing population.\(^4\)\(^5\)\(^6\)\(^7\) However, Ethiopia is still known for very low densities of health workers at 0.96 per 1000 population far below the African density of health workers of 2.2/1000 population and five times less than World Health Organization’s (WHO) minimum threshold of 4.45 per 1000 population to meet the SDG health targets.\(^7\)\(^8\)\(^9\) According to Haileamlak’s study finding Ethiopia has less than 100,000 health workforce and the health workforce in 2022 is estimated to be less than 120,000.\(^10\) Upon this, Ethiopia should require more than 200,000 health workers to reach the density of African health workforce and 500,000 to fulfill Universal Health Care (UHC) minimal requirements. This means that in order to achieve UHC by 2030, Ethiopia expected to train more than 65,000 health workers (medical doctors, health officers, nurses, and midwives) per year for the next eight years.

However, the work to increase Ethiopian Health workers has increased variations on cultural factors such as language, ethnicity, food, clothing, religion, etc., both within their colleagues and patients.\(^11\) Alternatively, to minimize health gaps, scholars suggested that the Ethiopian government revise its health education curricula, training modality, trainee recruitment, and quality of teachers to go in line with CC international standards, parallel to health workers increasing tasks.\(^12\) Whereas, increasing diversity, conglomerated with limited numbers of health workers, unavailability of CC tools to assess health practices, limited CC attention in health policy, and education curricula have hampered delivering quality healthcare, obstructed closing health disparity, and prevented achieving national and global health goals.\(^13\)

Despite this, the international health education curriculum overwhelmingly incorporates CC issues and conducts a comprehensive assessment of health workers as prime priorities of health research, policy, education, and practice.\(^14\) Because, designing mechanisms to improve in-service training to equip health-workers with knowledge, attitude and skill of CC is helpful to deliver quality healthcare by reducing societal dissatisfaction. However, CC as one of mitigating strategies to Ethiopian health challenges, particularly gaps resulted from increasing variations of cultural factors, has gained little attention from both the past and current Ethiopian governments.

In order to bring high-quality health outcomes to satisfy health demands of culturally diverse patients, Ethiopia health system should design strategies to enhance health workers or providers CC, and conduct comprehensive CC assessments to understand existing health gaps as essential components of health practices have emerged as suggestions.\(^15\)\(^16\)\(^17\) According to studies, Ethiopia should emphasize CC education, on-the-job training, interpreter service, and comprehensive assessment of health workers to deliver culturally appropriate care.\(^18\)

However, beyond CC limitations evidenced worldwide; the Ethiopian health system has demonstrated additional gaps. For instance, nurses reported as if they do not hear the term cultural competence both in their schooling and practices, rather they accustomed to clinical competence.\(^19\) Another study evidenced a considerable numbers of Ethiopian people believed that modern health services showed unacceptable practice.\(^20\) A growing numbers of studies overseas recommend the need for CC assessment of health workers to address patients’ health demand.\(^21\)\(^22\)\(^23\)\(^24\)\(^25\) Not surprisingly, improving the CC of health professionals by formulating CC policy and conducting comprehensive CC assessment has become the most recurring theme in Ethiopian health. According to Bonsa (2018), it is critical to conduct nationally representative surveys of CC from patient, provider, and professional respondents. The cultural difference of western countries as low context culture and non-western countries as high context culture have limited the universal and consistence use of CC instruments to measure CC of health workers.\(^26\)\(^27\)\(^28\) For instance, the cultural belief of health or illness acceptable in one culture may not be acceptable in another culture.\(^29\) A study on the health or illness belief of Ethiopians evidenced that spiritual healing is an important treatment for illness, prefer injections to tablets, and prefer translators from the same ethnic group.\(^30\)

Additionally, Ethiopia’s proven existence of a high number of cultural factors, limited attention to CC in health practices, ambivalent or complex nature of CC, and inconsistencies in the rationales of CC tool construction overseas might make using a specific CC tool to assess the CC of Ethiopian health workers unsound.\(^31\)\(^32\)\(^33\)\(^34\) Therefore, the CC tool for use in Ethiopia should be comprehensive in nature, let alone other cultural factors, and it should address the existing marginalization resulted by prejudice or stereotypes within same culture. Marginalization of individuals by their perceived dominant counterparts within same ethnic group; body contacts with them are not allowed, and touching should be believed as bedraggled that heightens patients’ dissatisfaction on the healthcare service.

The proven existence of health disparity of Ethiopia resulted from such and other cultural factors have made imperative to conduct such types of study prior to developing, adapting and validating CC tools. Therefore, conducting this type of study is becoming mandatory not only to achieve the objectives but also crucial to use as a steppingstone for the
development of Ethiopian Health Workers CC scale. The next sections will discuss the conceptualization and theoretical frameworks of CC.

**Conceptualization and theoretical frameworks of CC**

There is no single agreed-upon definition of CC, continues as complex, ambivalent, and ambiguous to conceptualize it. Scholars prefer to denote it interchangeably using various terms; ‘culturally congruent’, ‘interculturality’, ‘transculturality’, and ‘crossculturality’. The major defining features and concepts deduced from these terms might not contradict with each other. For example, the terms commonly share the necessity to address the needs of clients who are essentially different from the one who provide services to them by one or more cultural factors. The aforementioned terms were used in this study to represent the constructs and concepts addressed in CC.

In spite of its lack of consensus, CC is conceptualized as a set of congruent behaviors, attitudes, and policies that come together to enable a system or among consumer providers and professionals to work effectively in cross-cultural situations. Again, Padopoulos et al., defines CC as the study of cultural differences and similarities in health and illness, underpinning societal and organizational structures to understand current health care practice and contribute to future development in a culturally responsive way. The ambiguity of the concept is manifested in the definition stated as mechanisms of giving attention to patients’ prior healthcare experiences in their country and abroad, to possible migration backgrounds and experience therein.

The differences in defining features may become sources of contradiction to use a single model or theory on various health cultural contexts. Therefore it is mandatory to identify most widely used definitions, and the Campinha-Bacote’s definition of CC, and its sub-constructs are found to be appropriate to use as operational definitions of this study.

Beyond the differences, the definitions commonly strive to achieve three core components of global health goals; health as human rights, attention to the entire world’s perspectives, and use of an interdisciplinary approach. Regardless of the continued ambiguity, scholars continued to develop, adapt, and evaluate CC measuring tools in the context of health using theories, models and conceptualizations.

Among the classic theories of culture and culturally appropriate health, the most notable one is Leininger’s theory; she worked as a clinical nurse specialist with worried children in the USA and noticed recurrent behavioral variations among children. It motivated her to formulate a hypothesis stating whether such variations have a cultural basis or not. To test it, she studied and conducted intensive research to understand the existing nature of care outside the Western context. She came to Africa, Papua New Guinea, and studied the culture of the country, and she later synthesized her grand theory, called the Cultural Care Theory (CCT).

CCT aims to equip health professionals with the capacity to care for culturally diverse people as crucial elements of the healthcare field. During the study, she gave emphasis to the transcultural capabilities of nurses in delivering culturally appropriate care to clients. She devoted considerable time not only developing the CCT but also worried about making use of it in most cultural contexts. Leininger’s philosophical roots of CCT include: extensive and diverse nursing experiences; anthropological insights; life experiences and values; and reflection upon spiritual insights and beliefs. Leininger represents her theory using sunrise model, symbolizes rising of the sun to represent care. Her CCT is depicted in Figure 1.

The figure depicts the four foci of the Sunrise Enabler as a cognitive map of the culture care theory; within the foci in the upper portion of the model are components of social structure and worldview factors that influence care and health. The figure depicts how health as a service made available to diverse clients should consider cultural variations manifested on individuals, families, groups, and communities when providing services, delivering health education, and designing curriculum.

Leininger’s theory primarily emphasizes culturally based factors such as technological, religious, and philosophical; kinship and social; and cultural values, beliefs, and life-way’s in the second foci. Leininger argues that delivering culturally congruent health care is not gained by any single factor, but rather by a complex interaction of social, organizational, and internal and external dimensions and factors of health providers and professionals.

The other model of CC is Giger and Davidhizar established Model of Transcultural Assessment and Intervention (MTAI) first published in 1988. The MTAI explores variations that exist in caregivers’ responses and perspectives relative to cultural diversity presented in a specific country. This theory can be used to examine the evaluation, decision-making, and implementation approaches of health professionals regarding CC in health.
Some improvement is observed in 2000 and 2005 in Purnell’s model of CC in recapping the limitations of MTAI and clarifying CCT’s three activities and decisions. The factors in Leininger were identified as 12 domains in Purnell’s model that include overview/heritage, communication, family roles and organizations, workforce issues, bio-cultural ecology, high-risk behaviors, nutrition, pregnancy and childbearing practices, spirituality, healthcare practices, and healthcare practitioners. The most widely used CC model is developed by Campinha-Bacote; making many improvements to her initial model that aimed to recapture the ambiguous nature of CC. Her model is called the Process of CC in the Delivery of Health Care Services. Her 1991 model incorporated a limited number of constructs such as cultural awareness (CA), cultural knowledge (CK), cultural skill (CSK), and cultural encounters (CE) without expressing their interdependence at the same time. She then decided to revise her model, adding a fifth construct called cultural sensitivity (CS) and changing the pictorial representation from linear to a set of circles. The Campinha-Bacote’s Model is depicted in Figure 2.

The five CC constructs in the model share more common features. Campinha-Bacote’s model assumptions include: CC is a process, not an event; there is more variation within ethnic groups than across ethnic groups (intra-ethnic variation);
levels of competence of health workers are related to their ability to provide culturally responsive health care services; and it consists of the aforementioned five sub-constructs or factors.9

For this study, the concept of CC and sub-constructs defined by Campinha-Bacote’s 2003 work were used as operational definitions.6 This is because the validation, adaptation, and development of most CC instruments were highly informed by the sub-constructs of CC identified by Campinha-Bacote’s model.30,79,80

Cultural competence (CC) is health workers’ ability to deliver culturally appropriate care to culturally diverse patients or the ability to synthesize and apply previously acquired cultural awareness (CA), cultural knowledge (CK), cultural skill (CSK), cultural encounter (CE), and cultural sensitivity (CS) to provide culturally appropriate patient care.

Cultural encounter (CE) is an ongoing process of interacting with patients from diverse cultural backgrounds. It aims to validate, refine, or modify existing values, beliefs, and practices of a cultural group.

Cultural desire (CD) is the motivation of the healthcare professional to ‘want to’ engage in the process of becoming culturally competent, not to ‘have to’.

Cultural awareness (CA) is deliberate self-examination and in-depth exploration of one’s biases, stereotypes, prejudices, assumptions, and ‘-isms’ that one holds about individuals and groups who are different from them.

Cultural knowledge (CK) is the process of acquiring a solid educational foundation about culturally and ethnically diverse groups.

Cultural skill (CSK) is the ability to collect culturally relevant data regarding the patient’s presenting problem as well as accurately perform a culturally-based physical assessment in a culturally sensitive manner.

**Methods**

**Ethical statement**

Hawassa University, College of Natural and Computational Sciences, Research Ethics Review Committee (RERC) has approved the implementation of the research on September 11, 2022, with reference number RERC 10010/22 (Letter of Research Ethics Review Committee). The Committee confirmed that the research met the criteria, which included minimizing subject risk, selecting subjects equitably, determining respect for person-informed consent, documenting it appropriately, and maintaining privacy and confidentiality. Written informed consent was obtained from all participants.

**Sources of information and search strategy**

We searched the term ‘review of CC measuring tools in health settings’ in the Google Scholar database for publications from November 2021 up to August 2022. The search was narrowed down again by searching the statement ‘health workers’ Cultural Competence Measuring Instruments: a systematic review’ within the first search. Then, 347 articles or documents published in English were downloaded for further evaluation to select relevant materials. Making screenings based on COSMIN checklist manual,81 three articles such as “measures of cultural competence in nurses: an integrative review”51; “psychometric properties of instruments used to measure the cultural competence of nurses: a systematic review”,24 and “cultural competence: analyzing the construct”25 were chosen for intensive review to further narrow down the selection of articles.

Based on the insights gained, full-text screening of studies related to CC in health settings and factor analysis of sub-constructs, models, and theories with target groups of health professionals, health workers, nurses, nursing students, physicians, and health providers were conducted. Articles on CC assessment, scale, tools, instruments, and questionnaires were also considered for inclusion in this study.

Finally, 103 articles or documents of two types in the databases of Academic Medicine, Sage Pub, Academia, Course Hero, Science Direct, Transcultural Care Net, Quiz Let, Semantic Scholar, Simply Psychology, Springer, and Research Gate which were available on Google Scholar and published between 1972 to 2022 were identified and reviewed.

The first type of papers, papers from overseas outside of Ethiopia, including: methodological materials; review of culture and CC; developing, adapting, and validating CC tools; and CC concepts, theories, and frameworks. The second type of paper, papers based in Ethiopia or about Ethiopian topics, including: sociolinguistics, population, and human resources; health policy, strategy, goals, diversity for outcome disparities, and standards; and sociocultural health beliefs, languages, and ethnicity.
Procedure
Eight carefully selected experts participated in this study. The experts were selected based on their relevant contributions to the issue under study through in-depth discussions held with college deans, department heads, and researchers. There were three experts in the first group; one physician with an MD degree, teaching health-related courses for health students at Dilla University and delivering health services at the university’s referral hospital for more than five years; one sociologist (PhD) with more than ten years of teaching experience at the same University; and one clinical nurse working at the university’s referral hospital for more than a decade. The two health workers of this group were selected as participants upon the recommendations of the health college dean of Dilla University. The third expert in this group is a sociologist, whom the authors chose after carefully evaluating the profiles of the teachers in the departments of social anthropology and sociology. Experts in this group served as judges’ in the selection and inclusion of CC tools, sub-con structs, and respondents’ personal information. Besides, these experts have had a chance to independently select participants from health workers to participate as experts in the second group of the study, i.e., the screening of items from their agreed-upon CC tools.

Upon the recommendations of the first group of experts, the other three experts in the second group included a physician working at Black Lion Hospital; a clinical nurse holding a master’s degree and who was working at Jemmo Health Post in Addis Ababa, and a lecturer in public health working at Jimma University. These experts have more than seven years of proven experience in the health setting of Ethiopia and participated in the screening and selection of items for further evaluation of CC instruments initially decided for inclusion by the first group of experts.

Following this, the third group were set up for the final screening. To select experts who participated in the third group, the authors have had in-depth discussions. They arrive at a consensus to select these experts from psychology department. Because courses related to item development, adaptation, and validation have been delivered to undergrad and postgrad students at psychology department. The group constituted two experts with MA degrees in psychology who specialize in measurement and evaluation that aimed to equip learners with the necessary knowledge, skill, and value for valid and reliable assessment instrument construction, appraisal, and administration with proven experience in teaching test construction and related courses. Both experts in the third group were lecturers at the Department of Psychology, College of Education, Dilla University.

For the purpose of obtaining their willingness to participate in the study, experts in the first and third groups were communicated with through both face-to-face contacts and telephone calls. These experts in both groups were communicated with for data gathering at Dilla University, Odaya Campus, in the lecture hall of the college of education and behavioral sciences at different time points. While experts in the second group were contacted via phone calls, data was gathered at the Office of Excellence Consultancy and Training Center (ECTC) in Addis Ababa. The willingness expressed during our contacts has been taken as their initial consent to participate in this study. Beyond the discussions held at different time points, the experts in the entire three groups have independently engaged in activities related to the tasks assigned in the respective groups of the study. Finally, the written consent of experts in the entire three groups has been obtained during the face-to-face contact made during data gathering. The data collections were held in Dilla and Addis Ababa starting from September 2022 up to November 2022. Figure 3 depicts the procedure followed to guide this study.

To guide the subsequent stages, rating criteria have designed based on review of prominent scholarly works on principles of scale development and item construction; review of articles on CC instruments; and the COSMIN checklist manual. The rating criteria were synthesized and guided by the prominent work of McAlister et al., on qualitative coding to assess Inter-Raters Reliability (IRR) through experts discussion in iterative processes. The rating criteria for raters to rate have prepared by reviewing the technical manuals of CC instruments, and relevant articles.

The rating criteria used to select CC instruments for inclusion in the study include the adequacy of the number of sub-constructs; the use of a clear framework; the appropriateness of the target group; an adequate sample size; methodological flows; originality; and reports of validity and internal consistency. The personal characteristics were selected based on raters identified from CC measuring tools, literature reviews, CC frameworks, and practical experience on personal, professional, and job-related cultural factors believed to have significance in delivering culturally congruent care to culturally diverse Ethiopian patients. The sub-constructs were decided from the number of factors identified in the selected CC tools, literature review, CC frameworks, and factor analysis results of related studies. The prepared rating criteria to guide this study are indicated in Table 1.

The experts at each stage have prepared codes to use as rating criteria guided by the prominent work of McAlister et al., on qualitative coding to assess Inter-Raters Reliability (IRR) through experts discussion in iterative processes. As well,
the preparation of rating codes were also supported by test construction materials, three articles on systematic reviews of CC instruments, and COSMIN checklist manual.81 Three experts in the first group designed rating codes mainly formulated from the COSMIN checklist and test construction material. According to the checklist, in such studies, health measuring instruments should be included if they meet the following criteria: internal consistency or reliability, content validity, structural validity, and cross-cultural validity. As well, they further enrich their rating codes with test construction materials and the three reviewed articles. They agreed upon the number of sub-constructs in the cultural competence instrument, appropriateness of the target group, comprehensiveness in addressing prominent cultural diversity factors, attention to the use of a clear study framework during instrument development, or validation or adaptation as their rating criteria to guide their instrument selection decision.

Out of the selected CC tools, the first group of experts decided on 20 relevant measuring tools to pass to the next stages of this study. Next, IRR rating codes were further enriched and agreed on adequacy of number of sub-constructs, clear theoretical frameworks, appropriateness of the target group, sample size adequacy, internal consistence report of the full scale and sub-scales, indicating relevance validity report, methodological flows, and originality for their decision by rating “1” to agree with the inclusion of the CC instrument or “0” to disagree with the inclusion of the CC instrument. The prepared rating criteria to guide this study are indicated in Table 1.

In general, the three health experts in the first group have participated to decide inclusion of number of CC tools, identify sub-constructs and determine relevance personal information for the subsequent parts of this study, i.e., pilot testing, validating and factor analysis. As well, they have decided the number of CC instruments to pass to item selection, and evaluation subsequent stages.

To select items from agreed CC instruments, three experts in the second group have further designed IRR codes based on McAlister et al., work. They prepared the codes after attaining orientation on good personality item writing, and independent readings. After synthesizing the principles to select items from COSMIN81 and test construction materials,82–84 Expert of this group have agreed to these items incorporated from relevance sub-constructs, appropriateness to the target population, and adherence to principles of item construction as rating codes for their decision of inclusion. For instance, the codes to guide experts of both in the second and third group to for item selection and evaluation respectively were used test construction materials of Kline (2015), DeVellis (2012), and Nunnally and Bernstein (1994).82–84

Figure 3. Procedure of the study.
### Table 1. Inclusion, and Exclusion Rating Criteria of CC Measuring Tools, Sub-Constructs, Socio-Demographic Characteristics, and Item Screening (Selection and Evaluation).

<table>
<thead>
<tr>
<th>Sources and Short Brief</th>
<th>Study Construct</th>
<th>Definition</th>
<th>Ratings of Inclusion</th>
<th>Ratings of Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 CC tools, 19-23,27-29,84-98</td>
<td>CC Measuring Instrument</td>
<td>Twenty CC Tools developed Adapted and Validated to Measure the CC of Health Professionals</td>
<td>Instruments in which raters decide based on their overall decision of inclusion relied on the rating agreements on the presence or absence including: adequacy of number of sub-constructs; use of clear framework; appropriateness of the target group; sample size adequacy, internal consistence report of the full scale and sub-scales; methodological flaws; originality</td>
<td>These tools who miss one or more of the criteria stated in ratings of inclusion</td>
</tr>
<tr>
<td>Socio-Demographic Characteristics</td>
<td>The factors that signifies in identifying the personal, professional, and job related variables for Ethiopian health worker.</td>
<td>The raters identified from CC measuring tools, literature review, CC frameworks, and practical experience on personal, professional, and job related cultural factors believed to have significance impact in delivering cultural congruent care to culturally diverse Ethiopian patients</td>
<td>These factors missed the agreement of one or more raters during in-depth discussions were excluded</td>
<td></td>
</tr>
<tr>
<td>Sub-Constructs</td>
<td>The number of factors identified in the selected 20 CC tools, and there relevance guided by literature review, CC frameworks, and factor analysis</td>
<td>Most widely incorporated sub-constructs from CC measuring tools raters decided for inclusion in this study.</td>
<td>Least incorporated sub-constructs were prone to exclusion</td>
<td></td>
</tr>
<tr>
<td>Items from six CC tools such as CCET, 29, CCA, 88 CCAI-UIC, 27 NCCS, 28 K-CCS-CN, 23 CCS, 96 As well, the TSET 87 and IAPCC 27 to guide the study</td>
<td>Item Screening and Selection</td>
<td>The process of choosing relevance items from agreed CC tools for Inclusion by three raters</td>
<td>Includes agreed sub-constructs, target group, addressing major cultural factors, relevance to the rich cultural diversity of Ethiopia, perceived quality of the item</td>
<td>Excludes lack of target group, perceived to lack quality, factual type of items, ambiguity, multiple ideas</td>
</tr>
<tr>
<td></td>
<td>Item Evaluation</td>
<td>The process of raters to make either inclusion or exclusion decision based on CC items adherence or violation of principles of item writing, their expertise, and research experience</td>
<td>Items adherence to principles of item writing, including: relevance to the target group, use of clear language clarity, unambiguous or easily understood by subjects; refer specific rather than general behavior, ask only one question, written to be relevant to Ethiopian health workers, respondents must be capable of responding, and use straightforward wording</td>
<td>These items raters agreed to be prone to acquiescence, social desirability, endorsement, double barring, double negatives, multiple interpretation, above 20 words, terms of feeling, brevity should be excluded</td>
</tr>
<tr>
<td>Methodological materials, 23,28,29,85,86 and six CC tools.</td>
<td>Contextualization</td>
<td>The process of making selected items to become appropriate in the cultural health setting of Ethiopia through discussion</td>
<td>These items which have words or phrases were replaced by relevance world appropriate in the context of Ethiopian Health</td>
<td>No item is excluded at this stage</td>
</tr>
</tbody>
</table>
According to DeVellis, the criteria to select or evaluate the utility of personality items, including: don’t make stuff in the past tense; make items with only one idea; keep away from double negatives; items with a simple and direct sentence structure are preferred; avoid employing absolute words as only, just, always, and none; avoid items with many interpretations; avoid items that are likely to be endorsed by everyone; use straightforward language; and keep items to a maximum of 20 words.82 As well, rating criteria such as the acquiescence response set, social desirability response set, a preference for responding to things in the center or ‘uncertain category’, and a subject’s preference for responding to the extreme response categories of Kline were considered as codes of rating in both the selection and evaluation of items.84

Finally, the following item construction principles suggested by Nunnally and Bernstein were also taken into consideration during both item selection and evaluation.83 Their suggestion include: items must be clear; use straightforward wording; all respondents will be able to offer information reliably when asked for an item; each respondent must be able to provide an answer to each question; and write specified short items.82 Therefore, experts in the second group and third group were taken into consideration in selecting and evaluating items by using the aforementioned item writing suggestions to guide their ratings and use as coding criteria. Table 1 depicts the coding criteria to be used for inclusion, and exclusion ratings.73

According to McAlister et al.,86 preparation of criteria is an initial step towards establishing IRR between raters within a team of experts. IRR was computed to determine the percentage of agreement between raters. IRR indicated that each rater might not have to rate either for inclusion or exclusion of same number of items, tools, sub-constructs, and demographic variables.

The IRR was calculated as the number of agreed ratings over the total number of ratings on the issue under study as per required. The computed IRR, which gives the fractional percent of the number of rated materials that agree, should be compared to the desired 80%-90% agreement. The overall IRR was calculated the number of ratings that all three raters agreed on divided by the number of total ratings. In addition, IRR between each individual pair of raters computed the total number of rated materials or sections divided by the number of ratings that both raters agreed on. Each set of two raters had two IRR values: (1) the number of times rater one agreed with rater two divided by the total number of ratings used by rater one, and (2) the number of times rater two agreed with rater one divided by the total number of ratings used by rater two.

Cohen’s Kappa was computed to know the IRR of two raters that evaluate the quality of the final items before being included. Cohen’s, symbolized by the lowercase Greek letter, is a robust statistic useful for either inter-rater or intra-rater reliability testing. Similar to correlation coefficients, it can range from -1 to +1, where 0 represents the amount of agreement that can be expected from random chance, and 1 represents perfect agreement between the raters.87

The computed reliability for the inclusion of instruments, CC sub-constructs, item selection, evaluation, and demographic variable identification were used as IRR coefficients86 and inter and intra-rater reliability testing standards.87 For computed IRR, the fractional percent of the number of rated materials that agree should be compared to the desired 80%-90% agreement.

Two raters were screened and evaluate items in this study. The IRR is computed using Cohen’s Kappa. The probability of number of items both raters of experts in the third group agree on inclusion (a) and exclusion (b) were calculated. Then, for the items IRR Cohen’s Kappa was computed. The IRR result should be evaluated against indicated reliability index indicated. The two experts again discussed on contextualizing items further for inclusion, then an increase on the computed Cohen’s Kappa is expected to attain. The final number of screened items was given for two experts for further evaluation.

Finally, the computed reliability has been compared with McHugh’s standard of reliability coefficients and level of agreement, it is supposed to be: none for 0-.20; minimal for .21-.39; weak for .40-.59; moderate for .60-.79; strong for .80-.90; and excellent for above .90.

Results
A total number of 103 articles or documents were included in this study. For the subsequent analysis of this study, there were mainly 20 cultural competence measuring instruments of 20 to 83 items were incorporated. These instruments are the cultural competency instrument (CCI),88 ethnic competency skills assessment inventory (ECSAI),89 cross-cultural evaluation tool (CCET),90 Italian nurses’ cultural competence questionnaire (INCCQ),91 cultural knowledge scale (CKS),92 cultural diversity questionnaire for nurse educators (CDQNE),80 cultural self-efficacy scale (CSES),93 cultural
The first group of experts selected sub-constructs, and personal information that are relevant to capture cultural competence in the Ethiopian context. From these instruments agreed for inclusion at least by one expert in the first group. As well, experts in the second and third groups were selecting and evaluating items, respectively, from these CC tools entirely agreed for inclusion by the three experts in the first group.

However, to use as a guiding framework for the selection and evaluation of items, the first group of experts and researchers have agreed to keep two CC tools such as IAPCC17 and TSET.98 This is because most the existing CC instrument development, validation, and adaptations are informed either by the two instruments or the frameworks they used to develop their respective tools.30,67,80,94,101

In the following sections, a review of CC measuring tools, selection of CC tools, identification of sub-constructs, socio-demographic variables, and item screening and evaluation: item wordings to contextualize to Ethiopia’s cultural setting will be presented.

Review of CC measuring tools

Before presenting findings of the study, the reviewed results of CC measuring tools were presented in Table 2.73

These 20 instruments include: Nurse CC Scale (NCCS) of Perng and Watson27; Cultural Self-Efficacy Scale (CSES) of Bernal and Froman93; CC Clinical Evaluation Tool of Jeffreys and Dogan95; Cultural Competency Instrument (CCI) of Kosoko-Lasaki et al.91; Inventory for Assessing the Process of Cultural Competency (IAPCC and IAPCC-R) of Campinha-Bacote27; Cultural Knowledge Scale (CKS) of Brathwaite and Majumdar90; CC Assessment Tool (CCAT) of Vasiliou and Kouta19; Transcultural Self-Efficacy Tool (TSET) of Jeffreys and Smodlaka87; Cultural Diversity Questionnaire for Nurse Educators—CDQNE of Sealey89; Ethnic Competency Skills Assessment Inventory (ECSAI) of Napholz89; CC Assessment Instrument University of Illinois at Chicago (CCAI-UIC) of Suarez-Balcazar et al.28; CA Scale (CAS) of Rew et al.95; CC Assessment (CCA) of Schim et al.88; Cross-Cultural Evaluation Tool (CCET) of Hughes and Hood30; Cultural Sensitivity (CS) of Chang, Yang, and Kuo94; Korean CC Scale for Clinical Nurses (K-CCS-CNs) of Chae and Lee23; Italian Nurses CC Questionnaire (INCCQ) of Gol, Dorosti, and Montazer98; Critical CC Scale (CCCS) of Almutairi and Dahinten79; Cross Cultural Evaluation Tool (CCET) of Terrance52; and CC Inventory for Nurses in China (CCINC) of Cai et al.97

The 20 CC instruments were developed, adapted, and validated between 1987 and 2019. 11 of these were for use in the United States; two for Taiwan; and seven for use in Cyprus, Canada, Korea, Saudi Arabia, China, Italy, and Louisiana. The CC tools consist of 20 to 83 items.

Four out of the 20 CC tools consist of 20 items each. Three tools were developed for use in the USA, including the CCI developed for clinical researchers and research nurses88; the CCET validated for nursing students,90 and the CCET developed for nurses.100 One CC tool, the INCCQ, was adapted for Italian nurses.91

ECSAI developed by Napholz89 for nursing students of the USA consists of 23 items. Two tools consist of 24 items each, including the CCAI-UIC developed for occupational therapy practitioners in the United States,28 and the CSK, adapted for public health nurses in Canada.92

There are three tools for assessing CC, each with 25 items. The IAPCC-R was developed for Registered Nurses in the USA,102 the CCA developed for Hospice nurses and caregivers in the United States19; and the CS validated for Community Health Nurses in Pingtung Country of Taiwan.76

Eight CC tools consist of 26 to 72 items, including the 26 items of the CCCS developed by Almutairi and Dahinten79 for healthcare workers of Saudi Arabia; the 29 items of the CCINC adapted by Cai et al.,97 for nurses and nursing students of China; the 30 items K-CCS-CNs adapted by Chae and Lee23 for nurses of Korea; the 36 items of CAS developed by Rew et al.95 for undergraduate and graduate nursing students of the USA; the 40 items of the CCAT developed by Vasiliou and Kouta19 for community nurses of Cyprus; the 41 items NCQS validated by Perng and Watson77 for on-the-job
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Subjects</th>
<th>Description of instrument</th>
<th>Stated conceptual framework</th>
<th>Reported reliability and validity</th>
<th>Score Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Nurse CC Scale (NCCS), Perng SJ and Watson R</td>
<td>On the job nursing students of Taiwanese</td>
<td>41 items, measuring 4 constructs: CA, knowledge, sensitivity, and skill.</td>
<td>Based on the literature of Campinha-Bacote, Jeffreys, and Literature Review</td>
<td>Cronbach's alpha of 0.73 to 0.87. Significant alpha increases on test-retest. Subjects submitted to factor analysis, by inspection for the concept cross-cultural interaction, not reported</td>
<td>Survey—5-point Likert-type rating scale assessing behaviors with range from always exhibited to never demonstrated</td>
</tr>
<tr>
<td>(2) Cultural Self-Efficacy Scale (CSES), Bernal Hand Froman R</td>
<td>Community health nurses of the USA</td>
<td>30 items testing knowledge of cultural concepts comfort in performing cultural nursing skills and knowledge of cultural patterns for specific ethnic groups.</td>
<td>Consistent with Bandura Self-Efficacy (Social Learning Theory) and Leininger; otherwise not linked with an overarching conceptual model</td>
<td>Cronbach's alpha 0.97 overall with subscales ranging from 0.85 to 0.98. Content validity determined by factor analysis confirming construct validity</td>
<td>Survey—5-point Likert-type scale with 1—very little confidence and 5—quite a lot of confidence</td>
</tr>
<tr>
<td>(3) CC Clinical Evaluation Tool, Jeffreys and Dogan</td>
<td>Nursing Students, Nursing Teachers of the USA</td>
<td>83 items with three sub-scales such as Provision of Cultural Specific Care, Cultural Assessment, and CS.</td>
<td>Jeffery's CC and Confidence Model. It is an Assessment tool adapted from TSET</td>
<td>Reported simply the internal consistency is computed using Cronbach alpha. Content validity is evidenced using CVI</td>
<td>Survey—10-Point with indifferent Rating Scales</td>
</tr>
<tr>
<td>(4) Cultural Competency Instrument (CCI), Kosoko-Lasaki et al.</td>
<td>Clinical researchers and research nurses of the USA</td>
<td>20 items, assesses knowledge, attitudes, and sensitivity to other culture.</td>
<td>Not reported.</td>
<td>Not reported.</td>
<td>Survey—multiple choice</td>
</tr>
<tr>
<td>(5) Inventory for Assessing the Process of Cultural Competency (IAPCC and IAPCC-R), Campinha-Bacote</td>
<td>Registered nurses of the USA</td>
<td>Original scale of 20 items revised scale of 25-item questionnaire based on 5 constructs, knowledge, skill, awareness, and encounters.</td>
<td>Based on The Process of CC in the Delivery of Healthcare Services</td>
<td>Reliability of original version reported as a “limitation.” Content validity is determined by 5 national health care experts examining the instrument for reliability. Psychometric values not reported.</td>
<td>Survey—4-point Likert-type scale with differing range for responses</td>
</tr>
<tr>
<td>Instrument</td>
<td>Subjects</td>
<td>Description of instrument</td>
<td>Stated conceptual framework</td>
<td>Score Methods</td>
<td>Reported reliability and validity</td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>---------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>(6) Cultural Knowledge Scale (CKS), Brathwaite and Majumdar</td>
<td>Public health nurses of Canada</td>
<td>24 items with 2 knowledge sub-scales—items taken from IAPCC-R and CSES.</td>
<td>Based on The Process of CC in the Delivery of Healthcare Services Campinha-Bacote</td>
<td>Survey—5-point Likert-type scale—strongly agree to strongly disagree and no opinion</td>
<td>Cronbach's alpha overall was .92 with subscales reliability of .75-.93 Content and face validity were determined by 2 expert panels validity testing by factor analysis with two factors loading over .40 and contrasted group validity was determined Construct validity Not reported</td>
</tr>
<tr>
<td>(7) CC Assessment Tool (CCAT), Vasilio and Kouta</td>
<td>Community Nurses of Cyprus</td>
<td>40 items with four subscales such as CA, CK, CS, and Cultural Practices. The response option is not specified</td>
<td>The Papadopoulos, Tilki &amp; Taylor Model for Developing CC</td>
<td>Some items are rated as Yes or No. And other items are rated as frequently, often and rarely.</td>
<td>Reliability is computed using a small number of subjects; 28 for the first test and 10 for the retest. The Cronbach Alpha ranged from .643 to .786. Generally, the validity and reliability of the tool is not reported precisely.</td>
</tr>
<tr>
<td>(8) Transcultural Self-Efficacy Tool (TSET), Jeffreys and Smolak</td>
<td>Nursing students of the USA</td>
<td>83-item questionnaire with 3 subscales: cognitive, practical, and affective.</td>
<td>Albert Bandura’s Social-Cognitive theory, concepts of transcultural nursing</td>
<td>Survey—10-point Likert-type scale. For all items 1—not confident to 10—totally confident</td>
<td>Total alpha .97 and .98 with subscale alpha ranging from .90 to .98 Content validity by 6 member expert panel followed by factor analysis confirming construct validity</td>
</tr>
<tr>
<td>(9) Cultural Diversity Questionnaire for Nurse Educators—CDQNE, Sealey</td>
<td>Nursing faculty Members of Louisiana</td>
<td>72 items with 2 subscale measuring 5 constructs of desire, awareness, knowledge, skill, and encounters—included items from the IAPCC-R.</td>
<td>Based on The Process of CC in the Delivery of Healthcare Services Campinha-Bacote,</td>
<td>Survey—5-point Likert-type scale—strongly agree to strongly disagree</td>
<td>Not reported Content validity was determined by a panel of experts</td>
</tr>
<tr>
<td>(10) Ethnic Competency Skills Assessment Inventory—(ECSAI), Napholz</td>
<td>Nursing students of the USA</td>
<td>Originally designed for another discipline; used to examine self-report CC of nursing. 23 items with no description of items.</td>
<td>Not reported</td>
<td>Survey—5-point Likert-type scale—never to always</td>
<td>Coefficient alpha .9444 Validity Not reported</td>
</tr>
<tr>
<td>Instrument</td>
<td>Subjects</td>
<td>Description of instrument</td>
<td>Stated conceptual framework</td>
<td>Score Methods</td>
<td>Reported reliability and validity</td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>---------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>(11) CC Assessment Instrument—University of Illinois at Chicago (CCAI-UIC), Suarez-Balcazar et al</td>
<td>Occupational therapy practitioners of the USA</td>
<td>24 items, measuring 4 constructs: CA, knowledge, practices and skill.</td>
<td>Based on <em>The Process of CC in the Delivery of Healthcare Services</em> Campinha-Bacote</td>
<td>Self-Report 4 point likert-type Scale strongly Agree to Strongly Disagree</td>
<td>Cronbach’s alpha for 4 scales ranged from 0.75 to 0.82. Professional literature and expert panel for content validity. EFA was conducted for construct validity and four factors accounted for more than 40% of the variance. Not reported.</td>
</tr>
<tr>
<td>(12) CA Scale (CAS), Rew et al</td>
<td>Undergraduate and graduate nursing students of the USA</td>
<td>36-item questionnaire with 5 subscales: general education experience, cognitive awareness, research issues behavior/comfort with interactions and patient care issues.</td>
<td>Based on the <em>Pathways Model</em> and consistent with the <em>Purnell Model of CC</em></td>
<td>Survey—7-point Likert-type scale—strongly agree to strongly disagree</td>
<td>Cronbach’s alpha .82 for faculty and for students. .91 on overall test. Content validity determined by 7 member expert panel. Reported with caution.</td>
</tr>
<tr>
<td>(13) CC Assessment (CCA), Schim et al</td>
<td>Hospice nurses and caregivers of the USA</td>
<td>25 items with subscales—Cultural attitude and sensitivity, cultural diverse experiences, and CC behaviors.</td>
<td>Based on <em>CC Model of Schim and Miller</em>. High influence on the Campinha-Bacote the Process of CC Model and IAPCC. This is manifested in their use of different Likert-Scale Rating Options</td>
<td>Survey—8-point Likert-type scale—strongly agree to strongly disagree and no opinion or 8-always to 1-not sure</td>
<td>Cronbach’s alpha overall was .92 with subscales reliability of .75-.93. Content and face validity determined by 2 expert panels. Validity testing by factor analysis with two factors loading over .40 and contrasted group validity was determined. Construct validity not reported.</td>
</tr>
<tr>
<td>(14) Cross-Cultural Evaluation Tool (CCET), Hughes and Hood</td>
<td>Nursing students of the USA</td>
<td>20 items.</td>
<td>Based on items from unpublished presentation of Freeman,</td>
<td>Survey—5-point Likert-type rating scale assessing behaviors with range from always exhibited to never demonstrated</td>
<td>Validity and reliability not reported.</td>
</tr>
<tr>
<td>(15) Cultural Sensitivity (CS), Chang, Yang and Kuo</td>
<td>Community Health Nurses Pingtung Country of Taiwan</td>
<td>25 items with five subscales (interaction, engagement, respect for cultural difference, interaction confidence, interaction alternative).</td>
<td>Not clearly stated or reported</td>
<td>Response option in interquartile range 0-70</td>
<td>Cronbach’s alpha overall was .90. Content validity is reported.</td>
</tr>
<tr>
<td>Instrument</td>
<td>Subjects</td>
<td>Description of instrument</td>
<td>Stated conceptual framework</td>
<td>Score Methods</td>
<td>Reported reliability and validity</td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>---------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>(16) Korean CC Scale for Clinical Nurses (K-CCS-CN), Chae and Lee</td>
<td>Nurses of Korea</td>
<td>Around 74 Initial items later reduced into 33 items measuring 4 constructs: CA, knowledge, sensitivity, and skill.</td>
<td>Based on The Process of CC in the Delivery of Healthcare Services Campinha-Bacote together with Papadolous et al Tikli and CC model</td>
<td>Survey—5-point Likert-type scale—strongly agree to strongly disagree</td>
<td>Cronbach's alpha for 4 scales ranged from 0.78 to 0.96. Face validity established through reviews by 4 experts. Reported but yet not reviewed.</td>
</tr>
<tr>
<td>(17) Italian Nurses CC Questionnaire (INCCQ), Gol, Dorosti and Montazer</td>
<td>Nurses of Italy</td>
<td>20 items in 3 sub-scales (learning and teaching, awareness and knowledge, skill).</td>
<td>Not reported</td>
<td>Survey in 5-Point Likert Scale</td>
<td>Cronbach's alpha for the whole scale is .91. Content validity is established but not clearly reported.</td>
</tr>
<tr>
<td>(18) Critical CC Scale (CCCS), Almutairi</td>
<td>Healthcare workers of Saudi Arabia</td>
<td>26 items measuring 4 constructs; Critical Awareness, Critical Knowledge, Critical Skills. The response scale for the CCCS.</td>
<td>The theoretical base—the Critical CC (CCC) model—which was developed through empirical research and subsequent theoretical refinement. The study was informed by Campinha-Bacote's CC framework</td>
<td>a seven-point Likert scale, anchored from 1 (&quot;strongly disagree&quot;) to 7 (&quot;strongly agree&quot;)</td>
<td>Cronbach's alpha for 4 scales ranged from 0.75 to 0.82. Professional literature and expert panel for content validity, EFA conducted for construct validity and Four factors accounted more than 40% of variance. Not reported.</td>
</tr>
<tr>
<td>(19) Cross Cultural Evaluation Tool (CCET), Terrance</td>
<td>Nurses of the USA</td>
<td>20 items.</td>
<td>Professional literature and expert panel for content validity</td>
<td>5 point Likert scale strongly agree to strongly disagree</td>
<td>Cronbach's alpha .83 for the full scale. Pretest and posttest reliability ranged from .73-.87. Tested by Attendees at cultural diversity workshop. Concurrent validity, and Contrasting groups is not found (this common for most instruments. Factor analysis with principal component analysis for construct validity. Four factors accounted around 50% of variance in cross cultural interaction scores.</td>
</tr>
<tr>
<td>(20) CC Inventory for Nurses in China (CCINC), Cai et al.</td>
<td>Nurses and Nursing Students of China</td>
<td>29 items with five subscales such as CA, skill, understanding, knowledge, and respect).</td>
<td>Literature Review</td>
<td>Survey-5-point Likert Scale</td>
<td>Cronbach's alpha for the instrument was .94 with a range of .79-.92 for each dimensions. Construct validity was estimated by exploratory factor analysis.</td>
</tr>
</tbody>
</table>
nursing students of Taiwan, and the 72 items of CDQNE developed by Sealey[^80] for nursing faculty members of Louisiana.

Two of the 20 CC tools with 83 items were developed for use in the USA to the target group of students and teachers are the CCSET of Jeffreys and Dogan[^35] and TSET of Jeffreys and Smodlaka[^27]. The sections below raise conceptual frameworks, response options, and reliability of CC instruments.

The conceptual framework used by eight CC tools, such as IAPCC-R, CKS, CCAT, and CCAI-UIC, is based only on the Process of CC in the Delivery of Healthcare Services by Campinha-Bacote. The remaining four CC tools were highly informed by this model and with one or more additional frameworks. These include NCCS, which was based on Jeffreys’ Confidence Model and Literature Review; CCA was based on the CC Model of Schim and Miller; K-CCS-CN was based on Padopoulos, Tilki and Taylor’s CC model, and CCCS was based on the Critical CC (CCC) model of Almutairi and Dahinten.[^97]

CSES was based on the Consistent with Bandura Self Efficacy (Social Learning) Theory and Leininger’s Cultural Care Theory; the CCSET was based on Jeffery’s CC and Confidence Model and it is an assessment tool adapted from TSET; CCAT was based on the Papadopoulos, Tilki, and Taylor Model for Developing CC; TSET was based on Albert Bandura’s Social-Cognitive Theory; concepts of transcultural nursing; CAS was based on the Pathways Model and consistent with the Purnell Model of CC. TSET was based on Albert Bandura’s Social-Cognitive Theory and concepts of transcultural nursing.[^34,98] CSES, CCET, and CAS were based on the Purnell Model of CC. The theoretical basis of ECSAI, CS, and INCCQ was not reported.

Two instruments, such as IAPCC-R and CCAI-UIC, have four-point Likert-Type rating response options. Nine CC tools such as NCCS, CSES, CKS, CDQNE, ECSAI, CCET, K-CCS-CN, INCCQ, and CCINC have a survey of a five-point Likert-Type rating scale. The CAS and CCET have seven-point Likert-Scale response options. One exceptional CC tool, CS, has response options in the interquartile range of 0-70. CCET has a survey with a 10-point Likert-type scale with indifferent rating scales, and TSET has a survey with a 10-point Likert-type scale, for all items with 1—not confident to 10—totally confident.

CSES, CKS, TSET, CCA, CCET, and CCINC full-scale reliability ranged from 0.83 to 0.98, and sub-scales reliability ranged from 0.73 to 0.98. NCCS, CCAI-UIC, K-CCS-CN, and CCCS sub-scales reliability ranged from 0.64 to 0.96, but the full scales reliability was not reported. On the contrary, the full-scale reliability of ECSAI, CAS, CS, and INCCQ ranged from .82 to .94, but the reliability of the sub-scales was not reported. However, both the full-scale and sub-scale reliability of CCSET, CCI, and CDQNE were not reported.

In subsequent sections, study findings on IRR for inclusion and exclusion of CC measuring tools, socio-demographic characteristics, CC sub-constructs, item screening, and evaluation of items were presented.

**IRR of study constructs**

The 20 carefully selected CC tools were further evaluated to select the final instruments, determine sub-constructs incorporated, identify major socio-demographic variables, select relevance items, and evaluate items. Table 3 depicts the aforementioned issues, and the IRR of each themes of the study.[^27]

**Selection of CC measuring tools**

For the selection of CC measuring tools, the IRR of three experts’ ratings were computed. The results show that IRR was found to be between 75-87.5%. In other words, the raters’ disagreement ranged between 12.5%-25% owing to raters differing interpretation in evaluating tools that might be originated from lack of prior experience in executing similar tasks.

Eight CC tools, including; NCCS, IAPCC-R, TSET, CCAI-UIC, CCA, K-CCS-CN, CCCS, and CCET have gained 100% agreement of the three raters. Whereas, four instruments, including; CCI, ECSAI, CCET, and INCCQ have gained 100% disagreement of inclusion by the three raters. As well, six instruments including CSES, CCCET, CCAT, CAS, CS, and CCINC have attained 67.7% agreement of two raters, and disagreement of one rater. Finally, two instruments, the CKS and CDQNE have gained 33.33% of agreement by one rater, and disagreement by two raters.

**Identifying sub-constructs**

The computation of IRR here is conducted for the major six factors incorporated in the remaining 16 CC measuring tools based on the agreement of inclusion by one or more raters. The IRR of CS is 38%, CD is 44%, CE is 50%, both CK and CA is 75%, and CSK is 81%.
<table>
<thead>
<tr>
<th>Name of the Construct</th>
<th>Rater One</th>
<th>Rater Two</th>
<th>Rater Three</th>
<th>Agreement?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measuring Instrument. These 20 CC tools indicated in Table 1</td>
<td>NCCS, 28, CCCET, 95, IAPCC, 95, CAT, 95, TSET, 97, CCAI-UIC, 95, CAS, 94, CCA, 94, CS, 94, K-CCS-CNs, 95, CCCS, 96, CCET, 95, CCIN 97</td>
<td>NCCS, 28, CCCET, 95, IAPCC, 95, TSET, 97, CCAI-UIC, 95, CAS, 94, CCA, 94, K-CCS-CNs, 95, CCCS, 96, CCET, 95, CCIN 97</td>
<td>NCCS, 28, CSES, 98, IAPCC, 95, CAT, 95, TSET, 97, CCAI-UIC, 95, CAS, 94, CCA, 94, K-CCS-CNs, 95, CCCS, 96, CCET, 95, CCIN 97</td>
<td>Total Coded=16, Rater One total codes:13 (IRR=13/16=81.25%), Rater Two total codes: 14 (IRR=14/16=87.5%), Rater Three total codes: 12 (IRR=12/16=75%). Eight instruments attained 100% of three raters’ agreement. Four instruments attained 100% of three raters disagreement.</td>
</tr>
<tr>
<td>Sub-Constructs</td>
<td>The three Raters agreed upon factors or sub-constructs such as CK, CA, CSK, CS, and CE and 16 CC instruments. As well, the factors such as Practical, Affection, Cognition are agreed as broader representation of CC sub-constructs</td>
<td>Similar with the other two raters</td>
<td>Similar with the other two raters</td>
<td>The rating is not computed by raters, rather against 16 multidimensional CC tools. Agreed CC tools: 16, CSK: 13 (IRR=81%), CA: 12 (IRR=75%), CK: 12 (IRR=75%), CE: 8 (IRR=50%) CS: 7 (IRR=44%), CS: 6 (IRR=38%) (McAllister et al. 2017).</td>
</tr>
<tr>
<td>Socio-Demographic Characteristics</td>
<td>Age, religion, gender, education, experience, cultural diversity encounter, cultural training, level of CC, organisational type, ethnic-self identification</td>
<td>Age, gender, education, experience, cultural diversity encounter, cultural training, language use</td>
<td>Unit/department, age, religion, gender, language skill, education, experience, cultural diversity encounter, cultural training, level of CC, organisational type, performance evaluation result</td>
<td>Total Rated demographic factors=14, Rater One total rated:10 (IRR=9/14=64%), Rater Two total rated:7 (IRR=7/14=50%), Rater Three total rated:12 (IRR=12/14=86).</td>
</tr>
<tr>
<td>Item Selection</td>
<td>Agreed for the inclusion of 25 items from 41 items of NCCS, 38 items from 24 items of CCAI-UIC, 15 items from 26 items of CCCS, 10 items from 20 items of CCE, 17 items from 25 items of CCA, 22 items from 33 items of K-CCS-CNs</td>
<td>Agreed for the inclusion of 22 Items from 41 items of NCCS, 12 items from 24 items of CCAI-UIC, 10 items from 26 items of CCCS, 9 items from 20 items of CCE, 13 items from 25 items of CCA, 24 items from 33 items of K-CCS-CNs</td>
<td>Agreed for the inclusion of 24 items from 41 items of NCCS, 17 items from 24 items of CCAI-UIC, 14 items from 26 items of CCCS, 16 items from 20 items of CCE, 19 items from 25 items of CCA, and 20 items from 33 items of K-CCS-CNs</td>
<td>Agreement: 65, Rater One total rated: 105 (IRR=65/105=62%), Rater Two total rated: 97 (IRR=65/97=67%), Rater Three total rated: 110 (IRR=65/110=59%).</td>
</tr>
<tr>
<td>Item Screening and Evaluation</td>
<td>Agreed for the inclusion of 48 items and disagree on 17 items out of the total of 65 items</td>
<td>Agreed on the inclusion of 50 items but disagree on 15 items out of the total of 65 items</td>
<td>No Rater Three</td>
<td></td>
</tr>
<tr>
<td>Reasons for either Exclusion or inclusion with some modifications to items</td>
<td>Double barring, simplicity, endorsed, brevity, lacks clarity, Reconsidering the item,</td>
<td>Endorsed, double barring item, absoluteness word, use of more than 20 words, Absoluteness word ‘Usually’ &amp; ‘Completely’, Direct without idioms, Multiple interpretation, use of jargon word, use of more words, lacks target group, include self-referencing phrase or word</td>
<td>No Rater Three</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. IRR of Instruments, Sub-Constructs, Demographic Variables, item screening, and item evaluation.
Beyond raters’ agreement, IRR was computed with the exclusion of CAS and CCET due to unrelated sub-scales and limited number of sub-constructs respectively approximately reached 69%. Likewise, the percentage of each of the sub-constructs incorporation in these tools approximately ranged from 43% for CS to 93% for CSK. The CA and CK sub-constructs have equally incorporated by 86% in the aforementioned instruments. As well, CS, and CE were incorporated by 50% and 57% in the Reviewed instruments.

Socio-demographic variables
Regarding socio-demographic variables, the IRR was found to be 50-86%. From this, gender, age, religion, type of organization, department, language, professional category, education, language use, language skill, professional category experience, cultural diversity encounter, cultural training, and levels of cultural competence, ethnic self-identification, and performance evaluation result were decided to be incorporated to measure the socio-demographic characteristics of health workers in the cultural setting of Ethiopia.

Item selection
The IRR for the selection of items were computed similarly and found to be 59% for rater three, 62% for rater one, and 67% for rater two. The first three experts were agreed to keep TSET and IAPCC from the selected instrument to use as the framework of the subsequent study. Then, six CC measuring tools, including; the 41 items of NCCS, the 24 items of CCAI-UIC, the 26 items of CCCS (excluding the 17 items of critical empowerment), the 20 items of CCE, the 25 items of CCA, and the 33 items of K-CCS-CN’s were included for the purpose of item selection. Therefore, the raters were expected to select relevance items from the total of 169 items compiled from the aforementioned six CC instruments. Based on this, the raters were agreed and selected 65 items to pass to the next section of evaluation and item screening (See extended data - appendix A).73

Item screening and evaluation: Rational of exclusion and inclusion of items
The two raters both agree on the inclusion of 46 items and exclusion of 13 items. Then, the computed IRR by Cohen’s Kappa was found to be 0.75. The IRR result should be agreed to be moderate reliability. The two experts again discussed on contextualizing items for inclusion and both agreed to incorporate 10 additional items. Following this, the screened 65 items have further evaluated to decide the final included and excluded number of items.

The IRR was computed using the quotient on the number of items violated test construction principles that already identified in Table 2 of this paper (n=12) and the number each rater identified. Based on this, IRR of rater one is found to be 50% and rater two is 67%. Based on agreed appraising criterion indicated in Table 2, both raters have arrived in a consensus to discarded 13 items out of 65 items. The 65 selected items from the six CC tools, items evaluated by experts, the final 56 included items, and the 11 excluded items are indicated in Appendices A, B, C, and D, respectively.73

The items such as ‘health care knowledge and the patient’s comprehension of interpretation of health/illness are usually different systems’, ‘even if a patient’s use or adoption of a treatment method differs from my professional knowledge’, ‘I usually don’t prohibit it’, and ‘knowledge of cultural differences about the decision maker in the family’ were automatically discarded due to violation of one or more item construction principles.73

Items such as ‘I can explain the possible relationships between the health/illness beliefs and culture of the patients’, ‘I do not feel that I have the skills to provide services to ethnic minority clients’, and ‘there are no cultural variations within a cultural group of people’ were rejected, because they are endorsed by every respondent. Items such as ‘I openly discuss with others’ issues I may have in developing multicultural awareness’, and ‘Awareness of culture’s impact on perception on health and illness’ were rejected because they are prone to multiple interpretations.

Whereas the item ‘cultural and linguistic differences could compromise healthcare provider’s well-being’ was rejected due to violations of both double barring and brevity principles. The item, ‘it is not important to assess a patient’s preferences in terms of healthcare services if I am knowledgeable about their culture’ was excluded on violation of item writing principles such as the item should be short or comprehensible and brevity.

The items ‘thinking what it would be like if I were a foreign patient’, ‘trying to say simple phrases in the foreign patient’s language when I care for them’, and ‘knowledge of cultural differences about sensitivity to pain’ were rejected due to lack of clarity, in appropriateness of the target group, and lack of both self-referent and clarity respectively.

As well, both experts of the third group agreed to accept 11 items with some minor modifications as per the nature of the items. Some of the items both experts agreed for inclusion with some modification by mainly incorporating self-referent phrases include: ‘I understand a patient’s cultural background can promote the quality of Health care’, ‘I can compare the
health or illness beliefs among patients with diverse cultural background’, I have the Knowledge of cultural/religious belief that limits caring depending on gender’, ‘I use a variety of sources to learn about the cultural heritage of other people’, and ‘I have knowledge of cultural differences about touch or Space’.

Therefore, 13 items were excluded, 11 items modified for inclusion, and remaining 52 items passed for further evaluation.

**Item wording to contextualize to the cultural setting of Ethiopia**

The words or phrases in the items selected from six CC tools by replacing by contextual words or phrases to make items more appropriate for use in the health cultural context of Ethiopia. These words or phrases, including: ‘goals’, ‘most people’s’, ‘it is difficult replaced’, ‘race’, ‘a service provider’, ‘provider’s’, ‘compromise’, ‘people’, ‘all’, ‘members’, ‘AND’, ‘identify’, ‘school’, ‘relate’, and ‘foreign patients (FP)’ replaced or changed by ‘activities’, ‘culturally diverse patients’, ‘I face difficulties’, ‘health professional’, ‘health professional’s’, ‘hinder’, ‘patients’, ‘different’, ‘patients’, ‘OR’, ‘encounter’, ‘place’, ‘treat culturally diverse patients or patients whose ethnic is different from mine’ respectively. The noun ‘nursing’ has been replaced by ‘clinical’, and ‘health care’ depending on context. The word ‘client’ has been replaced by the ‘culturally diverse patient’, and ‘patients’.

The experts’ reasons to replace or remove words/phrases have included: to address lack of targets; to avoid items likely to be endorsed by everyone; to avoid items with multiple interpretations; to improve the clarity of items; to make items independent as possible; to avoid jargon words of items; to articulate the items intended factors to measure precisely; to restate items to keep proper arrangements, and to improve the manifestation of patterns of associations. As well, words like ‘only’, ‘usually’, and ‘completely’ that denote absoluteness of the items have been also removed. Furthermore, those items violated one or more principles of item constructions including double barring, lacks target group, might be open for endorsement, prone to multiple interpretations, caused brevity, included absoluteness words, stated as long sentences or with more than 20 words, lacks simplicity, stated to direct with idiomatic expressions, written too short, and proven lack of comprehensibility has been rewritten to improve the quality of the items.

These double barring items were rewritten either into two independent statements or stated as a single item based on the significance of the items. For instance, the item which stated ‘Can have appropriate verbal and nonverbal communications with foreign patients’ was rewritten as two statements; ‘I can have appropriate verbal communications with patients whose ethnicity or culture is different from mine; and ‘I can have appropriate nonverbal communications with patients whose ethnicity or culture is different from mine’ after the necessary modifications were made by experts. Another double-barring item, ‘I find ways to adapt my services to individual and group cultural preferences’ has been restated as a single item and written as ‘I find ways to adapt my services to individual cultural preferences.’

Lastly, the phrase ‘ethnic minority clients’ has been replaced by phrases such as ‘my culturally diverse patients’, and ‘patients whose ethnicity is different from mine’. On the other hand, the words or phrases like ‘behavior’, ‘ways’, ‘usually’, ‘completely’, ‘will’, ‘disability’, ‘being flexible, empathetic, and non-judgmental,’ and ‘group’ have been decided and removed.

Finally, 56 items were selected for further validation of the items to be included in the final version of the newly developed CC Scale to use in the context of Ethiopian health care settings.

**Discussion**

The objectives of this study include: identifying CC tools and demographic factors; detecting sub-constructs in CC tools; screening items in the agreed CC tools; and evaluating items’ adherence to test construction principles. For this end, 20 CC tools consists of 20-83 items were reviewed. The CC tools consist of 20 items are CCI of Kosoko-Lasaki et al., ECSAI of Napholz, INCCQ of Gol et al. Also, the CCAI-UIC of Suarez-Balcazar et al., CCA of Schim et al., CCCS of Almutairi AF and Dahinten VS, CSES of Bernal H and Froman R, K-CCS of Chae and Lee, CAS of Rew et al., NCSS of Perng and Watson consist of 24, 25, 26, 30, 34, 36, and 41 items respectively. Both TSET of Jefferys and Smolakia and CCCET of Jefferys and Dogan consist of 83 items. The number of ratings of the reviewed CC tools ranged from four to 10 Likert-Type Scale options. 12 out of the 20 CC tools use a five-point Likert-Type option.

The predominant CC sub-constructs incorporated in the reviewed 20 tools are CSK, CK, CS, and CA. The sub-constructs CS and CE are least incorporated in the reviewed tools. Whereas, factors denoted in different terms but believed to be similar with the aforementioned stated factors. For instance, TSET of Jefferys and Smolakia, the sub-construct cognitive intended to measure knowledge, understanding, application, synthesis, analysis, and evaluation capabilities of health heath workers as indicated in prominent work of Bloom. According to Bloom, the TSET’s affective sub-construct intended to guide health professionals’ way of responding, organizing the values, and internalizing the proven
cultural variations that help him/her in delivering culturally congruent care to culturally diverse patients; and the practical sub-construct measures how a health professional effectively manipulates with precision to handle the expected health services. Therefore, 83 items of TEST are believed to address six of the most commonly identified sub-constructs in comprehensive, reliable, and valid ways.

Generally, most CC tools have incorporated three sub-constructs such as CSK, CA, and CK. CE and CS were significantly incorporated into the CC tools next to the three sub-constructs. The sub-construct, CD is the least incorporated sub-construct in CC tools.

Specifically, factors incorporated in reviewed instruments of the IAPCCs’ of Campinha-Bacote, Campinha-Bacote’s CC and confidence model, the Purnell model of CC, the critical CC model, and literature review were used by professional experts in their development processes of the tools. On the contrary, a significant number of the aforementioned tools were not reported models or theories used to guide their tool construction process.

Concerning the conceptual frameworks of the CC tools, the most commonly used frameworks were Campinha-Bacote’s cultural care theory and Campinha-Bacote’s social cognitive theory, Jeffery’s CC and confidence model, the Purnell model of CC, the critical CC model, and literature review were used by professional experts in their development processes of the tools. On the contrary, a significant number of the aforementioned tools were not reported models or theories used to guide their tool construction process.

Therefore, for the subsequent study, Leininger’s cultural care theory and Campinha-Bacote’s process of CC model have been recommended and agreed to be used for the synthesizing of the then to be developed Ethiopian CC Scale for Health Workers study’s conceptual framework. This has been made necessary because of a lack of consensus on the definition of culture and CC. The term ‘CC’ has appeared regularly in health care literature in the past six decades, but there is no single, commonly accepted definition and it continues with ambiguity.

Therefore, the experts’ engagement is required to mitigate the lack of agreement on what exactly constitutes culture and CC in the health setting context by grasping the major tents of culture and CC through the models, theories, and literature. Besides, the stated limitations, together with a lack of attention to culture and CC issues in the health policy and practices of Ethiopia, have become added reasons to explore more theories, models, tools, and literature that are believed to help the adoption of a synthesizing conceptual framework more appropriate in the rich cultural diversity of Ethiopia.

The CCT of Leininger revealed that her work has created new insight into the nature of culture, care, and the transcultural health profession by presenting ways to describe, interpret, discover, and understand the diverse meanings and patterns of health and illness, care and healing, survival, and ways of facing and understanding death or disability by culturally diverse people. Scholars have recommended studies of CC in countries with people citing various languages used, the existence of proven ethnic variability, a high number of foreign patients, internal displacement, and increasing numbers of immigrants as common reasons for using CCT and believe that it properly guides similar studies of measuring health professionals’ CC.
Most CC theories and models have been constructed or adapted using Leininger’s pioneering ideas by blending them into their works. Based on the scholars’ recommendation, the experts of this study agreed to select the contents of items to measure health professionals’ level of CC. They should have been guided by CCT factors that encompass: worldview; beliefs, values, and practices; technology; education; economics; communication; ethno-history; kinship, family, and social; religious, philosophical, and spiritual; political and legal; and generic forks. Even though CCT identified the major components or sub-constructs to be incorporated in the adaptation, development, and validation of health workers’ CC, but it lacks clarity in conceptualizing their use as the sub-constructs.

Therefore, experts agreed to use Campinha-Bacote’s process of CC model, believed to be operationalized concepts of CCT actions, and the decision modes as sub-constructs of health professionals’ CC measurement. This model has been recommended as the proper model for such types of studies that adopt Leininger’s three actions and decisions into workable CC sub-constructs. According to Campinha-Bacote, the three CCT actions and decisions were further categorized into five sub-constructs such as CK, CA, CSK, CE, and CS to make it plausible to conceptualize and operationalize CC underlining manifestations.

Most CC measuring tools have used Campinha-Bacote’s the process of CC model and mainly incorporated CA, CK, CS, CSK, and CE as the major sub-constructs in the measuring tools of health providers and professionals’ CC. As well, the model of Padopoulos et al. identified CA, CK, CSK, and CS as four levels or stages, and they categorized CC as the final stage in which the previous stages are integrated and applied to skills, including recognizing and challenging all forms of discrimination.

Most of the reviewed CC tools assess the CC of health workers in a self-administered manner and are based on individuals’ perceptions. As well, most instruments developed to measure CC are commonly used, the Campinha-Bacote’s IAPCC. The tools included CA, CK, CSK, CE, and CS as the major sub-constructs of CC.

Generally, eight instruments were selected by experts to be incorporated for the selection of relevant items from each instrument. This group of three experts selected relevant items that were appropriate in the context of Ethiopia’s health setting. Upon their selections, a total of 65 items were initially selected by them.

The numbers of items selected are 13 from NCCS, 12 from CCAI-UIC, eight from CCCS, six from CCE, 10 from CCS, and 17 from K-CCS-CN. Then the items are evaluated for inclusion to pass to the next stage. Based on the experts’ judgment, 13 items were automatically excluded. For instance, items like ‘healthcare knowledge and the patient’s comprehension of interpretation of health/illness are usually different systems, even if a patient’s use or adoption of a treatment method differs from my professional knowledge’ and ‘I usually don’t prohibit it, and knowledge of cultural differences about the decision-maker in the family’ were automatically discarded. This is because of the violation of item construction principles and their similarity with other items decided to be incorporated in the initial version of the tools. While 11 items were accepted with minor modifications,

The remaining items were modified and contextualized to make appropriate to the cultural context of Ethiopia. This by addressing the words or phrases in the existing instruments has been replaced by other words or phrases that make the items more relevant and appropriate in the healthy cultural setting of Ethiopia.

Conclusion

According to the reviewed CC instruments and experts’ judgment, Campinha-Bacote’s process of cultural competence model in the delivery of healthcare services to culturally diverse patients and cultural care theory was the most widely used model. The experts thus recommended the adoption of Campinha-Bacote’s model along with theory that informs it the subsequent stages of developing the Ethiopian CC measure. More specifically, the experts suggested that, beyond the selected 56 items from the six existing instruments, additional items should be generated to address the rich cultural diversity of Ethiopia. The prominent among them are TSET and IAPCC recommended guiding the item development phase in order to tap the rich cultural diversity in Ethiopian health service delivery settings.

Most cultural competence measuring instruments included sub-constructs of cultural competence such as cultural skill, cultural awareness, and cultural knowledge. Cultural sensitivity, cultural desire, and cultural encounter were incorporated into 50% of the cultural competence instruments. This research suggested that the commonality that exists in the cultural encounter, cultural desire, and cultural sensitivity might have originated from the broad categorization of behavior such as cognitive, affective, and practical. These three factors were categorized under the affective behavioral domain that encompasses awareness, receiving, responding, valuing, organizing, and characterization.
Therefore, during cultural competence instrument development, authors might think that they can address the three sub-con structs when they construct items for the cultural awareness factor. The overlap in these three sub-constructs might make it preferable to develop cultural competence tools by rethinking the educational taxonomy of Bloom in his identified three major domains; cognitive, affective, and psychomotor.

As well, from the selected 65 items, some of the items were overlooked major and minor rules of test construction principles or guidelines. For instance, the items with proven existence of double-barring ideas, endorsed, long sentences, prone to multiple interpretation, encounter brevity, and use of ambiguous words were among the observed limitations of the items in the reviewed cultural competence instruments. This might be mitigated through the participation of experts in the field of psychometric. Most instruments were developed by health professionals, and most of them overlooked the ground rules of personality test construction that are highly recommended by prominent thinkers in the field.

Finally, generating additional items of cultural competence, the translation, and adaption of these 56 items, content validation and factor analysis should be conducted to have valid and reliable tools that measure the Ethiopian health workers’ cultural competence. Besides, to make the items more appropriate to the cultural health setting of Ethiopia. The content validation of the selected items and the generated items will be held together to refine the final items to be administered for pilot testing, and conduct CFA, and EFA.

Data availability

Underlying data


The project contains the following underlying data:

- **Figure 1. pptx** portrays the cultural care universality and diversity theory.
- **Figure 2. pptx** indicates the Campinha-Bacote’s the processes of cultural competence in the delivery of health care services.
- **Figure 3. pptx** demonstrates the procedure of the study.
- **Table 1. xlsx** depicts sources, short summaries, study constructs, operational definitions, rating codes for inclusion criteria, and rating codes for exclusion.
- **Table 2. xlsx** lists the instrument, subjects, instrument description, conceptual framework, scoring method, reliability, and validity.
- **Table 3. xlsx** includes issues related to the use of three raters’ independent ratings to compute the IRR of instruments, sub-constructs, demographic variables, item screening, and item evaluation.

Extended data


The project contains the following extended data:

- **Appendix A. xlsx** depicts number of CC items selected from six instruments for further evaluation.
- **Appendix B. xlsx** incorporates evaluation of items selected from CC instruments.
- **Appendix C. xlsx** reveals selected and modified items in the study.
- **Appendix D. xlsx** shows number of excluded items in the study.

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).


The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com