



# The words between us

HOW WELL DO ENUMERATORS UNDERSTAND THE  
TERMINOLOGY USED IN HUMANITARIAN SURVEYS?

**A STUDY FROM NORTHEAST NIGERIA**

November 2018



**TRANSLATORS**  
WITHOUT BORDERS

# Acknowledgments

Translators without Borders extends its sincere gratitude to the organizations and individuals that supported or participated in this assessment.

This study is funded by European Union Civil Protection and Humanitarian Aid (ECHO) through the International Organization for Migration (IOM).

Translators without Borders (TWB) envisions a world where knowledge knows no language barriers. The US-based non-profit provides people access to vital knowledge in their language by connecting non-profit organizations with a community of language professionals, building local language translation capacity, and raising awareness of language barriers. Originally founded in 1993 in France (as Traducteurs sans Frontières), TWB translates millions of words of life-saving and life-changing information every year. In 2013, TWB created the first-ever crisis relief translation service, Words of Relief, which has responded to crises every year since.

The European Union and its Member States are the world's leading donor of humanitarian aid. Relief assistance is an expression of European solidarity with people in need all around the world. It aims to save lives, prevent and alleviate human suffering, and safeguard the integrity and human dignity of populations affected by natural disasters and man-made crises. The European Commission ensures rapid and effective delivery of EU relief assistance through its two main instruments: civil protection and humanitarian aid. Through its civil protection and humanitarian aid operations department (ECHO), the European Commission helps over 120 million victims of conflict and disasters every year. With headquarters in Brussels and a global network of field offices, the Commission's civil protection and humanitarian aid operations department provides assistance to the most vulnerable people on the basis of humanitarian needs. For more information, please visit the European Commission's website.

*This document covers humanitarian aid activities implemented with the financial assistance of the European Union. The views expressed herein should not be taken, in any way, to reflect the official opinion of the European Union, and the European Commission is not responsible for any use that may be made of the information it contains.*



# Effective humanitarian programming depends on high-quality data

Data collection is the foundation of humanitarian programs, informing humanitarian response plans, program design, and safeguarding measures. If the data is inaccurate, humanitarian action may be misconceived, unaccountable, inefficient, and ultimately lacking in impact for those most in need. Language plays a critical but largely overlooked part in data quality.

Translators without Borders (TWB) is working to ensure language is not a barrier to gathering high-quality data. One approach is to measure how well enumerators understand the terminology used in the surveys they conduct. This can help gauge the accuracy of the data collected and identify areas for improvement in training and support for enumerators.

This report presents findings from three workshops on language barriers in humanitarian data collection in northeast Nigeria. TWB conducted the workshops between June and September 2018. These four-hour sessions involved guided discussions, group work, a survey, and a structured comprehension assessment.

Fifty-two enumerators and monitoring and evaluation (M&E) team members from three organizations took part in the comprehension testing (21 women and 31 men). Of 11 mother tongues reported, Hausa and Kanuri were the most common, spoken by 33 percent and 29 percent of respondents respectively. For each group, the test used 10 key terms from the organizations' recent surveys. Three terms were the same across the organizations, meaning 27 terms in total were tested. The surveys covered the WASH, shelter, food security, livelihoods, and demining sectors. Each individual was shown the terms in writing in the language used in the surveys - in this case English. We asked the participants to explain what each term meant in their own words. Their answer was scored as either 'correct' or 'incorrect' against a predetermined answer key.

## We found that:



**Enumerators report language challenges at every stage of the data collection process.**

**10%**

**Teams of enumerators understood at best 80 percent and at worst 10 percent of key terms from surveys they administer.**



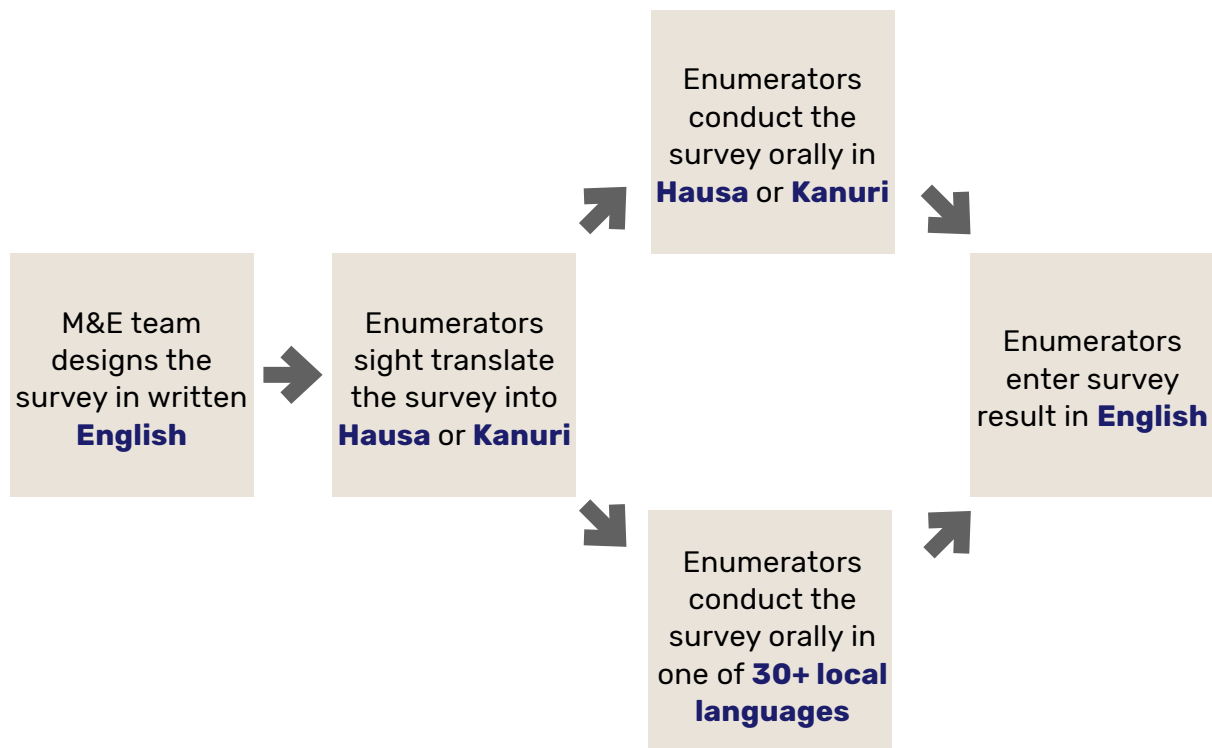
**Language is not a routine consideration in survey design.**



**Data quality inevitably suffers as a result.**

## Language impacts all stages of the **data collection** process

A survey during the workshops revealed that 84 percent of male and 95 percent of female enumerators feel that language is a major barrier in their work. Workshop discussions explored the challenges of conducting surveys in linguistically diverse northeast Nigeria. The flowchart below highlights the successive stages in which language impacts the data collection process.



Participants identified five key points where language barriers often manifest in the data collection process. A survey during the workshops revealed that 84 percent of male and 95 percent of female enumerators feel that language is a major barrier in their work. Workshop discussions explored the challenges of conducting surveys in linguistically diverse northeast Nigeria. The flowchart below highlights the successive stages in which language impacts the data collection process.

1. All but two people claimed to read English 'well' or 'very well,' even though none listed it as their mother tongue. However, many of the participants reported difficulty understanding complex or technical English terminology often used in the surveys. Abbreviations were especially problematic.

2. Many participants said that even when they understood the questions they often had difficulty sight-translating the questions accurately into local languages. Surveys are often only made available in English, placing the onus on data collectors to translate questions and answers on the spot. Many English concepts do not have precise equivalents in other languages.

concepts do not have precise equivalents in other languages.



3. Many of the participants did not speak a local language besides Hausa or Kanuri. Yet recent data from the International Organization for Migration (IOM) identifies over [30 primary languages](#) used at sites for internally displaced people. When enumerators encounter someone who does not speak Hausa or Kanuri, they often rely on a neighbor or relative (sometimes a child) to act as an informal interpreter. This presents a further series of challenges. Volunteer interpreters often struggle to communicate complex or technical questions accurately. They are also unfamiliar with the survey content, and may attempt to influence or speak for the interviewee. One enumerator explained that when he can't find someone to interpret, he communicates survey questions using "a very basic version of Hausa combined with gestures." This is clearly an ineffective mechanism for administering complex humanitarian surveys.



4. When using informal interpreting support, many enumerators also felt that the answers to survey questions are not accurately relayed. One man who had been working as an enumerator for over two years gave an example: “It’s very common to hear the [interviewee] and the interpreter having a long conversation. The interpreter will then turn to me and simply say ‘yes.’ When I hear this, I just have to wonder how much is being lost.”

5. Humanitarian surveys are designed to ensure monitoring and evaluation teams can process and analyze the data quickly. When open-ended questions are asked, the enumerator typically has to choose from a list of possible answers those that most closely match the response received. However, some enumerators said they do not always understand the possible answers in English, and they find it hard to identify which one best fits the response received. When that occurs, they select answers they are confident they understand.

This feedback indicates that, in the absence of language support, data collectors are resorting to a range of problematic practices to cope with language barriers. Some raise ethical concerns. Relying on neighbors and relatives – particularly children – to interpret surveys on topics such as gender-based violence can compromise confidentiality and expose those involved to risks. All the practices have negative implications for data quality. However, few are likely to be visible to monitoring and evaluation teams unless language is an explicit consideration.

“When I hear this,  
I just have to  
wonder how much  
is being **lost**.”

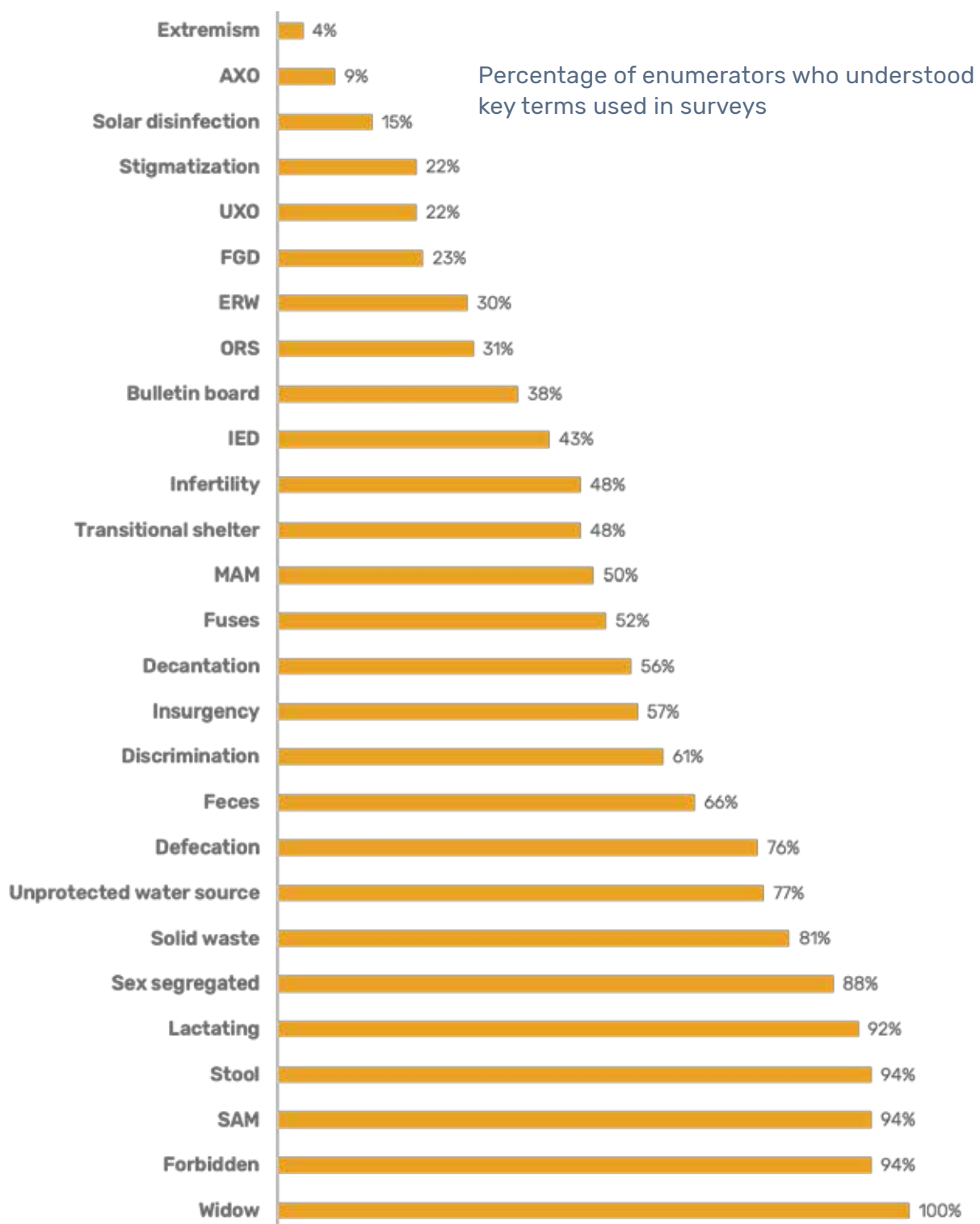
- A male enumerator, Maiduguri,  
Nigeria



## Measuring comprehension is an important step in ensuring **high-quality data**

Comprehension testing is a way of quantifying some of the challenges identified at the survey design and data entry stages (steps one and five in the figure above). If enumerators do not fully understand the English terminology used in the survey, it is fair to assume they are not accurately sight-translating the questions. If they do not understand all the terms listed as possible answers, we can assume the corresponding options are not accurately reflected in the survey results.

The graph below shows the percentage of respondents across the three organizations who understood the selected terms commonly used in surveys. Each word was tested with between 13 and 29 participants.



Abbreviations proved especially difficult. Just 31 percent of respondents understood 'ORS' for oral rehydration salts, a common treatment for dehydration. Less than half (43 percent) of the respondents correctly understood 'IED' (improvised explosive device). 'Extremism' had the lowest measured rate of comprehension: only one out of a group of 24 enumerators was able to explain what it meant.

Experience with a given organization or subject area did affect comprehension scores, but did not guarantee understanding. As the graph below shows, enumerators with more than one year's experience scored higher for comprehension than their less experienced colleagues across all three organizations. More seasoned data collectors at organization B scored more than five times as well as those with less than one year's experience. But even they only understood 55 percent of the terms they were shown.

Average comprehension test scores by enumerator experience





## Poor understanding of **terminology** can lead to underreporting and inaccuracy

Question 7.3 below is taken from one of the surveys used as the basis for this comprehension test. It illustrates how terminology forms the foundation of the entire data collection process. Here the enumerator is prompted to ask the respondent a question about their understanding of the effects of injury from explosive devices. The answer is open-ended and the enumerator enters the response by ticking all the boxes that apply.

Accurate data entry relies on the enumerator understanding several complex or nuanced English terms: isolation, dependency, burden, careless, luck, infertility, divorced, psychological suffering, second-class citizen, stigmatization, and discrimination. Words in blue in the table below were tested in this survey. Of the 24 enumerators shown these words, only 57 percent correctly understood 'insurgency,' 48 percent understood 'infertility,' and 22 percent understood 'stigmatization.' This implies that these three terms are being underreported as factors affecting individuals injured by explosive devices in this survey.

7.3 What would be the effect on an individual person if they were injured by one of these items? (Don't read the answers to the person, tick all that she / he mentions)		
<input type="checkbox"/> Isolation	<input type="checkbox"/> People will view them as careless	<input type="checkbox"/> Will be viewed as a second class citizen
<input type="checkbox"/> Wife can be rejected by family	<input type="checkbox"/> People will think they have bad luck	<input type="checkbox"/> Stigmatization
<input type="checkbox"/> Dependency	<input type="checkbox"/> Infertility	<input type="checkbox"/> Discrimination
<input type="checkbox"/> Can't go to school	<input type="checkbox"/> Women would be divorced	<input type="checkbox"/> No effect
<input type="checkbox"/> Can't travel for any reason	<input type="checkbox"/> Girls would not be able to get married	<input type="checkbox"/> Don't know
<input type="checkbox"/> People will view them as a burden	<input type="checkbox"/> Psychological suffering	<input type="checkbox"/> No response
<input type="checkbox"/> Other:		

# Four easy language solutions will improve the quality of humanitarian data

## 1. Use plain language to maximize comprehension.

Apply plain language principles when designing surveys. Maintain an average sentence length of 15-18 words. Choose clear and unambiguous terms, and use them consistently throughout the survey, even if that seems repetitive. Avoid abbreviations, or write the term in full on first mention, followed by the abbreviation in parentheses. For example, ‘improvised explosive device (IED):’

## 2. Translate surveys to ensure minority language speakers are not excluded.

English is not always the most effective language for conducting surveys. Of the enumerators in this sample, 19 percent preferred to read surveys in Hausa. Translating surveys can be time consuming, but online humanitarian surveying tools such as KoboToolbox make it possible to [design surveys in multiple languages](#). The results feed into a single database where staff who don’t speak those languages can still analyze them.

If translating surveys is not feasible, or enumerators prefer an English text, then a glossary of terminology specific to the sector or the organization is a helpful reference. Glossaries can be created as a separate tab on an Excel sheet where the survey is hosted, or as a [mobile application](#) available for offline viewing when needed. These should be used as part of training to prepare enumerators for administering the survey in local languages.

## 3. Make comprehension testing a standard survey design practice.

The survey design team is responsible for ensuring that enumerators understand it. Asking a group of enumerators, “Do you understand the wording of this question?” is not an accurate way to measure comprehension. Instead, a basic comprehension test such as the one used in the examples above can highlight potentially problematic terms or questions. TWB has developed a simple methodology for this test for download [here](#). Depending on how many words you choose to assess, the survey takes five to 10 minutes per enumerator. Testing comprehension during the design or early training phase leaves time to adjust the survey or training before data collection begins.

## 4. Provide data collectors with language support to communicate in local languages.

Preparation and translation in the planning stages cannot entirely avoid language challenges when a survey is administered. Enumerators also need language support when gathering data in a multilingual environment. Humanitarian organizations should provide vetted and trained interpreters to ensure the questions and answers are accurately communicated and to preserve confidentiality and safety.

Mobile data collection tools facilitate the use of audio recording to improve comprehension of questions and answers. Pre-recorded questions in multiple languages can be played to respondents if there is any confusion about the meaning. They can also speak open-ended answers into a recorder for data collection teams to transcribe. This has the added benefits of enabling subsequent verification and obtaining not only a more accurate record but also a more nuanced response.



For questions about this report or for language support, contact [info@translatorswithoutborders.org](mailto:info@translatorswithoutborders.org)

**Additional resources**

ACAPS, [How to design a questionnaire for needs assessments in humanitarian emergencies](#)

KoboToolbox, [Adding another language to your form](#)

OHCHR, [A human rights-based approach to data](#)

TWB, [Enumerator comprehension testing methodology](#)

TWB, [Humanitarian glossaries for northeast Nigeria](#)

TWB / PIN, [Rapid guide to localizing and translating survey tools](#)



**TRANSLATORS**  
WITHOUT BORDERS