

Researchers

Rachel Glennerster
UK Department for International Development (DFID)

Is It Time to Rethink How We Measure Women's Household Decision-Making Power in Impact Evaluation?

This is the first in a series of blog posts summarizing the discussions from a researcher gathering on measuring women's empowerment in impact evaluations co-hosted by Innovations for Poverty Action (IPA) and the AbdulLatif Jameel Poverty Action Lab (J-PAL) in May 2017.

By Rachel Glennerster and Claire Walsh

One of the first rules of thumb you learn about developing survey questions is that they should be specific and time-bound. In other words, it's better if a question is about a specific event or behavior rather than a vague idea so respondents are less likely to interpret it in different ways, and it should include a clear timeframe so that their responses are comparable.

Yet the most common survey questions for measuring women's participation in household decisions are not specific or time-bound. The questions, often adapted from USAID's Demographic and Health Survey (DHS), go like this:

"Who usually makes decisions about [healthcare for yourself]/ [major household purchases]/ [visits to your family or relatives]: you, your husband/partner, you and your husband jointly, or someone else?"

These questions are an important part of the DHS women's empowerment modules and are widely used by researchers and practitioners outside the DHS. At a recent IPA and J-PAL roundtable on measuring women's empowerment, more than half the researchers present had used these kinds of questions in impact evaluations before.

Several, however, had concerns. In practice, these questions can be hard to answer accurately because they are vague and require people to make a quick guess about general trends in decision-making at home. As one researcher put it, "They don't pass the 'Can I answer my own survey question?' test."

A simple alternative could be to ask people about how they would make a decision in a concrete scenario that's relevant in their context.^[i] Instead of asking, “Who usually makes decisions about your healthcare,” we could ask, “If your child is sick and needs immediate healthcare, but your husband is not home, what would you do?” Or, “If you ever need medicine for yourself (for a headache, for example), could you go buy it yourself?”

Our team asked both the standardized question and the more specific questions ... We got very different answers.

In an [evaluation](#) one of us (Rachel) is conducting on girls' empowerment in Bangladesh, our team asked both the standardized question and the more specific questions above. We got very different answers.

In response to the standard question, 16 percent of women said they usually make decisions about their healthcare alone or jointly with their husbands. Given this response, we would call this group more empowered—yet nearly a quarter of this group also said they could not take a sick child to the doctor until their husbands came home.

We also found discrepancies in the other direction: over half of the women who appeared disempowered according to the standard question said that they could take a sick child to the doctor on their own, and even more telling, could buy medicine for themselves.

These data should make us concerned that the standard questions are not picking up the characteristics we think they are. However, one test is not enough to jettison the DHS-style questions, which have other benefits.

First, there is value in asking questions in multiple countries over many years. For one, it allows us to benchmark a study to the broader literature, and to do meta-analyses of studies using a common indicator. They are also easier and more convenient to add to surveys than developing new questions.

The hope is that a more general question can fit many contexts, whereas specific questions may be more context-dependent. “Who decides whether and what type of health insurance to purchase for the family?” might be relevant in the United States, but not many other countries. “If you had a headache, could you purchase medication?” might provide a useful diversity of responses in Bangladesh, but not in the US, where most women can purchase cheap over-the-counter drugs.

So when we ask a general question like “Who usually makes decisions about your healthcare”, respondents arguably will adjust it to be about whatever the relevant health decisions are in their context. The downside is that we usually don't know exactly what kind of decision the woman is thinking about when she answers, and different women are likely thinking about different decisions. If we don't know the decisions she's thinking about, and whether they are important to her or not, is hard to judge whether any change we see in this general indicator is meaningful.

However, there have been cases when general questions led to more accurate responses

than specific ones. For instance, [de Mel, McKenzie, and Woodruff](#) found that simply asking small-scale entrepreneurs what their profits were was more accurate than asking them to report detailed revenues and expenses. Women and men may similarly have a good-enough sense of decision-making at home so that even if there is measurement error, the standard decision-making questions may still pick up something that's correlated with the underlying truth.

One indication that this could be the case comes from Markus Goldstein, head of the [Africa Gender Innovation Lab](#) at the World Bank, who shared an analysis comparing women and men's responses to the DHS decision-making questions at our recent roundtable. It is now available in a [working paper](#) by him and co-authors Donald, Koolwal, Annan, and Falb. They find that women who reported having greater sole or joint-decision making power were also more likely to own land, work outside the home, earn more than their husbands, and not condone domestic violence—outcomes we typically think of as signs of empowerment.[\[ii\]](#)

Yet even if responses to the standard household decision-making questions can be correlated with empowerment outcomes, it may not make sense to use them *in impact evaluations* without carefully working through whether they're relevant to the program being tested or the context.

Several researchers at the recent IPA and J-PAL roundtable observed that they have rarely seen significant changes in household decision-making indicators in their own or others' impact evaluations. It could be that these changes take longer than most evaluations. Another possibility is that the program wasn't likely or designed to change these decisions in the first place. When this is the case, it is probably better to use other questions more specific to the program.

It's also important to check that our survey questions are relevant to the context. Gender roles and dynamics can vary widely even within small geographic areas and change over time.

Beyond the program, it's also important to check that our survey questions are relevant to the context. Gender roles and dynamics can vary widely even within small geographic areas and change over time. Before starting an evaluation of an empowerment program, we typically conduct formative research in the field to collect qualitative and quantitative data about where women lack the ability to make strategic life choices that they want to make. Based on these data, we identify locally relevant indicators of empowerment and develop new survey questions to pick them up.

It can be valuable to use standardized questions in impact evaluations if they're relevant to the program and context, but we think it is equally, if not more important to include context-specific questions about what the women in our study communities can and want to change in their lives.

More broadly, a fruitful area for future measurement research is to conduct more validation exercises comparing different methods for asking about tricky concepts like agency and

decision-making (see a useful recent [example](#) from IFPRI that makes the case for calibrating questions to specific contexts). More validation exercises could help us identify whether there are improvements or additions to current standard questions that are worth making. For example, can we develop more specific questions that are relevant in many contexts—such as, “If your child is sick and needs immediate health care, but your husband is not home, what would you do: seek immediate care, ask for permission from someone, wait for your husband....”?

There will likely never be an effective one-size-fits-all set of survey questions to measure women’s decision-making power or agency, but we are optimistic about the potential to improve on current practice. We’re always looking for more research on this, so if you’re aware of useful validation exercises that have already been completed or are currently in the works, please send them to [Claire Walsh](#) and we’ll update this post with relevant links.

[i] Non-survey instruments like [structured community activities](#) or [purchase decisions](#) could be a useful alternative because they allow us to observe a real decision but these can be expensive to run.

[ii] Most definitions of empowerment emphasize agency and gaining the ability to make strategic life choices. Many draw on Sen’s concept of an agent as “someone who acts and brings about change, and whose achievements can be judged in terms of her own values and objectives,” (1999), and/or Kabeer’s definition of empowerment as “the process by which those who have been denied the ability to make strategic life choices acquire such an ability” (1999). Sources: Sen, Amartya. 1999. *Development as Freedom*. New York: Alfred A. Knopf. Kabeer, Naila. “Resources, agency, achievements: Reflections on the measurement of women's empowerment.” *Development and Change* 30, no. 3 (1999): 435-464.

September 06, 2017