

## **Abstract**

Objectives. Autism research in low- and middle-income countries (LMIC), including Kenya, remains limited despite its growing global prevalence. This instrument development pilot study paper describes a process for developing culturally grounded survey instruments to assess the knowledge and beliefs of caregivers and educators of children with autism in western Kenya.

Methods. Our approach blended evidence-centered design, a psychometric framework for designing measurement tools that are valid for their intended use, with community-based participatory research, which emphasizes equitable stakeholder involvement.

Results. In this study, we detail the measure we created and, more importantly, the process used to blend evidence-centered design and community-based participatory research. We demonstrate how merging these two traditions can enhance the cultural and contextual validity evidence around survey instruments in low- and middle-income countries.

Conclusions. Our work offers a replicable model for researchers aiming to develop participatory, culturally responsive tools for deployment in underrepresented contexts and contributes to efforts to improve autism knowledge and support in Kenya and LMICs.

*Keywords:* Survey design, low-and middle-income countries, autism, community-based participatory research

## **Developing Culturally Responsive Surveys on Neurodevelopmental Disabilities: Lessons from the Kenyan Context**

Autism spectrum disorder (referred to as “autism” throughout) is a neurodevelopmental disability with a growing global prevalence, currently estimated at 1 in 100 children (Zeidan, et al., 2022). Autism is characterized by deficits in social communication and the presence of restricted and repetitive behaviors and interests (APA, 2023). Children with autism often require specialized support including behavioral and educational services (Kasari, Gulsrud, Freeman, Paparella, & Hellemann, 2012). Unfortunately, access to behavioral and educational intervention varies widely and is severely limited in low- and middle-income countries (LMIC), such as Kenya. These conditions are in stark contrast to those in more developed countries like the United States, with base-case spending estimated at \$80.6 billion annually as of 2021 for support of individuals with neurodevelopmental disabilities (NDD; Tanis, 2021). Though 80% of the world’s children are born in LMICs (World Bank 2022, as cited by Bitta et al, 2017), most of the research on autism has occurred in high-income countries (Gona et al., 2016). The relative lack of autism research in LMICs contributes to disparities in autism knowledge, diagnosis, education, and support services (Ruparelia et al., 2016). These disparities may contribute to misapplication of research findings from western cultures. In the context of our own work, we wanted to understand the knowledge and beliefs of educators and caregivers of children with autism in Kenya, basic information that is essential to serving children, families, and educators affected by autism, yet is virtually non-existent in Kenya and many other LMICs.

Partially in response to disparities in autism spending and support between LMICs and higher-income countries, researchers have formulated international partnerships to share knowledge and practices across countries. Some of the primary tools to explore knowledge, beliefs, and attitudes in international health and education studies are survey instruments. Surveys

have been used to understand autism awareness (Chansa-Kabali, Nyoni, & Mwanza, 2019), perceptions of autism knowledge (Harrison et al., 2024), and family quality of life of children with autism in Africa (Schlebusch, et al., 2017). Yet, studies have also pointed to the lack of research on autism in sub-Saharan African countries, including highlighting a lack of psychological measures and instruments specific to the African context (Chambers, 2017; Sousa & Rojjanasrirat, 2011; Taliep & Florence, 2012). Gaps in these measures include a lack of context-specific instruments related to prevention and interventions (Kakooza-Mwesige et al., 2014), diagnosis and specialist care (Smith et al., 2017), and quality education for developmentally disabled children (Genovesi et al., 2022). More specific to Kenya, Greer et al., (2022) mention that autism is poorly understood but widely diagnosed in Kenya.

Oftentimes, when measures do exist, they have been developed by researchers conducting the study. A key assumption underlying international surveys is that they measure the construct of interest consistently across cultures, including that ideas from one partner's country translate to those of another partner's country. In the case of autism research, a survey might implicitly assume that Western terms and ideas translate to other languages and cultures, an assumption that ample research suggests may not always be the case (Collins et al., 2018). We developed a survey instrument measuring the knowledge and beliefs of educators and caregivers of children with autism in Kenya because we did not want to rely on peripherally related surveys from other LMICs that may or may not have been true to the Kenya context. Simply assuming that related surveys from other LMICs would also be valid in the Kenyan context could lead to a fundamental misunderstanding of the knowledge and beliefs we care about in our work.

Several existing methods can be employed to help avoid such assumptions, including those derived from fields like psychometrics that are specifically concerned with the validity of a

measure for an intended purpose with a particular group of respondents. For example, in our survey design project, we employed evidence-centered design (ECD; Mislevy & Haertel, 2006; Mislevy et al., 1999), a powerful approach to “validity by design” in which measurement researchers answer three basic questions in sequence: What are the constructs to be measured? What behaviors or performances should reveal those constructs? What tasks or situations should elicit those behaviors? (Messick, 1994, p. 17). These behaviors are often (though not always) defined through engagement with individuals who are the target of measurement (in our case, parents and teachers of children with autism). Beyond psychometrics, community based participatory research (CBPR) is a research approach that emphasizes equitable partnerships to reduce health disparities, particularly among marginalized groups (Wallerstein, et al., 2017) used across fields including health and education (Ortiz et al., 2020). In the context of our own work, CBPR could be used to develop a culturally grounded understanding of educational experiences of children with autism/NDD and their families.

Despite important links between approaches like ECD and CBPR, they have rarely been discussed in tandem (Farrell et al., 2024) and certainly have not been blended in the context of research to support children with autism in LMICs. In addition, few studies lay out a process and potential best practices for survey design in the context of LMICs, international autism research, or both (e.g., Afshari, 2013). With these gaps in mind, the current study documents the process used to develop a survey instrument measuring the knowledge and beliefs of educators and caregivers of children with autism in Kenya. That instrument is called the Autism in the Context of Education – Kenya Survey, or ACE-KS. The validity of a measure is interwoven with its intended use; ours was to gain information on the attitudes about, and understanding of, autism among teachers and parents affected by it such that we could understand how best to support their needs in caring for

children with autism. In our efforts to develop the instrument (with this intended use and accompanying validity argument in mind), we employed both ECD and CBPAR. The purpose of this paper is to describe the context in which this work occurred, outline the survey design process employed, and discuss lessons learned. By describing our approach, we hope to provide a useful resource both to autism researchers conducting international work, and to education and health care researchers interested in cross-cultural, participatory based instrument design. We hope this study helps other researchers move beyond the common practice of simply adopting an existing measure from a more developed country, or even another LMIC that may have important cultural differences with the site of the research. We would argue this “off-the-shelf” approach can not only undermine the validity of using an existing survey in a different context but may also lead researchers to conclusions that are antithetical to the reality of the LIMC they intend to study (Soland, 2021; Xu & Soland, 2024).

## **Background**

### **The Context: Autism in Kenya**

The most recent estimates suggest .63% of adolescents and young adults in Kenya present traits associated with autism (Mamah, et al., 2022). This is likely an underestimate of prevalence driven by underdiagnosis and underreporting of disability due to factors including widespread poverty, limited governmental resources related to disability rights and education, disability stigma, and lack of data collection on prevalence, services and outcomes for autistic children (Human Rights Clinic at UVA Law, 2024). Many Kenyan communities have limited awareness or understanding of autism/NDD and believe disabilities to be a result of witchcraft, curses, or bad omens (Masaba et al., 2021). Blame for the onset of disability is often placed on the mother along with shame for her parenting skills (Greer, Sood, & Metcalfe, 2022). When a child is believed to

have a disability, often the entire family faces stigma from the community, which can result in loss of employment, loss of access to village support and resources, abandonment of the mother and children by the father (Bunning et al., 2017), and even injury or death to the child (Bayat, 2015).

There are limited resources for screening and diagnosing disability in Kenya (Abubakar et al., 2016). Diagnostic tools commonly used in western countries have not been adapted or validated for their intended uses in Africa (Ruparelia et al., 2016). Furthermore, national and county governments often do not have policies, practices, or guidance on diagnosis (Kamau, 2017). Healthcare and educational assessment workers frequently lack training in recognizing and diagnosing autism/NDD (Bakare & Munir, 2011; Elder & Kuja, 2019). Compounding the issue, families may delay bringing their child for evaluation due to stigma, hope that the symptoms will resolve with time, struggle with financial burdens and lack of health insurance coverage for autism/NDD intervention and lack awareness of educational opportunities and possible outcomes for children with disabilities (Bakare & Munir, 2011).

The relation between diagnosis and access to subsequent educational interventions and supports in Kenya is unknown. Kamau (2017) conducted interviews and focus groups with caregivers and professionals in Nairobi and found that, following a diagnosis of autism, many families sought guidance from religious leaders or traditional healers rather than from the education system. The author posits this may be due in part to perspectives on the cause of autism/NDD and implications for care, limited access to schools, and a near total absence of educational intervention research for autism/NDD in Kenya.

### ***Research on Educational Services for Children with Autism/NDD in Kenya***

Children with autism/NDD who do attend school are likely to attend Special Schools for children with intellectual disabilities (Karisa, McKenzie, & De Villiers, 2021). Special Schools are

segregated, often residential facilities where the focus is generally on custodial care rather than educational intervention (Muuya, 2002). This custodial focus leads to low student expectations, limited instruction, and ultimately poor educational outcomes for students with disabilities (Abubakar, et al., 2022). The custodial focus is exacerbated by disability stigma, lack of educational funding, costly fees, and minimal teacher special education training (Mukuria & Korir, 2006).

In 2002 Muuya surveyed headteachers (i.e. principals) in Special Schools in Nairobi and the surrounding areas. They found that headteachers viewed personal care and child obedience as the primary goals of special education and noted the lack of alignment between school staff practices and national policy towards disability rights and education in Kenya. Given disability stigma and limited teacher training in autism/NDD, inclusion of children with disabilities with non-disabled peers is rare (Elder & Kija, 2019). Genovesi et al. (2022) systematically reviewed research examining stakeholder experiences related to inclusive education of children with NDD in sub-Saharan Africa. The authors found that, while stakeholders supported the idea of inclusion of children with disabilities, barriers included unclear educational policies, limited awareness of educational opportunities, lack of teacher training, and community stigma towards people with NDD. The review did not include any studies from Kenya nor perspectives of families or teachers.

### ***Research on Families of Children with Autism/NDD in Kenya***

Caregivers of children with autism/NDD in Kenya have reported experiencing stigma, as well as financial and emotional burden (Masaba et al., 2021). Stigma has led caregivers to experience exclusion from religious services (Gona et al., 2016), community rejection (Chabeda-Barthe et al., 2019) and changes in family structure (Kamau, 2017). Families have reported that access to care is not available in rural areas and the cost of traveling for services is prohibitive

(Elliott, 2019). In some cases, the challenges in caring for a child with autism have led to child abandonment and significant caregiver distress (Cloete & Obaigwa, 2019). Mkabile et al. (2021) systematically reviewed the literature on African family experiences raising a child with NDD. Their review contained only nine studies, once again highlighting the paucity of research related to prevalence, diagnosis, education, and impact of care for individuals with NDD in Africa. The authors identified variation in understanding the cause of NDD, stigma and discrimination, significant caregiver burden, and lack of specialized educational services. Rispoli et al., (in preparation) reviewed the literature on caregiver training in autism/NDD intervention in sub-Saharan Africa and identified only seven empirical studies. This finding underscores the critical need for research exploring autism services and care in Africa.

### **Survey Research in LMICs**

Surveys are a primary tool used to understand neurodevelopmental disorders, like autism, as well as health challenges more generally in LMICs. To employ a survey for such a purpose, they must be supported by validity evidence for their intended use with a given population of interest. Messick once described validity as “an integrated [on-going] evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on... scores” (Messick, 1989, p. 13). There are several key aspects to this quote that still undergird current approaches to validation efforts. First, a measure and its scores are not simply valid or invalid; rather, validity is inextricably linked to not only the intended use of the scores, but also the population being measured. In the context of LMICs, measures supported by hundreds of studies in more developed Western countries may not, and likely are not, valid for use in LMICs, including from one specific LMIC to the next. Second, because a measure is neither simply valid nor invalid, validity is about an ongoing accrual of evidence that one is



justified in employing the measure for its intended use. Pieces of evidence can include theoretical formulations, subject matter expertise, statistical analyses of empirical data, and qualitative results around how a measure is interpreted and perceived. Thus, in what follows, we describe how we engaged in a process guided by theory, driven by Kenyan researchers and stakeholders, and assisted through the ongoing, evaluative analysis of data (quantitative and qualitative). In short, we used a validity by design approach that is at the heart of ECD and that, we would argue, is implicit in the broader practices embodied by CBPR.

### ***Community-Based Participatory Research***

While CBPR is not an approach designed narrowly for survey design alone, it nonetheless yields a potential way to produce culturally grounded survey instruments. To promote equitable participation and benefit, such partnerships often involve capacity building, sustainable development, research with bidirectional benefits, and humility (Plamondon et al., 2021). The CBPR framework engages key stakeholders as co-researchers within a collaborative partnership centered on reciprocal learning (Sors et al., 2022). The CBPR framework is designed to share knowledge among partners, and to develop culturally appropriate measures and analyses with the goal of deep understanding of community circumstances to facilitate the development and adaptation of practices to reduce disparities (Viswanathan et al., 2004). CBPR features include: (a) recognizing community autonomy, (b) building on community strengths and resources, (c) promoting bidirectional learning among partners, (d) balancing research and action to benefit all partners, (e) focusing on community defined problems, (f) iteratively developing and maintaining community partnership, (g) disseminating research findings to and by all partners, and (g) forming a long-term commitment among all partners (Holkup et al., 2004). In the context of autism/NDD in Kenya, CBPR could be used to generate new knowledge on the educational experiences of

children with autism/NDD in western Kenya, as well as the experiences of their families and teachers.

### ***Evidence Centered Design***

More specific to psychometrics and survey design, ECD is an approach that is only just being combined with principles of CBPR (Soland, Penuel, Farrell, & Wellberg, 2024). ECD is a principled approach that involves multiple stages to support the close correspondence between the claims about a given construct that a measure is designed to evaluate and the quality of the evidence that is used to support inferences drawn from resultant scores (Herman & Linn, 2015). A key strength of ECD is that it involves development of assessments in a manner that considers and collects validity evidence from the onset of the instrument design, not simply after the fact. ECD is built on the principle that a survey or other instrument is a measurement tool designed to be used to gain particular types of understanding (or, in ECD parlance, to make “claims” about a group of people being studied), and that a good instrument follows those claims and is in alignment with the beliefs, attitudes, opinions, and behaviors of the individuals being measured (Arieli-Atalli et al., 2019).

While ECD can involve many steps and stages, we focus on the three broad steps articulated by Ferrell et al., (2024), who also blended ECD with CBPR. First, instrument developers consult existing research, available measures and instruments, and input from stakeholders to define constructs of interest (Domain Analysis). Second, instrument developers gather observable evidence of what beliefs, attitudes, opinions, or behaviors are associated with different levels of development on those constructs (note that this activity, called Tool Design and Testing, can be highly participatory). Third, developers analyze how well the instrument produces measurable, valid, and reliable evidence for the construct (Observations) and provide guidance on

interpretation and use of results for different purposes (Interpretation and Use). ECD unfolds iteratively and collaboratively, with measurement instruments being refined as new insights emerge. A key element of ECD is that one does not start with the instrument itself, or an existing instrument. Rather, the design of an instrument flows directly from the attitudes, behaviors, and knowledge of the population that is the target of measurement, oftentimes in collaboration with those individuals being measured.

Increasingly, participatory research practices are being infused into ECD. For example, Farrell et al. (2024) used a participatory ECD approach to develop measures of research practice partnerships (RPPs). A key facet of their ECD process involved regular interaction with, and input from, stakeholders, including researchers and educational practitioners regularly engaged in RPP work. However, this melding of ECD and CBPR is still a relatively new phenomenon. In the following sections, we attempt to describe the ECD and CBPR blended process we used to develop a survey on the knowledge and beliefs of teachers and caregivers of children with autism in western Kenya. As this is an instrument development pilot study and not a validation study, in this article we present the survey development process and not the survey results.

### **Instrument Development Process Blending ECD and CBPR**

The instrument development process described below follows the activities related to ECD and infuses CBPR principles at each step. For example, during Domain Analysis, we developed a list of constructs and made construct maps in partnership with the Kenyan research team, our community advisory board, a group of village elders, teachers, and caregivers of children with autism from Eldoret, Kenya. Similarly, we collected evidence of what beliefs, attitudes, opinions, or behaviors are associated with the constructs of interest by having Kenyan team members brainstorm with members of the advisory board, all with relatively little involvement from Western

researchers. The survey items emerged from that brainstorming work. Following the survey item development, our Kenyan partners conducted cognitive interviews with a subset of the advisory panel to ensure the items elicited the information they were intended to capture. We detail these steps in the ECD process that blend CBPR principles in the next sections, with primary focus on the first two stages of ECD (Domain Analysis and Tool Design/Testing). A visual for this process is provided in Figure 1.

### **Step 1. Construct Development (Domain Analysis)**

#### ***Initial Construct Identification***

Building from our team's prior work and knowledge of the literature in autism education in Kenya and in LMICs, our international team met to identify initial constructs to be measured with this survey. By jointly developing initial constructs to be measured, we focused on CPBR principles of recognizing the local community as a unit and on identifying and prioritizing community-driven issues (Minkler, Blackwell, Thompson, & Tamir, 2003). In developing the Autism in the Context of Education – Kenya Survey (ACE-KS), we sought to explore knowledge and beliefs about autism from the perspectives of caregivers and educators. Initial constructs identified by the team included (a) knowledge of causes of autism, (b) beliefs about educational and life outcomes for autistic people, (c) knowledge of access to behavioral and educational services, and (d) beliefs regarding how autistic children should be taught and cared for. These initial constructs were then refined through a review of the literature, a review of existing measures and feedback from the community advisory panel. We detail that literature review in the supplemental materials. However, we found very few instruments related to autism/NDD developed specifically for sub-Saharan Africa, and none developed specifically for Kenya (Bakare et al., 2008, 2009).

### ***Initial Team Interviews/Discussion***

Following the review of existing measures, the first and second authors revised the constructs and shared those with the team for feedback and revision during weekly videoconference meetings held via Zoom. This CBPR process of iterative refinement of constructs supported participation and bidirectional learning from all team members and allowed us to focus more clearly on the community-defined constructs of interest. During our weekly meetings, each construct was presented and team members provided feedback and refinements to the construct with a focus on cultural relevance and appropriateness, and potential interpretation by Kenyan respondents. Team members offered suggestions on refining constructs, expanding constructs, and distinguishing potential overlapping constructs. Weekly team discussion continued for three weeks until the team reached consensus on the following constructs: (a) knowledge of autism and causes, (b) knowledge of diagnostic process, (c) beliefs related to optimal outcomes for children with autism, (d) beliefs about how autistic children should be taught, (e) beliefs about stigma and attitudes toward people with autism, and (f) how caregivers should care for autistic children.

### ***Construct Mapping***

Next, the team created a construct map for each construct of interest. A construct map is a theory-driven and thoroughly researched ordering of qualitatively different levels of a given construct focusing on one characteristic at a time (Wilson, 2009). At each level, the researcher might identify the attitudes, perceptions, beliefs, or knowledge of respondents that are unique to that level. In simple terms, a construct map defines what is to be measured in terms general enough to be interpretable to the public (including those taking the survey), but specific enough to guide the development of the instrument. While the idea of a construct map stems from the field of

achievement testing, it is also employed in survey design (e.g., Maul et al., 2008; Nuansri et al., 2016).

Creating a construct map yields many benefits in the instrument design process, detailed by Wilson (2009). First, the map can help ensure the construct is defined succinctly, including in terms of a developmental progression on that construct. Second and related, a construct map can provide insights into dimensionality. Concepts that are too complex to be captured straightforwardly in a single map can be broken apart into separate constructs and, therefore, maps. Third, the map can be used to streamline the item writing process. Specifically, once attitudes and behaviors at each qualitative level are articulated, items can be written to match those attitudes and behaviors. Fourth, one can ensure the items adequately cover the range of the construct by ensuring items correspond to every level of the map. (See Appendix A for two of our construct maps). Finally, once data are collected, one can check that the rank ordering of items along the map matches the rank ordering of items based on the proportion of actual respondents endorsing each item. Example construct maps for two of our constructs can be found in the supplemental materials.

The initial construct map was shared by the first author with the team and refined over the course of two Zoom meetings with Kenyan collaborators. Revisions focused on cultural adaptation of language. For example, the word “construct” was viewed by the team as being too abstract for future survey respondents and was replaced with the term “ideas.” Other adaptations to language included reducing jargon and technical terms. For example, the construct of “optimal outcomes for children with autism” was changed to “best possible life for children with autism.” The construct map was further culturally adapted along the dimensions of metaphors and content (Bernal et al., 1995) to align the constructs and behaviors with cultural values, customs, and practices.

Specifically, the construct of causes of autism underwent significant revision to not only encompass the wide range of local beliefs of disability etiology, but also to ensure respondents would not feel judged for selecting behaviors at the extreme ends of the continuum. The revised construct map was then shared with the Kenyan community advisory board to finalize construct development and construct mapping. The construct maps provided a blueprint for the CBPR process of brainstorming beliefs, attitudes, and behaviors associated with levels of each construct without constraining our community member participants in terms of the constructs being measured or what defines them.

## **Step 2. Instrument Design (Tool Design and Testing)**

### ***Community Advisory Board Convening***

The collaboration with a community advisory board was essential for our approach of blending CBPR into the ECD process. The advisory board helped ensure Kenyan research partners and community members were involved in all stages of the project from construct development to data collection, to data interpretation, and finally to results dissemination. By gathering community members, we focused on community autonomy and respect for their needs, values, strengths, and culture (Mikesell, Bromley, & Khodyakov, 2013).

We convened a community advisory board consisting of caregivers of children with autism, education leaders with backgrounds in special education, and community leaders from sub-counties within a large county in western Kenya. Kenyan team members led the nomination of individuals to serve on the community advisory board. A key element of CPBR is long-term commitment to partnership goals from all partners. The initial list was discussed and expanded with additional nominations of community members who had previously engaged with the team through community outreach and previous research activities demonstrating a commitment to

goals of understanding and supporting people with autism and their families. The community advisory board consisted of 13 individuals including four village chiefs and elders, three caregivers of autistic children, one autistic self-advocate, two head teachers/school directors of special schools serving children with disabilities, one member of the Ministry of Education, one pediatrician, and one program director of a teacher training program on autism.

The community advisory board met with the research team for a two-day meeting. The meeting was conducted in English, which is an official language in Kenya. During small group activities, some board members spoke briefly in Swahili, also an official language of Kenya. While the whole research team welcomed and encouraged use of both languages, board members often prompted each other to speak in English so that their thoughts could be understood by researchers who did not understand Swahili. The board meeting was structured based on Kenyan cultural practices to align with CBPR principles of recognizing community autonomy and building on community strengths and resources. Cultural practices included an opening prayer, individual and extensive self-introductions of persons in the meeting in order of position/hierarchy in the community, a morning tea break, and orientation to the project. During the introduction, the US and Kenyan principal investigators jointly presented on their history of collaboration and the research and outreach already undertaken in the county, the purpose of the project, and the goals for the community advisory board meeting. The first author then presented an overview of activities and processes to be undertaken during the meeting with respect to construct development and brainstorming beliefs, attitudes, and behaviors aligned with the construct maps.

The structure of the community advisory board meeting was culturally adapted according to the Ecological Validity Framework construct of concepts and context (Bernal et al., 1995). Specifically, to engage the community advisory board in construct development and construct



mapping, community advisory board members were organized into homogeneous groups based on their role of community leaders, education leaders, or family leaders, at the guidance of our Kenyan colleagues. This homogenous grouping was essential for ensuring all voices were heard and that different perspectives could be shared and recognized and aligned with the CBPR features of leveraging community strengths and promoting bidirectional learning among community advisory board members and between the board members and the research team.

Kenyan facilitators were present at each small group to facilitate discussion, record notes, and maintain focus on the topic at hand. Facilitators used a script developed collectively by the whole research team and rehearsed the script with the team prior to the advisory board meeting. (See Appendix A for session scripts.) A single prompt was provided to the small groups who then reflected on this prompt through dialogue and note taking on chart paper. These small groups then reported to the larger group. The first prompt requested board members to reflect on whether the initial constructs were comprehensive, distinct, and appropriate. In general, the group suggested only minor revisions to these construct definitions and maps (e.g., using synonyms for particular words that might be better understood in the community). A consensus across groups was quickly reached with approval of these initial constructs.

The community board members then worked in the homogenous small groups to brainstorm all the possible attitudes, behaviors, and levels of knowledge associated with each construct. We began by having community board members reflect on their own attitudes, behaviors, and knowledge, then asked them to broaden the brainstorming to include the same for members of their communities. For example, how might someone who sees little if any stigma attached to autism behave toward someone with autism and how might someone who has very negative attitudes and beliefs about autism behave towards someone with autism?

After each small group session, we reconvened as a whole group to share out the lists of beliefs/attitudes/behaviors/knowledge and to add any new beliefs/attitudes/behaviors/knowledge identified by the groups. With this more comprehensive list of beliefs, attitudes, behaviors, and knowledge, the whole group was asked to rank each entry on each construct from most to least common in the community. The intent of this process was to help situate these behaviors on the construct maps while prioritizing community autonomy. Examples of two of these construct maps are available in the Appendix.

In keeping with the CBPR framework, our team also wanted to ensure research benefits were fairly distributed across the community. Thus, at the conclusion of day two, the whole group was again convened where community board members reflected on the board meeting, offered suggestions for next steps, and shared ideas for how this project could be of mutual benefit to their communities, families, and to research. The community advisory board shared the importance of partnering with trusted community leaders for participant recruitment. They also shared the need for immediate community benefit while the survey is being developed and distributed. This discussion led to the creation of a virtual community for autism information and resource sharing on the platform WhatsApp.

Community advisory board members were essential in reviewing the constructs and brainstorming attitudes, knowledge, and behaviors associated with the constructs. Though this process was time-consuming, it improves the likelihood that the measures developed will elicit what is intended (Farrell et al., 2024). Further, when collaboratively generated by those with in-depth understanding of autism's role in the community (including in families and schools), the process constitutes a form of validity evidence regarding how constructs are represented in the survey. In short, the process helps ensure validity by design, in line with ECD principles.

### ***Item Development***

Item writing developed iteratively over several stages. First, in the days immediately after the stakeholder convening, we gathered as a research team to identify the behaviors, attitudes, and knowledge identified by the stakeholders that seemed essential to include in the survey. This process was driven by Kenyan team members, who had a better sense of the authenticity of the behaviors, attitudes, and knowledge being considered. Part of the conversation was ensuring that the behaviors, attitudes, and knowledge selected covered the range of the construct map for a given construct.

After that initial culling, subsets of the research team drafted items without response categories. The focus was on ensuring the item stems for each block of items were consistent and clear. Following multiple rounds of whole-team discussion about the items, especially their wording and clarity, response options were added for each block of items. Grounded in a CBPR iterative development process, we conducted several additional rounds of whole-team discussion about those items with response categories. This process yielded an initial draft of the survey scales for piloting.

### ***Piloting***

***Cognitive Interviews.*** We initially piloted the survey using a cognitive interview framework. The purpose of cognitive interviews is to gather information from respondents about their cognitive processes as they view the survey (Willis, 2005), and to then change the instrument as needed (Karabenick et al., 2007). Following the community advisory board meeting, two Kenyan research assistants trained in cognitive interviewing conducted two rounds of cognitive interviews using an interview protocol developed by the team (See Appendix B) and informed by guidance in cognitive interviews for instrument develop by Knafl et al. (2007). Because the survey

was intended to be administered to caregivers and educators, in each round of cognitive interviews, we interviewed a caregiver of a child with autism and a special needs educator or educational leader. The first round of interviews included two members of the community advisory board. The second round included two community members who had not engaged with our team in previous research or outreach activities. The aim of these interviews was to determine if potential survey respondents interpreted the survey items as intended. Cognitive interviews were administered individually by a Kenyan research assistant. Respondents were asked to “think aloud” as they reviewed the survey wording, response option format, range of response options for each item, and whether most teachers or caregivers would be able to answer each item accurately. Interviews were audio recorded and transcribed. Transcriptions were analyzed to identify themes related to clarity and cultural appropriateness of items and response options.

The survey was then revised based on these themes. Revisions pertained to the ordering of items, wording of individual items, and clarification and adjustment of response options. With respect to ordering, the demographics section was moved to the end of the survey and a rationale for why personal information was requested in this section was provided. Item wording changes were suggested to reduce potential confusion or to select a word that was more commonly used in the community. For example, “downtime” was replaced with “free time,” “evil spirits” was replaced with “witchcraft,” and “trade skills” was replaced with “vocational skills.” Respondents expressed concern with items with large number of response options and with differentiation between these options. For this reason, some response options were reduced from five to three or from four to two. The final survey included 84 items and six scales with response options ranging from binary to four-point Likert.

### ***Translation into Kiswahili***

The process of translating the survey from English to Kiswahili involved a forward and backward translation approach to ensure accuracy and cultural relevance. First, the survey was translated from English to Kiswahili using Google Translate, after which native Kenyan Kiswahili-speaking members of the research team reviewed and refined the translation to ensure clarity, appropriate terminology, and alignment with local linguistic nuances. Once finalized, the Kiswahili version was then backtranslated into English using Google Translate to assess whether the original meaning was preserved. The research team compared the back-translated version with the original survey, identifying any discrepancies and making necessary adjustments to the Kiswahili translation to improve accuracy. This iterative process helped ensure that the final Kiswahili survey maintained the intended meaning while being understandable and contextually appropriate for caregivers and educators in Kenya.

### ***Initial Psychometric Analyses and Results***

After completing cognitive interviews and making further refinements, then translating the survey, we piloted the instrument with approximately 30 participants. With this early data collection, we conducted basic psychometrics analyses and, to the extent appropriate, refined the survey once again before final administration for this project.

***Psychometric Analyses.*** While small sample sizes precluded us from engaging in comprehensive latent variable modeling to help establish the validity of the measure for its intended uses, we nonetheless conducted descriptive analyses of the item responses. Many (though not all) of these analyses are either rooted in classical test theory or are rough approximations of the parameters estimated in factor analytic models. We provide details on some of the results in the supplemental materials (See Appendix E).

One potential concern when designing a survey is that items responses will vary only minimally. For example, if nearly everyone selects Strongly Agree for an item, then the item responses will have low variability. To examine variability in the item responses, we plotted histograms of responses for each item and examined summary statistics such as standard deviations, skewness, and kurtosis. Items with particularly leptokurtic distributions (i.e., items that did not elicit much variability among responses) were considered for removal at this stage. We also collapsed response categories if they were not used (e.g., no one selected “strongly disagree”).

Next, we examined inter-item correlations, particularly among items from the same scale (i.e., measuring the same construct), which should be strongly correlated with each other. Similarly, correlations should be lower across constructs, which also speaks to the factor structure of the constructs being measured. We further examined item-total score correlations, separately within each scale and over the entire set of items. Item-total correlations quantify the association between an item and the total score (without that item) for all items in a scale. Higher item-total score correlations indicate that a given item is a more reliable indicator of the latent variable of interest. Items with negative or near-zero inter-item correlations within a construct and item-total score correlations less than .30 were considered for removal. Finally, we computed Cronbach’s alpha for each scale to quantify each scale’s internal consistency. Given the small sample size, no statistical result was alone used as grounds to remove an item; rather, we used these results as a mechanism for reviewing the items one last time with our Kenyan research team.

***Psychometric Results.*** In general, the results were very promising. Except for one survey scale, all Cronbach’s alpha values were above .70. Further, only two items produced item-total correlations that were below .30, which generated conversation about refinement or removal. As

for factor structure/dimensionality, inter-item correlations were indeed higher for items within the same scale than across scales.

Given our very small sample size, we did not remove items based on their psychometric properties alone. However, items flagged during the analyses were brought before the research team and we discussed why the item might be performing at a suboptimal level. For some items, we decided to leave them unchanged under the assumption that small sample sizes could have been a primary culprit. For other items, we made additional (and final) changes based on those conversations.

Perhaps the most substantive change resulted from examining the variability of the item responses. We found that, for the scale measuring caregiver priorities in caring for an autistic child, almost all respondents agreed that each task was important. This likely occurred because most of the priorities involved tasks of daily living (toilet training, providing shelter, taking the child to school, etc.). In response, we removed three items with very low variability in the item responses and replaced them with items related to slightly more “controversial” priorities. These new items were derived from data collected during the advisory board convening, and involved actions like helping the child (upon reaching adulthood) find employment and meet a potential spouse. The final version of the ACE-KS is available in Appendix F.

## **Discussion**

The purpose of this study was twofold. First, we provide details on the ACE-KS, a survey instrument designed to understand beliefs about, and knowledge of, autism among Kenyan caregivers and educators. Second and more broadly, by detailing our design process, we show how surveys can be constructed in participatory ways that emphasize validity by design and, we would argue, produce measures most likely to capture the reality of the participants being measured for

the survey's intended use. Thus far, we have described the specific approach to using both ECD and CBPR to design a survey to understand caregiver and teacher beliefs about, and knowledge of, autism in the context of Eldoret, Kenya. In the discussion we widen the lens to consider broad takeaways from this work. Specifically, we ask what aspects of that blending process did, and did not, go well in the pursuit of developing a survey that is valid for its intended purpose in the Kenyan context.

### **Which Aspects of this Process Worked Well?**

Throughout our description of the instrument design process, we describe the blending of ECD and CBPR. ECD aligns closely to the principles of CPBR. As such, we were able to embed CBPR approaches within the ECD design process. For example, something akin to the advisory board meeting has been used in some of our previous instrument design work (Farrell et al., 2024; Soland et al., 2024). However, in contrast to prior work, in our approach, the researchers with less knowledge of the local context (those from outside Kenya) played a minimal role in the advisory board proceedings. US team members served primarily as facilitators to keep the conversations on topic and to ensure activities were completed. The US team's focus on time management and linear progression through tasks did not always coincide with Kenyan culture. In fact, at one point, one of our Kenyan colleagues gently pointed out that we needed to let people spend more time speaking off topic in certain cases. "Sometimes," he said, "you just have to let people talk so they can heal." This redirection resulted in the US team stepping even further back from leading the advisory board meeting, while nonetheless monitoring that best practices for ECD and CBPR were being followed (and, in many ways, directly adhering to the principles of CBPR [Wallerstein, et al., 2017]).



The Kenyan advisory board not only provided the final definitions of the constructs being measured, but generated all the behaviors, attitudes, beliefs, and knowledge that became the items in the survey. That idea generation was curated almost exclusively through conversations overseen by our Kenyan research team. In line with the principles of ECD, we did not show up with survey items pulled from measures given in other LMICs for other purposes. Rather, we encouraged Kenyans with extensive knowledge of their communities populate the ideas that appear in the current survey instrument. Similarly, when conducting cognitive interviews to make sure the items were being interpreted as intended, we supported the Kenyan team in crafting a rough interview protocol that adhered to the strongest scientific principles, but they conducted the interviews independently, in their own language, and in ways they felt best probed likely survey takers on the interpretation of items.

While some of the final evidence on the success of this blended ECD-CBPR approach will come down to the ultimate psychometric analyses (the subject of a future study), we nonetheless see the potential for this blending to enhance our validity argument for the survey and its intended use. We feel relatively confident that the measure is appropriate not only for its intended use, but also for measuring Kenyan teachers and caregivers in the country because teachers and caregivers generated its content, sat for cognitive interviews, and advised at every step of the process. The ongoing, evaluative nature of the process is also in line with not only ECD, but broader validity theory (Messick, 1989). Our instrument was adjusted at multiple stages during the process based on stakeholder feedback, local researcher expertise, and psychometric analyses using pilot data (coupled with conversations with our Kenyan team on why some items might be performing worse than others). Further, the initial psychometric analyses provide additional evidence that our

theories about the measure—including how many constructs are being measured and whether items are reliable indicators of those constructs—are supported by the data.

### **Which Aspects of This Process Did Not Work Well?**

While many aspects of the blended ECD-CBPR approach worked well, we nonetheless encountered challenges along the way. For example, while our advisory board members are intimately familiar with their communities and consist of teachers and caregivers (the target of the survey), we could not be sure they represented the full range of perspectives we hope the survey will capture. We are interested in the beliefs and knowledge of parents who live in remote rural areas, may not have accepted their child's diagnosis, or both. While advisory board members hypothesized about the knowledge and beliefs of those parents based on experience and anecdote, we did not have the full range of potential survey takers in the room. Similarly, upon administering the survey, we often need to rely on connections with village elders to identify potential survey takers or, alternatively, to recruit from parents coming to the local hospital. Neither group involved in this snowball recruitment method is likely representative of the population of interest. In short, the absence of piloting in more diverse and, in particular, rural samples is a limitation for external validity.

Another challenge involved reconciling different opinions about the survey across the vast stakeholder input we solicited. In some cases, one group thought a particular belief or attitude about autism seemed plausible among likely survey takers while others did not. For example, some advisory board members had heard of children with autism being killed by parents in certain cases, while others had not heard of such behaviors and did not find them plausible. To address these discrepancies, we tried to only populate the survey with consensus views. Yet, in so doing, we may

have removed items that could be appropriately matched to respondents who hold particular views about autism.

Finally, some of the process strengths may also constitute limitations. For instance, given lower literacy rates in some locations, limited English, or both, our Kenyan research team administered the survey verbally to respondents. When respondents were confused by certain ideas, our Kenyan team members clarified those terms. On the one hand, such an approach is responsive to the context and, we believe, conducive to collecting the best data possible. On the other, we cannot rule out that such flexibility may introduce some form of bias into the results.

### **Conclusion**

This study presents a framework for developing culturally responsive survey instruments designed to capture the perspectives of caregivers and educators of children with autism in western Kenya. By integrating evidence-centered design (ECD)—a systematic approach to ensuring instruments yield valid, interpretable results—with the principles of community-based participatory research (CBPR), which prioritizes inclusive and collaborative engagement, we aimed to co-create a tool that reflects the lived experiences of those it is meant to assess. Rather than focusing solely on the final instrument, we emphasize the iterative, collaborative process that underpinned its development. Our findings illustrate how combining ECD and CBPR can strengthen the cultural relevance and validity of survey tools used for their intended purposes in low-resource settings. In doing so, this work offers a model for designing contextually meaningful instruments and supports broader efforts to advance understanding of autism in Kenya and other LMICs.

## **Declarations**

### **Author Contributions:**

Conceptualization: James Soland, Mandy Rispoli

Funding acquisition: Mandy Rispoli, Chelagat Saina

Investigation: Emmanuel Mwendwa, Carolyn Boke

Methodology: James Soland, Gillian Xu, Jamie Pearson, Megan McHenry

Project Administration: Mandy Rispoli, Mark Omollo

Supervision: Mandy Rispoli, James Soland, Chelagat Saina

Writing Original Draft: James Soland, Mandy Rispoli

Writing Review and Editing: Carol Stauffer, Tonia Hassinger, Rebecca McNally Keehn

**Funding:** The research reported in this article was made possibly by a grant #202400130 from the Spencer Foundation. The views expressed are those of the authors and do not necessarily reflect the views of the Spencer Foundation.

**Data Availability:** Data are restricted due to privacy or ethical considerations.

**Ethics Approval:** Both the University of Virginia Institutional Review Board for Social and Behavioral Sciences and the Moi Teaching and Referral Hospital Institutional Research and Ethics Committee approve this study. Informed consent was obtained in accordance with these approvals.

**Conflict of Interest:** The authors declare no conflicts of interest.

## References

- Afshari, R. (2013). What is the “Best Research” for low income countries?. *Asia Pacific Journal of Medical Toxicology*, 2(1), 1-1.
- American Psychiatric Association. (2023). Diagnostic and statistical manual of mental disorders (5th ed., text rev.; DSM-5-TR). American Psychiatric Publishing.
- Arieli-Attali, M., Ward, S., Thomas, J., Deonovic, B., & Von Davier, A. A. (2019). The expanded evidence-centered design (e-ECD) for learning and assessment systems: A framework for incorporating learning goals and processes within assessment design. *Frontiers in Psychology*, 10, 853.
- Bakare, M. & Munir, K. M. (2011). Excess of non-verbal cases of autism spectrum disorders presenting to orthodox clinical practice in Africa- a trend possibly resulting from late diagnosis and intervention. *South African Journal of Psychology*, 17(4), 118-120.
- Bayat, M. (2015). The stories of ‘snake children’: Killing and abuse of children with developmental disabilities in West Africa. *Journal of Intellectual Disability Research*, 59, 1-10.
- Bernal, G., Bonilla, J., & Bellido, C. (1995). Ecological validity and cultural sensitivity for outcome research: Issues for the cultural adaptation and development of psychosocial treatments with Hispanics. *Journal of Abnormal Child Psychology*, 23(1), 67–82.  
<https://doi.org/10.1007/BF01447045>
- Bitta, M., Kariuki, S. M., Abubakar, A., & Newton, C. R. J. C. (2017). Burden of neurodevelopmental disorders in low and middle-income countries: A systematic review and meta-analysis. *Wellcome Open Research*, 2, 121.  
<https://doi.org/10.12688/wellcomeopenres.13540.3>
- Chambers, N. J., Wetherby, A. M.,

- Stronach, S. T., Njongwe, N., Kauchali, S., & Grinker, R. R. (2017). Early detection of autism spectrum disorder in young isiZulu-speaking children in South Africa. *Autism, 21*(5), 518–526. <https://doi.org/10.1177/1362361316651196>
- Chansa-Kabali, T., Nyoni, J., & Mwanza, H. (2019). Awareness and knowledge associated with autism spectrum disorders among university students in Zambia. *Journal of Autism & Developmental Disorders, 49*(9), 3571–3581. <https://doi.org/10.1007/s10803-019-04044-7>
- Cloete, L.G., & Obaigwa, E. O. (2019). Lived experience of caregivers of children with autism spectrum disorder in Kenya. *African Journal on Disability, 8*, a435.
- Collins, S. E., Clifasefi, S. L., Stanton, J., Straits, K. J. E., Gil-Kashiwabara, E., Rodriguez Espinosa, P., Nicasio, A. V., Andrasik, M. P., Hawes, S. M., Miller, K. A., Nelson, L. A., Orfaly, V. E., Duran, B. M., & Wallerstein, N. (2018). Community-based participatory research (CBPR): Towards equitable involvement of community in psychology research. *American Psychologist, 73*(7), 884–898. <https://doi.org/10.1037/amp0000167>
- Elder, B., & Kuja, B. (2019). Going to school for the first time: Inclusion committee members increasing the number of students with disabilities in primary schools in Kenya. *International Journal of Inclusive Education, 23*(3), 261-279.
- Elliot, R. (2019). Autism in Kenya: A GeoPoll and Kaizora study. *GeoPoll Survey Reports, 16*, 1-16.
- Farrell, C. C., Penuel, W. R., Arce-Trigatti, P., Soland, J., Singleton, C., Resnick, A. F., ... & Schmidt, D. (2024). Designing measures of complex collaborations with participatory, evidence-centered design. *Frontiers in Research Metrics and Analytics, 9*(1), 1210547.
- Genovesi, E., Jakobsson, C., Nugent, L., Hanlon, C., & Hoekstra, R. A. (2022). Stakeholder experiences, attitudes and perspectives on inclusive education for children with

- developmental disabilities in sub-Saharan Africa: A systematic review of qualitative studies. *Autism*, 26(7), 1606–1625. <https://doi.org/10.1177/13623613221096208>
- Greer, J. M. H., Sood, S. S. M., & Metcalfe, D. R. (2022). Perceptions of autism spectrum disorder among the Swahili community on the Kenyan coast. *Research in Developmental Disabilities*, 131, 104370. <https://doi.org/10.1016/j.ridd.2022.104370>
- Harrison, A. J., Naqvi, N. C., Smit, A. K., Kumar, P. N., Muhammad, N. A., Saade, S., Yu, L., Cappe, E., Low, H. M., Chan, S.-J., & de Bildt, A. (2024). Assessing autism knowledge across the global landscape using the ASK-Q. *Journal of Autism & Developmental Disorders*, 54(5), 1897–1911. <https://doi.org/10.1007/s10803-023-05936-5>
- Herman, J. L., & Linn, R. (2015). Evidence-centered design: a summary. University of California: Los Angeles, CA, USA.
- Human Rights Clinic at University of Virginia School of Law. (2024). An overview of Kenyan Disability Law: The history, reality, and way forward for protecting the rights of persons with disabilities in Kenya. [White paper].
- Kakooza-Mwesige, A., Ssebyala, K., Karamagi, C., Kiguli, S., Smith, K., Anderson, M. C., Croen, L. A., Trevathan, E., Hansen, R., Smith, D., & Grether, J. K. (2014). Adaptation of the “ten questions” to screen for autism and other neurodevelopmental disorders in Uganda. *Autism*, 18(4), 447–457. <https://doi.org/10.1177/1362361313475848>
- Kamau L. (2017). Autism spectrum disorders (ASD) in Kenya: Barriers encountered in diagnosis, treatment and management. *Journal of Research in Pharmaceutical Science* 3(7), 1-11.
- Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazeovski, J., Bonney, C. R., De Greoot, E., Gilbert, M. C., Musu, L., Kempler, T. M., & Kelly, K. L. (2007). Cognitive

- processing of self-report items in educational research: Do they think what we mean?  
*Educational Psychologist*, 42(3), 139-151. Doi:10.1080/00461520701416231
- Knafl, K. Deatrick, J., Gallo, A. Holcombe, G., Bakitas, M., Dixon, J., & Grey, M. (2007). Focus on research methods: The analysis and interpretation of cognitive interviews for instrument development. *Research in Nursing and Health*, 30, 224-234.
- Mamah, D., Mutiso, V., Gitonga, I., Tele, A., & Ndeti, D. M. (2022). A population-based survey of autistic traits in Kenyan adolescents and young adults. *South African Journal of Psychiatry*, 28, Article 1694. <https://doi.org/10.4102/sajpsychiatry.v28i0.1694>
- Maul, A., Penuel, W. R., Dadey, N., Gallagher, L. P., Podkul, T., & Price, E. (2017). Measuring experiences of interest-related pursuits in connected learning. *Education Tech Research Development*, 65, 1–28. <https://doi.org/10.1007/s11423-016-9453-6>
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.  
<https://doi.org/10.3102/0013189X023002013> (Original work published 1994)
- Minkler M, Blackwell AG, Thompson M, Tamir H. (2003). Community-based participatory research: implications for public health funding. *American Journal of Public Health*, 93(8): 1210–1213. DOI: 10.2105/ajph.93.8.1210
- Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice*, 25(4), 6-20.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (1999). Evidence-centered assessment design: A brief overview. Princeton, NJ: Educational Testing Service



- Muuya, J. (2002). The aims of special education schools and units in Kenya: A survey of headteachers. *European Journal of Special Needs Education*, 17(3), 229-239. DOI: 10.1080/08856250210162130
- Nuansri, M., Tangdhanakanond, K., & Pasiphol, S. (2016). Development of multidimensional construct map of responsible citizenship of lower secondary school students. *Procedia-Social and Behavioral Sciences*, 217, 537-543.
- Plamondon, K. M., Brisbois, B., Dubent, L., & Bisung, E. (2021). Assessing how global health partnerships function: An equity-informed critical interpretive synthesis. *Global Health*, 17(1), 73. <https://doi.org/10.1186/s12992-021-00726-z>
- Schlebusch, L., Dada, S., & Samuels, A. (2017). Family quality of life of South African families raising children with autism spectrum disorder. *Journal of Autism & Developmental Disorders*, 47(7), 1966–1977. <https://doi.org/10.1007/s10803-017-3102-8>
- Smith, L., Malcolm-Smith, S., & de Vries, P. J. (2017). Translation and cultural appropriateness of the Autism Diagnostic Observation Schedule-2 in Afrikaans. *Autism*, 21(5), 552–563. <https://doi.org/10.1177/1362361316648469>
- Soland, J., Penuel, W. R., Farrell, C. C., & Wellberg, S. (2024). Developing a measure to evaluate education research-practice partnerships. *Research Evaluation*, 33, 1–15. <https://doi.org/10.1093/reseval/rvae042>
- Soland, J. (2021). Is measurement noninvariance a threat to inferences drawn from randomized control trials? Evidence from empirical and simulation studies. *Applied Psychological Measurement*, 45(5), 346-360.
- Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline.

- Journal of Evaluation in Clinical Practice*, 17(2), 268–274. <https://doi.org/10.1111/j.1365-2753.2010.01434.x>
- Taliep, N., & Florence, M. (2012). Evaluating the construct validity of the Kidscreen-52 Quality of Life Questionnaire within a South African context. *South African Journal of Psychology*, 42(2), 255–269. <https://doi.org/10.1177/008124631204200212>
- Tanis, E. S., (2021). *State of the state* [Report] University of Kansas. <https://news.ku.edu/sites/news/files/2024-06/State%20of%20the%20State%20FinalJH6.25.24%201.pdf>
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.
- Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 46(6), 716-730.
- World Bank Group: Fertility rate, total (births per woman). <https://data.worldbank.org/indicator/SP.DYN.TFRT.IN/>
- Xu, R., & Soland, J. (2024). Beyond group comparisons: Accounting for intersectional sources of bias in international survey measures. *International Journal of Testing*, 24(3), 230-258.