



INTERNATIONAL
FOOD POLICY
RESEARCH
INSTITUTE

IFPRI

IFPRI Discussion Paper 02254

May 2024

Measuring Women's Empowerment in National Surveys

**Development of the Women's Empowerment Metric for National Statistical Systems
(WEMNS)**

Greg Seymour, Jessica Heckert, Agnes Quisumbing, Ruth Meinzen-Dick, Hazel Malapit,
Florescia Paz, Simone Faas, Emily Myers, Cheryl Doss, Sheela S. Sinharoy, Yuk
Fai Cheong, Kathryn M. Yount, Md. Zahidul Hassan, Md. Imrul Hassan, Sudhindra
Sharma, Pankaj Pokhrel, Mónica Dardón Sagastume, Shelton S. E. Kanyanda, Wilbert D.
Vundru, and Heather Moylan

Natural Resources and Resilience Unit
Poverty, Gender, and Inclusion Unit

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), a CGIAR Research Center established in 1975, provides research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition. IFPRI's strategic research aims to foster a climate-resilient and sustainable food supply; promote healthy diets and nutrition for all; build inclusive and efficient markets, trade systems, and food industries; transform agricultural and rural economies; and strengthen institutions and governance. Gender is integrated in all the Institute's work. Partnerships, communications, capacity strengthening, and data and knowledge management are essential components to translate IFPRI's research from action to impact. The Institute's regional and country programs play a critical role in responding to demand for food policy research and in delivering holistic support for country-led development. IFPRI collaborates with partners around the world.

AUTHORS

Greg Seymour (g.seymour@cgiar.org) is a research fellow in the Natural Resources and Resilience unit of the International Food Policy Research Institute, Washington, DC.

Jessica Heckert (j.heckert@cgiar.org) is a research fellow in the Poverty, Gender, and Inclusion unit of the International Food Policy Research Institute, Washington, DC.

Agnes R. Quisumbing (a.quisumbing@cgiar.org) is a senior research fellow in the Poverty, Gender, and Inclusion unit of the International Food Policy Research Institute, Washington, DC.

Ruth Meinzen-Dick (r.meinzen-dick@cgiar.org) is a senior research fellow in the Natural Resources and Resilience unit of the International Food Policy Research Institute, Washington, DC.

Hazel Malapit (h.malapit@cgiar.org) is a senior research coordinator in the Poverty, Gender, and Inclusion unit of the International Food Policy Research Institute, Washington, DC.

Florencia Paz (f.paz@cgiar.org) is a research analyst in the Poverty, Gender, and Inclusion unit of the International Food Policy Research Institute, Washington, DC.

Simone Faas (simone.faas@icf.com) was a research analyst in the Poverty, Gender, and Inclusion unit of the International Food Policy Research Institute, Washington, DC, when she wrote this work. She is currently a Survey Management Specialist in the Demographic and Health Surveys (DHS) Program at ICF, Rockville, MD.

Emily Myers (e.myers@cgiar.org) is a research analyst in the Poverty, Gender, and Inclusion unit of the International Food Policy Research Institute, Washington, DC.

Cheryl Doss (cheryl.doss@tufts.edu) is Professor of Economics at Tufts University, Medford, MA.

Sheela S. Sinharoy (ssinharoy@emory.edu) is Assistant Professor in the Hubert Department of Global Health at the Rollins School of Public Health at Emory University, Atlanta, GA.

Yuk Fai Cheong (ycheong@emory.edu) is Associate Professor of Psychology, Department of Psychology, at Emory University, Atlanta, GA.

Kathryn M. Yount (kyount@emory.edu) is Asa Griggs Candler Professor of Global Health, Hubert Department of Global Health and Professor of Sociology, Department of Sociology, at Emory University, Atlanta, GA.

Md. Zahidul Hassan (hassanzahidul@gmail.com) is Managing Director of Data Analysis and Technical Assistance (DATA), Dhaka, Bangladesh.

Md. Imrul Hassan (ihassan02@gmail.com) is System Analyst of Data Analysis and Technical Assistance (DATA), Dhaka, Bangladesh.

Sudhindra Sharma (sudhindra@ida.com.np) is Executive Director at Interdisciplinary Analysts (IDA), Kathmandu, Nepal.

Pankaj Pokhrel (pankaj@ida.com.np) is Research Associate at Interdisciplinary Analysts (IDA), Kathmandu, Nepal.

Mónica Dardón Sagastume (monicadardon@voxlatina.org.gt) is Project Coordinator at Vox Latina, Guatemala City, Guatemala.

Shelton S. E. Kanyanda (skanyanda@hotmail.com) is a Director of Agriculture and Economic Statistics, National Statistical Office, Zomba, Malawi.

Wilbert D. Vundru (wvundru@worldbank.org) is a consultant in the Development Data group of the World Bank, Washington D.C.

Heather Moylan (hmoylan@worldbank.org) is a senior economist in the Development Data Group of the World Bank, Washington D.C.

Notices

¹ IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by IFPRI.

² The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

³ Copyright remains with the authors. The authors are free to proceed, without further IFPRI permission, to publish this paper, or any revised version of it, in outlets such as journals, books, and other publications.

Contents

Abstract	vi
Acknowledgments	vii
Acronyms	viii
Introduction	1
Methods	6
Development of the WEMNS pilot questionnaire	6
Data collection	7
Indicator development	9
Quantitative methods	10
Women’s Empowerment Metric for National Statistical Systems	13
Domains of WEMNS	13
Intrinsic agency	15
Instrumental agency	16
Collective agency	17
Agency-enabling resources	18
Calculating WEMNS	20
Decomposing WEMNS	23
Results	25
Respondent characteristics	25
Aggregate results	26
Indicator-level results	27
Decomposition of WEMNS by subgroup	30
Robustness analysis	32
Missingness and non-response rates	32
Associations among WEMNS indicators	33
Pairwise associations of WEMNS with other outcomes	36
Estimated interview times	36
Conclusion	39
References	43
Appendix I	47
Appendix II. Description of cognitive interviewing in Bangladesh, Guatemala, Malawi, and Nepal	48
Appendix III	53

Tables

Table 1. WEMNS domains, indicators, definitions, and weights	13
Table 2. Respondent characteristics (% of women and men)	25
Table 3. WEMNS results by gender and pilot study setting	26
Table 4. Uncensored and censored indicator headcount ratios (% constrained in each indicator) by study setting and gender	28
Table 5. Decomposition of the WEMNS results by rural/urban status	32
Table 6. Non-response rates by indicator (% of women and men)	33
Table 7. Pairwise associations (Cramer's V) among WEMNS indicators	35
Table 8. Pearson's correlation between empowerment scores and other outcomes by pilot study setting and gender	36
Table 9. Median adjusted interview time (in minutes) by country and gender, interquartile range in parentheses	37
Table 10. Item sets, Sustainable Development Goal linkages, and sources of item sets included in WMENS	47
Table 11. Women and men sampled for cognitive interviewing by country	48
Table 12. Key cognitive interviewing findings by country	51
Table 13. Select survey items cognitively tested and revised survey items for WEMNS questionnaire	52
Table 14. WEMNS headline results by gender and country using alternative empowerment cutoffs	53

Figures

Figure 1. Kernel density estimates of empowerment scores by study setting and gender	27
Figure 2. Contribution of each indicator to disempowerment by pilot study setting and gender	29
Figure 3. Kernel density estimates for the adjusted overall interview times by study setting	37
Figure 4. Contribution of each indicator to disempowerment by pilot study setting and gender, based on 60% empowerment cutoff	53
Figure 5. Contribution of each indicator to disempowerment by pilot study setting and gender, based on 65% empowerment cutoff	54

ABSTRACT

Monitoring progress toward Sustainable Development Goal (SDG) 5—achieving gender equality and empowering all women and girls—remains challenging unless we incorporate women’s empowerment metrics into nationally representative and multi-topic surveys. To address this data gap, we designed the Women’s Empowerment Metric for National Statistical Systems (WEMNS) as a streamlined empowerment module suitable for the 50x2030 Initiative, a global partnership that aims to build capacity and close the agricultural data gap in 50 countries by 2030, as well as other large multi-topic surveys. WEMNS measures women’s and men’s empowerment and is applicable to urban and rural areas and a variety of livelihood strategies (farming, self-employment, wage labor) across countries in different stages of structural transformation. WEMNS is a counting-based, multidimensional index composed of four domains: intrinsic agency, instrumental agency, collective agency, and agency-enabling resources. Each domain is measured with binary indicators derived from question sets in the WEMNS module. In this paper, we describe the development and testing of WEMNS and its components, including: (1) WEMNS’s distinctiveness from other empowerment metrics; (2) the iterative approach used to develop and pilot the WEMNS module in Bangladesh, Guatemala, Malawi, and Nepal, using cognitive interviewing, phone surveys, and face-to-face surveys; (3) analysis of quantitative pilot data; and (4) a summary of the findings from the cognitive interviewing. The paper concludes with a discussion of lessons learned and possibilities for further development of WEMNS and other empowerment metrics.

Keywords: gender equality, women’s empowerment, national statistical systems, national surveys

ACKNOWLEDGMENTS

This work was undertaken as part of the Gender, Agriculture, and Assets Project Phase Two (GAAP2) and the CGIAR Research Programs on Agriculture for Nutrition and Health (A4NH) and Policies, Institutions, and Markets (PIM). Funding support for this study was provided by the Bill & Melinda Gates Foundation (BMGF) [Grant number: INV-008977], the United States Agency for International Development (USAID) [Grant number: EEM-G-00-04-00013-00], A4NH, and PIM. The Women's Empowerment Metric for National Statistical Systems (WEMNS) was specifically developed as an empowerment module for the 50x2023 Initiative, a global partnership to build capacity and close the agricultural data gap. Additionally, numerous individuals participated in consultations throughout the development of WEMNS. We are grateful to the 50x2023 Initiative partners and others for their feedback as we developed WEMNS. This work was conducted in partnership with Emory University, the University of Oxford, and the World Bank's Living Standards Measurement Survey and would not have been possible without the contributions of our data collection partners from Data Analysis and Technical Assistance Limited in Bangladesh, Interdisciplinary Analysts in Nepal, the National Statistical Office of Malawi, and Vox Latina in Guatemala. Several of these collaborators are co-authors on this report, and this work also benefited from the talented interviewers and other staff. Finally, we are grateful to the numerous survey respondents, without whom this work would not have been possible.

ACRONYMS

BBS	Bangladesh Bureau of Statistics
BIHS	Bangladesh Integrated Household Survey
CAPI	Computer-Assisted Personal Interviews
CSPro	Census and Survey Processing System
DATA	Data Analysis and Technical Assistance, Ltd.
DHS	Demographic and Health Surveys
ICT	Information communication and technology
IDA	Interdisciplinary Analysts
IFAD	International Fund for Agricultural Development
FAO	Food and Agriculture Organization of the United Nations
LMIC	Low- and middle-income country
LSMS	Living Standards Measurement Study
NSO	National statistical organization
NSO (Malawi)	National Statistical Office of Malawi
pro-WEAI	Project-level Women's Empowerment in Agriculture Index
pro-WEAI+HN	Pro-WEAI for health and nutrition
pro-WEAI+MI	Pro-WEAI for market inclusion
PSU	Primary sampling unit
SDG	Sustainable Development Goal
SDG5	Sustainable Development Goal 5
WEAI	Women's Empowerment in Agriculture Index
WEMNS	Women's Empowerment Metric for National Statistical Systems
WVS	World Values Survey

INTRODUCTION

The adoption of women's empowerment as part of Sustainable Development Goal 5 (SDG5) by the United Nations in 2015 motivated the need for better measurement. Monitoring global progress toward this goal will be impossible without relevant gender-disaggregated data, particularly in low- and middle-income countries (LMICs) where gender inequalities related to rights, resources, and responsibilities persist and limit women's agency, productivity, and well-being despite recent progress. Beyond tracking progress on SDG5, data is needed to identify factors that can contribute to empowerment and design effective interventions to improve empowerment and gender equality outcomes.

Widespread survey-based measurement of women's empowerment and gender equality requires a module that is relevant to a variety of cultures and contexts and is feasible to implement in large-scale, multi-purpose surveys, such as those administered by national statistical organizations (NSOs). For such a module to be useful for monitoring SDG5, it should be applicable to all adults within the country, both rural and urban, regardless of their livelihood strategies or life cycle stage and capture the full range of dimensions of empowerment. While no standardized tool can fully capture the nuance of women's experiences in each context, as a broad indicator of women's empowerment, the module should be comparable across countries and time (Yount et al. 2023). Additionally, the module should be relatively short and easy to administer, to incentivize its inclusion in national surveys. Finally, the tool should motivate and shape policy action; it should be able to be easily conveyed and understood by technical and non-technical audiences alike and serve as an aspirational target toward achieving SDG5.

A multitude of different measures of women's empowerment has emerged in recent years, but none is well suited to this particular purpose (Elias et al. 2021). Quisumbing et al. (2023) review the progress made since 2010 on the measurement of women's empowerment in the context of agriculture-dependent LMICs, pointing out great improvements in the direct, quantitative measurement of different aspects of women's agency. Multidimensional indices of women's empowerment based on the Women's Empowerment in Agriculture Index (WEAI; Alkire et al. 2013), including the project-level WEAI (pro-

WEAI; Malapit et al. 2019) and the specialized modules for health and nutrition projects (pro-WEAI+HN; Heckert et al. 2023) and market inclusion projects (pro-WEAI+MI; Malapit et al. 2023) have been widely applied.¹ Yet, because WEAI-based measures focus predominantly on the agricultural sector, they cannot capture the changes in women's empowerment as countries undergo structural transformation.

Other indices of women's empowerment that are not sector specific have been developed, including several constructed from Demographic and Health Survey (DHS) data, which provide wide geographic and historical coverage and are typically based on 20 or fewer survey questions (Ewerling et al. 2017; Rettig, Fick, and Hijmans 2020; Rettig and Hijmans 2022; Miedema et al. 2018). While such measures are useful for characterizing broad trends in the relationship between women's health and empowerment, their usefulness for informing policy outside this scope is limited. In practice, DHS-based indices of women's empowerment often focus on measures of resources and select dimensions of agency among women of reproductive age (Heckert and Fabric 2013). While extremely concise, these DHS-based measures do not offer the same breadth of measurement, across all dimensions of agency, which can be achieved with customized women's empowerment modules (Yount et al. 2020).

Despite this advantage, there are drawbacks to using measures based on customized survey questionnaires. Specifically, they tend to be costly to collect. Measures like those based on the WEAI often rely on many times the number of survey questions utilized by the DHS-measures discussed above, and require substantial time to implement, increasing costs and respondent burden (Alkire et al. 2013; H. Malapit et al. 2019). These high costs typically make such measures unsuitable for inclusion in nationally representative, multi-topic household surveys. Yet, those are exactly the types of surveys that appeal most to policymakers and NSOs because of their wide scope and geographic coverage. For these reasons, national multi-topic surveys are the best vehicle for wide-scale data collection on women's

¹ Since the launch of WEAI in 2012, WEAI-based measures have been applied by 251 organizations in 60 countries.

empowerment. No concise, yet comprehensive women's empowerment metric is routinely available to incorporate into these surveys.

To address this gap, we designed and piloted the Women's Empowerment Metric for National Statistical Systems (WEMNS). This paper documents our efforts to develop a measure of empowerment for women and men in urban and rural areas, with the flexibility to be relevant to a variety of livelihood strategies (e.g., farming, businesses, wage labor) across countries in different stages of rural and structural transformation. Because it is not lifecycle or livelihood specific, WEMNS allows for analysis of how women's empowerment differs across stages of structural transformation, as people engage in various sectors over time. This flexibility also makes it possible for WEMNS to measure national progress toward SDG5 across time and place.

We developed WEMNS primarily as an empowerment measure for the 50x2030 Initiative, which aims to build capacity and close the agricultural data gap in 50 countries by 2030. This initiative is implemented through a partnership of the International Fund for Agricultural Development (IFAD), the UN Food and Agriculture Organization (FAO), and the World Bank. Countries participating in the 50x2030 Initiative will have the option to include WEMNS in their surveys. We see potential for broader use of WEMNS in other large-sample and multi-topic surveys and have designed it with this complementary goal in mind. Having such a measure will allow users to generate meaningful data to track progress on the empowerment of women and men. WEMNS also can be used in smaller, more geographically focused surveys designed for research or impact assessment related to women's empowerment. Thus, in addition to the advantages of WEMNS being a concise instrument that captures multiple dimensions of empowerment, the ability to compare results across geographically focused impact assessments and nationally representative WEMNS data can provide insights about the external validity of the more focused studies that use WEMNS.

WEMNS is a multidimensional, counting-based index, similar in methodology to the Multidimensional Poverty Index (MPI; Alkire and Foster 2011) used by the United Nations Development Programme to track world poverty and the WEAI family of multidimensional empowerment indices. We

designed WEMNS following a normative approach, akin to the development of the MPI (Alkire et al. 2015a), and relied on multiple pieces of evidence throughout the development process. WEMNS is distinct from traditional psychometrically-derived scales, as its development did not include all steps nor consider all associated statistical evidence that the psychometric literature recommends (Boateng et al. 2018). Rather, throughout the development process we weighed qualitative and quantitative evidence, as well as theoretical, political, and practical concerns, to elevate the policy relevance and user-friendliness of WEMNS in the 50x2030 Initiative and beyond. Nevertheless, WEMNS underwent multiple phases of testing, including cognitive interviewing during questionnaire development, multiple rounds of piloting using alternative survey modalities, and comprehensive robustness analyses comparing alternative specifications for all index parameters. Additionally, the survey items used to construct WEMNS underwent psychometric assessment at sequential phases of questionnaire development, and those results are reported elsewhere (Sinharoy et al. 2023; Yount et al. 2023). As with any survey-based measure, there is a risk of measurement error in WEMNS. Even in the face of such error, so long as measurement error is not systematically associated with respondent characteristics, the WEMNS indicators still serve as meaningful proxies, which when measured together, provide valuable information about women's and men's agency.

Counting-based approaches have long been the standard in global poverty measurement, in large part because of how easy results based on these approaches (e.g., poverty headcounts) are to communicate (Atkinson 2019; Alkire and Foster 2011). The same is true for WEMNS, which conveys the percentage of women who are empowered and the intensity of disempowerment in a single headline figure. At the same time, this headline figure can be disaggregated into its component indicators, enabling users to easily identify the areas of disempowerment that require the most attention. The decomposable nature of WEMNS also facilitates comparisons of specific indicators across sub-populations, which is valuable for informing program design. Past experience with WEAI and pro-WEAI confirms that such findings, as well as the underlying methodology, are easy to communicate to policymakers, which is a major

consideration in getting the index adopted, the data collected and analyzed, and the findings ultimately influencing policy.

This paper focuses on the process of developing WEMNS. We describe our iterative approach. We present the index, discussing the conceptual foundation and formal definition for each WEMNS indicator. We present findings from the pilot face-to-face surveys in Bangladesh, Malawi, Nepal, and Guatemala. Finally, we conclude with reflections and recommendations for the future use of WEMNS in national surveys, national or sub-national longitudinal surveys, or impact assessments.

METHODS

Development of the WEMNS pilot questionnaire

To develop the questionnaire content, we reviewed all SDG Targets to identify links to women's empowerment and the relevant dimensions that should be included in WEMNS. Although only SDG5 focuses exclusively on the empowerment of women and girls, other SDG Targets have explicit gender components and identify ways in which women's empowerment may help achieve these goals. Appendix I shows the specific SDG Targets that were the focus of the WEMNS survey module, along with the item sets that we developed for each identified domain of empowerment and any existing items sets that informed what was used in WEMNS. Concurrently, we conducted semi-structured interviews with more than 70 key informants who were gender experts in Africa, Asia, and Latin America representing academic, civil society, and governmental agencies from 24 countries, primarily those eligible to participate in the 50x2030 Initiative. These interviews focused on identifying priority women's empowerment topics and current women's empowerment data collection efforts and the use of these data (Yount et al. 2023; Heckert et al. in development). Based on our review of the SDGs and consultations with key informants, we identified priority topics for WEMNS then drafted item sets for them. When possible, the item sets drew on existing material that had been successfully used or validated elsewhere, although often in a limited geographic scope. Additionally, we held two virtual workshops (July 2020 and March 2022) to which we invited interviewed national stakeholders, global gender experts, and partners in the 50x2030 Initiative. The first workshop further verified priority topics, and the second sought further refinement on the questionnaire content.

To strengthen and pilot the draft WEMNS module, we conducted cognitive interviewing, phone surveys, and in-person surveys (for details see Yount et al. 2023; Sinharoy et al. 2023). We collaborated closely with partners at institutions in Bangladesh (Data Analysis and Technical Assistance, Ltd., (DATA)), Guatemala (Vox Latina), Malawi (National Statistical Office (NSO of Malawi)), and Nepal (Interdisciplinary Analysts (IDA)). These partnerships were invaluable, especially in providing information on how content was interpreted linguistically and culturally. Cognitive interviewing is a

qualitative approach that aims to understand how respondents interpret and respond to survey questions and can identify and resolve potential sources of errors (Willis and Miller 2011b). We began this effort during the peak of the COVID-19 pandemic when in-person interviews were not feasible. Thus, cognitive interviewing and the first piloting and refinement of the surveys were conducted via phone surveys. In Bangladesh, Malawi, and Nepal we conducted cognitive interviewing prior to the administration of the phone surveys, and the results informed revisions to the phone surveys, which in turn informed revisions to the face-to-face surveys. A summary of the cognitive interviewing methods and results is found in Appendix II. Although the phone surveys were not part of the original plan and were a pandemic-related modification, conducting two phone survey rounds and a single face-to-face round incorporating refinements from the psychometric analysis was cost-effective. We conducted cognitive interviewing in Guatemala on the revised modules, after the completion of the phone surveys, and the results, along with preliminary results from the analysis of the phone-survey pilots (1 in Guatemala, 2 in the other three countries), informed the revisions to the in-person questionnaire.

Data collection

For the in-person surveys, which are the focus of the analysis in this paper, we aimed to sample 800 households in each country and to interview one woman and one man between the ages of 18 and 64 years in each household. The sampling design was intended to capture the socioeconomic, cultural, and geographic diversity of the populations in each country and to include urban and rural areas. In Guatemala, we also aimed to select a sample in which about half of the respondents spoke one of the two most common indigenous languages (K'iche' and Q'eqchi'). In all four countries we used a two-stage cluster stratified probability sample design and adhered to common principles to guide respondent selection. In all four countries, within each household, one woman and one man within the eligible age range were selected randomly from the household roster using a Computer Assisted Personal Interviewer

(CAPI) program (CSPPro in Guatemala and the World Bank Survey Solutions in the other countries)².

Thus, the respondents are not necessarily married to one another or the primary decision makers of the household but are drawn from all co-residing adult women and men. The process of identifying households, detailed below, differed slightly for each country.

In Bangladesh, we selected one division from each geographical region: Rangpur (north), Khulna (south), and Mymensingh (central). The three divisions were divided into rural and urban primary sampling units (PSUs).³ We selected 20 PSUs each from rural and urban areas. In each PSU, we conducted a household listing and randomly selected 20 households in each PSU from among those with at least one woman and one man in the eligible age range.

In Guatemala, using the most recent census, we identified municipalities where at least 40% of the population spoke either K'iche' or Q'eqchi' as their first language and dropped municipalities with security or logistical concerns, which limited the sampling frame to select municipalities in the Alta Verapaz, El Quiche, and Sololá districts. From the rural PSUs in these municipalities, we randomly selected 14 PSUs from among all PSUs in the municipalities where each language was dominant (28 total).⁴ Within each rural PSUs, we randomly selected 15 households from existing lists maintained by community leaders. Households that did not have both a woman and man in the target age range were dropped and replaced with the next eligible household. For urban areas in the same municipalities, in the absence of PSU boundary maps, based on the population size of each urban area and the average population per PSU, we calculated the corresponding number of PSUs for each urban area. In each urban area, we randomly selected a set number of geographic coordinates equal to the number of PSUs for the area. From among the list of coordinates, we randomly selected 13 sets of coordinates from municipalities

² Survey Solutions is a free software developed by the World Bank to design and implement surveys that rely on CAPI, computer-assisted telephone interviewing, and computer-assisted web interviewing. For more information, visit <https://mysurvey.solutions/en/>.

³ The rural sampling frame was based on the 2011 Population and Housing Census, and the urban sampling frame was based on the 2014 Census of Slum areas and floating populations. The Bangladesh Bureau of Statistics defined urban informal settlements as a cluster of compact settlements of 5 or more households. We considered urban informal settlements that have more than 60 households for constructing the sampling frame.

⁴ Survey rejection at the community level is common in rural Guatemala, and replacement communities were randomly selected.

where each language group was dominant (26 total) as starting points and used a random walk method to identify 15 eligible households.

In Malawi, we purposefully selected 12 districts from the central and southern regions. The districts were divided into rural and urban based on classifications by the NSO of Malawi. In each region, we randomly selected 36 PSUs with probability proportional to size. In each selected PSU, we conducted a household listing and randomly selected 11 households from among those with at least one woman and one man in the eligible age range.

In Nepal, the sample was selected from the Rupandehi and Kapilvastu Districts in Lumbini Province. PSUs were created by splitting or combining formerly delineated wards in these two districts into areas of roughly 200 households. Within districts, we randomly selected 27 PSUs, for a total of 54 PSUs across the two districts. In each selected PSU, we conducted a household listing and randomly selected 15 households from among those with at least one woman and one man in the eligible age range.

The final samples were as follows: Bangladesh 800 households, 800 men, and 800 women; Guatemala 811 households, 811 men, and 811 women; Malawi: 792 households, 792 men, and 791 women; and Nepal 813 households, 810 men, and 810 women.

Indicator development

The development of indicators for WEMNS involved several steps and considered multiple forms of qualitative and quantitative evidence. To begin, we combined related survey items (individual questions) to form composite (binary) indicators and proposed a range of potential cutoffs for each indicator based on the response structure(s) for the included items. Indicator cutoffs define the minimum level of achievement required for a respondent to be considered unconstrained in an indicator. Cutoffs are a necessary component of any counting-based index approach. For each indicator, we considered multiple potential item groupings, sometimes including all items from a particular question set, other times including a subset of items, and multiple cutoffs, ranging from those that most respondents would be expected to surpass to those that few respondents would be expected to achieve.

Drawing on guidance from (Alkire et al. 2015a), when grouping items together to form indicators and choosing cutoffs, we considered a range of factors, including qualitative evidence on the WEMNS module from expert review, key informant interviews, cognitive interviewing, and fieldwork experiences, including recommendations from the teams that implemented the survey, and quantitative evidence, described in detail in the following section. In addition, we bore in mind practical concerns related to implementation costs, interview times, and training requirements, as well as factors related to political will and feasibility, considering the processes of and stakeholders involved in the 50x2030 Initiative. Key to these political concerns was ensuring policy relevance—by broadly mapping indicators to the specific SDG Targets as discussed above—and communicability, e.g., the ease with which findings from WEMNS can be shared with diverse policymakers to inform resource allocation decisions or to support and monitor development interventions. No one piece of evidence factored primarily in our decision-making regarding the indicators. Rather, for each indicator, we considered all factors and selected the item grouping and cutoff that best balanced our concerns.⁵

Quantitative methods

Indicator development was informed by quantitative evidence on the psychometric properties of the WEMNS module, presented in Yount et al. (2023), and by additional statistical analyses, presented below, which assess the robustness of the item grouping and cutoff for each indicator.

Our choices regarding item groupings were informed by the dimensions and item sets assessed in the psychometric analysis of the Bangladesh, Malawi, and Nepal datasets (Yount et al. 2023). This analysis tested for measurement invariance across countries and between women and men to examine if item sets based roughly on all but one the WEMNS indicators performed similarly across these groups.⁶ When assessing cross-group measurement invariance using exact invariance testing approaches, it typically is recommended that item sets show evidence of scalar invariance across the groups, meaning that the factor

⁵ We recognize that one challenge of this holistic approach to indicator development is reproducibility, or the ability of others to implement independently the same process of indicator development to assess whether they would come to similar conclusions.

⁶ The item used to construct the secure property rights indicators were considered observed measures and, thus, inappropriate for factor analysis (Yount et al. 2023).

loadings and intercepts are similar across groups (Putnick and Bornstein 2016).⁷ The psychometric assessment of WEMNS relaxed this assumption and tested for at least partial scalar invariance (Yount et al. 2023). Because the counting-based approach followed in developing WEMNS differs fundamentally from psychometric approaches in its application of cutoffs to create binary indicators from item sets, we took general guidance from the psychometric assessment regarding the configural invariance of each, separately analyzed item set and then applied a different approach to indicator assessment. No study has established a gold standard level of invariance for a counting-based multidimensional index, such as WEMNS. Given that WEMNS is constructed from binary indicators, scalar invariance is an inappropriately strict criterion; the relative weight of different items is irrelevant if scales are not created. In such cases, configural invariance is a sufficiently appropriate criterion for consideration with multidimensional indices constructed from binary indicators. Judged in this manner, WEMNS performed adequately well in the psychometric analysis. For all indicators for which the underlying items were subject to invariance testing, only the items comprising the rejection of women's subjugation to sexual harassment indicator failed to achieve configural invariance across countries (Yount et al. 2023).

We applied a range of statistical methods to assess the robustness of the item grouping and cutoff for each indicator (Alkire et al. 2015b). When considering these results, we were guided by theoretical and practical principles relevant to the multidimensional measurement of women's empowerment (Alkire and Foster 2011; Alkire et al. 2013; Malapit et al. 2019). First, we compared indicator headcount ratios (the share of respondents identified as constrained in an indicator) and missingness (the share of respondents missing responses to one or more of the items used to construct an indicator) across the potential cutoffs for each indicator. In evaluating these cutoffs, the share of respondents identified as unconstrained in an indicator should not be so high as to preclude improvement over time (e.g., to avoid ceiling effects).

Second, missingness for all items used to construct indicators should be low to maintain sufficient sample

⁷ Ranging from the lowest to highest levels of invariance, configural invariance indicates that the same items load on the same factor across groups; metric invariance indicates equal factor loadings; scalar invariance indicates equal factor loadings and item intercepts; and residual invariance indicates the items also have equal residual variances across groups (Putnick and Bornstein 2016).

sizes for analysis. Low missingness was a priority because the Alkire-Foster approach requires complete information on all indicators for each respondent. Second, we estimated pairwise associations (Cramer's V) among WEMNS indicators and between indicators with other variables expected to be correlated with agency (see Yount et al. 2023 for a similar analysis using factor scores). Pairwise correlations among WEMNS indicators should be low to moderate to indicate that we are measuring distinct, but related agency constructs. Moreover, pairwise correlations should not be too high to suggest redundancy, which contributes to unnecessarily long interview times and would suggest the need to drop one of the indicators. Lastly, to support construct validity, indicators should be correlated with other measures expected to be correlated with agency.

WOMEN’S EMPOWERMENT METRIC FOR NATIONAL STATISTICAL SYSTEMS

Domains of WEMNS

While there are many possible definitions of (women’s) empowerment, we broadly follow Kabeer’s (1999) definition of empowerment as involving resources, agency, and achievements. We focus on agency, which we further classify into intrinsic, instrumental, and collective agency, corresponding to categories of generative forms of power: “power within”, “power to”, and “power with” (Kabeer 1994; Rowlands 1995; Ibrahim and Alkire 2007). Intrinsic agency includes concepts such as awareness of rights and self-efficacy. Instrumental agency is broadly conceptualized as enacting personal goals and often includes measures of decision making. Collective agency represents the power derived from working with others toward shared interests. We do not include indicators of achievements (such as health, nutrition, or freedom from violence) because these indicators are covered by other SDGs. The fourth domain, however, is “agency-enabling resources,” which addresses three fundamental types of economic resources. Although data on these resources, such as information and communication technologies (ICTs), financial services, and property rights may also be collected elsewhere, their relevance for women’s empowerment was underscored by the key informants interviewed in this project. While other types of resources are also important, these three categories have broader implications, as discussed below. The indicators, sub-indicators, and definitions are presented in Table 1.

Table 1. WEMNS domains, indicators, definitions, and weights

Domain	Indicator	Sub-indicator	Definition (number of required items)	Weight	
				Women	Men
Intrinsic Agency	Endorsement of women’s freedom in livelihood choices*		Believes that every woman should be free to choose to complete secondary school; work for pay; prioritize her work for pay over domestic duties; how to use any money that she earns; and purchase land, a house, or other valuable goods as demonstrated by fully agreeing with all five statements in this module (5)	1/12	Not included
			Believes that every woman should be free to choose when to get married; to divorce or end her marriage; whether and when to have children; and not to have any more children as demonstrated by fully agreeing with all four statements in this module (4)	1/12	Not included

	Rejection of women's subjection to sexual harassment*		Believes that it is never acceptable for a man to treat a woman as "lesser" because she is a woman; prevent a woman from doing certain kinds of work that she wants to do; spread unwanted rumors about a woman's sex life; try to have a romantic or sexual relationship with a woman against her wishes; and offer work-related benefits to a woman with expectation of receiving sexual favors (5)	1/12	Not included
Instrumental Agency	Influence over time allocation		Has a lot of influence on whether to spend their time on all activity categories in which they participated in the last seven days (from among household duties, caring for household members, going to the market to purchase essential items, non-agricultural work activities, agricultural production for sale, agricultural production for household consumption) (6 -12)	1/12	1/9
	Influence in spending decisions		Has a lot of influence over any source of income and over large household purchases (2)	1/12	1/9
	Influence in own health decisions		Has a lot of influence on own healthcare decisions (1)	1/12	1/9
Collective Agency	Participation in organizations		Participated in at least one of five possible types of community organization in the last 12 months (from among government councils or agencies, groups that provide local services, formal or informal savings or credit groups, groups related to livelihood activities, and other groups) (5)	1/12	1/9
	Leadership in organizations		Acted as a leader in at least one of five possible types of community organization in the last 12 months (same groups as above) (0 – 5)	1/12	1/9
	Confidence in women's/men's community engagement		Believes that women/men understand what is going on in their communities; can participate effectively in community activities; and can participate effectively in community decision-making; that it is important that women/men actively participate in local women's/men's issues; and that most community leaders would listen to women/men as demonstrated by full agreeing with all five statements in this module. (5)	1/12	1/9
Agency-enabling resources	Use of information communication technologies (ICT)		Daily use of mobile phone or internet	1/12	1/9
	Use of financial services and access to credit	Use of financial services	Has used at least one of four financial services (mobile money account, bank account, bank/ATM card, credit card) in last 12 months (4)	1/24	1/18
		Access to credit	Believes would be able to take a loan from at least one of five formal or semi-formal lending sources (bank/formal financial institution, cooperative, group-based micro-finance, informal credit/savings groups, other NGO program) (4)	1/24	1/18
	Secure property rights	Documented property rights	Name listed on land rights document for at least one land parcel or dwelling (7)	1/36	1/27
		Secure tenure rights	Not at all likely to involuntarily lose ownership or use rights to at least one land parcel or dwelling (2)	1/36	1/27
Secure transfer rights		Able to sell or bequeath at least one land parcel or dwelling (4)	1/36	1/27	

Note: * Indicators are not calculated for men; 56-67 items required for women; 42-53 items required for men.

Intrinsic agency

Intrinsic agency or “power within” is critical reflection that allows individuals to become aware of their own ability to make choices and is related to Freire's (1970) concept of critical consciousness and awareness of systems of oppression (Kabeer 1994; Rowlands 1995). WEMNS includes three indicators of intrinsic agency, each of which is only calculated for women.

The first indicator, *endorsement of women's freedom in livelihood choices*, captures the belief that every woman has the right to attend school and pursue a livelihood as she desires. Specifically, a respondent is unconstrained in this indicator if she fully agrees (as opposed to partially agrees or disagrees) with five statements related to this domain. The second indicator, *endorsement of women's freedom in family formation choices*, is constructed similarly from four questions reflecting women's belief in their rights related to marriage and childbearing.

The third indicator, *rejection of women's subjection to sexual harassment*, measures women's belief that sexual harassment is not acceptable. Specifically, empowerment on this indicator is achieved if she indicates that each of the five behaviors is never acceptable.

For each of these three indicators, we considered alternative indicator definitions, including cutoffs defined by a minimum mean score and less strict thresholds, such as only partially agreeing with each of the items in the first two indicators and considering sexual harassment behaviors acceptable under some circumstances. Conceptually, full support for these choices, without caveats and exceptions was deemed to be an important aspect of empowerment. Additionally, the selected cutoffs allow for improvements over time and help avoid potential ceiling effects.

To have comparable indicators for men would have required asking men if they believe men should have freedom to make choices about livelihoods and family formation, or if it is acceptable for men to be subjected to sexual harassment. Based on previous fieldwork experience, we expected that asking men these questions would have been problematic and unlikely to show much variation across men. This expectation was particularly true for sexual harassment—the concept of sexual harassment of men feels strange to many respondents. Because these questions were not asked, there are no corresponding

indicators of intrinsic agency for men. We did ask men whether they thought that women should have freedom of choice about livelihoods and family formation, and the acceptability of sexual harassment of women, but we do not use them to calculate men's empowerment. While men's responses to these questions may provide useful information about the enabling environment for women's empowerment, such as men's progressive support for women's rights or men's awareness of women's rights, they do not measure men's empowerment. Future research should consider designing alternative questions about men's perceived vulnerabilities and ability to overcome those vulnerabilities, toward the development of an equivalent set of intrinsic agency indicators for men.

Instrumental agency

Instrumental agency, or “power to,” is typically measured by participation in key decisions. WEMNS includes three instrumental agency indicators. *Influence over time allocation* reflects whether individuals had a lot of influence on how they spent their time during the past seven days across six broad categories of paid and unpaid work activities.⁸ In considering alternative definitions, cutoffs based on having “at least some” influence (as opposed to “a lot”) led to very few respondents being identified as constrained (<30% in all cases) for all subgroups (rural and urban women and men in each country), which we considered unreasonable given the global disparities in time use, especially unpaid care and domestic work. “At least some influence” covers a wide range, and may be subject to social desirability bias, where respondents do not want to say they have (absolutely) no influence (Meinzen-Dick et al. 2019; Nepa School of Social Sciences and Humanities 2017). We also piloted questions on how much influence individuals had over whether to not participate in the activities, but do not include them in the WEMNS indicators, as they were poorly understood according to field team reports.

The other two instrumental agency indicators are *influence in spending decisions*, for which individuals are unconstrained if they report a lot of influence over a source of income and over large

⁸ This item set differs from WEAI and pro-WEAI, which use time use data collected via 24-hour recall-based time diary. Such approaches take considerable time to implement, and qualitative research showed that it was often not so much the total hours worked that mattered, but perceived control over their time (Seymour, Malapit, and Quisumbing 2020; Eissler et al. 2022; Sinharoy, Cheong, et al. 2023). Thus, for WEMNS we refocused attention on a shorter module related to control over their time.

household purchases and *influence on own healthcare decisions*, which is unconstrained if individuals report a lot of influence over their own healthcare decisions. Like other indicators, alternative definitions in which “at least some” influence and mean scores were considered, led to extremely few women respondents being identified as constrained, which is inconsistent with current evidence from the pilot countries (Malapit et al. 2014; Quisumbing et al. 2022) We also considered combining these two indicators, but since they refer conceptually to different domains of decision making, we determined that they were better suited as separate indicators, particularly because combining them as a single indicator would define someone with influence in one area, but not both, as disempowered.

Collective agency

Collective agency, or “power with,” is a critical component of empowerment beyond the individual and interpersonal levels and reflects power gained from belonging to a group with shared interests. Within this domain, *participation in organizations* measures whether respondents are part of groups or organizations from which they may derive collective agency. *Leadership in organizations* is derived from the same module and measures whether they acted as a leader in these organizations from which collective agency is derived. Respondents are unconstrained in *Participation* if they participated in at least one type of community organization in the past 12 months and in *Leadership* if they acted as a leader in at least one type of community organization in the past 12 months.⁹ We considered indicator cutoffs based on belonging to at least two types of groups, but decided against them, both because nearly all respondents would have been constrained and because the questions ask about broad types of groups (not specific groups), which means that participation or leadership in multiple groups of the same type would not be counted. Moreover, it is not conceptually clear that being a member of two groups is necessarily more empowering than being a very active member of one group.

⁹ The interrelation between how these two indicators is constructed means that someone who does not belong to a group will be constrained in both indicators. While we normally try to avoid this type of correlation between indicators to maintain a more balanced index, in this instance, we believe it is justified due to conceptual significance of group leadership.

Confidence in women's/men's community engagement measures whether respondents believe that women (or men) are capable of participating actively in the community. The five statements that constitute this indicator are phrased differently for women and men. Women are asked about their beliefs regarding women's community engagement, and these items have been validated among women in Bangladesh (Yount, Khan et al. 2020). Men are asked a similar set of items about their beliefs regarding men's community engagement. We considered cutoffs defined based on a minimum mean score, as well as on partial agreement with the statements. For reasons like those described above for other indicators, we selected the threshold of fully agreeing with each of five statements for conceptual reasons and to avoid potential ceiling effects, as these indicators potentially improve across time.

Agency-enabling resources

Kabeer (1999) identifies the importance of material, human, and social resources for enabling individuals to exercise agency. Of these, we focus on economic resources. Under this domain, *use of ICT* indicator reflects daily use of mobile phone or Internet. This indicator is consistent with SDG 9.c: Significantly increase access to information and communications technology and strive to provide universal and affordable access to the Internet in least developed countries by 2020. We considered alternative definitions that might reflect a broader range of technologies or less frequent use, but we focused on Internet and mobile phone usage because these technologies allow for two-way communication, not just for receiving information. Two-way communication is critical for breaking the disempowerment effects of isolation. Abubakar and Dasuki (2018) show how ICTs can contribute to women's participation in not only economic, but also social and political activities, and that simply having access to ICTs is not enough when they are not used. We settled on daily use, because cutoffs based on less frequent usage resulted in very low levels of constraint, increasing the risk of ceiling effects over time.

Use of financial services and access to credit is comprised of two sub-indicators: *use of financial services* and *access to credit*. Such resources are key to economic participation. Respondents are

unconstrained in *use of financial services* if they have used at least one type of financial service in the past year. We considered having at least one type of financial service in the past year (but not necessarily use), as well as having two financial services and using one in the past year. We determined that use itself was important, as un-accessed accounts likely do little to enable agency. We elected not to consider multiple types of financial services, as these are often substitutes for one another. *Access to credit* indicates a belief that the respondent would be able to take a loan from at least one formal or semi-formal lending source. We considered including informal sources (friends, relatives, or informal lenders) as part of the indicator, but did not. Respondents had high levels of borrowing from friends and relatives, but the poor likely have similarly poor friends and family, suggesting these loans are typically small; moreover loans from friends and family are commonly used to smooth consumption, as opposed to means of investments (Garikipati et al. 2017). At best, taking these types of loans may temporarily improve women's influence in the household, but only because they take these loans on behalf of their husbands to save his honor and status (Garikipati et al. 2017). Additionally, some informal lenders, who serve as a last resort, may operate with predatory practices that are disempowering, and borrowing from these types of sources is often considered degrading (Guérin, D'Espallier, and Venkatasubramanian 2013).

The *secure property rights* indicator is comprised of three sub-indicators: *documented property rights*, *secure tenure rights*, and *secure transfer rights*. All three of these refer to either a parcel of land or a dwelling. Although we asked about multiple parcels and dwellings in the pilot surveys, the indicators are defined based on having these rights for at least one property, either land or dwelling. This approach is in-line with the SDG 1.4.2 definitions and allows for the indicator to be relevant for urban and rural respondents. "A durable shelter is one of the most basic requirements for people to be able to organize their thoughts, discipline their actions and undertake plans and programs for creative pursuits" (Huq and Sultan (1991), cited in Kabeer (1994), p. 239). *Documented property rights* and *secure tenure rights* follow the methodology used for SDG Indicator 1.4.2.¹⁰ However, while the SDG indicator is limited to

¹⁰ Proportion of total adult population with secure tenure rights to land, with legally recognized documentation and who perceive their rights to land as secure, by sex and by type of tenure.

land, we have broadened the WEMNS indicator to include dwelling so that it would be applicable in urban areas. *Secure transfer rights* was added to reflect whether respondents can sell, transfer, or bequeath their land or dwelling.

Calculating WEMNS

Below we describe the steps to compute WEMNS using similar notation as in previous work on the multidimensional measurement of poverty and women's empowerment (Alkire and Foster 2011; Alkire et al. 2013; H. Malapit et al. 2019).

(i) **Identify constraints.** For each of the indicators and sub-indicators described above, a person is identified as constrained or unconstrained. For each indicator j , we define an indicator cutoff, z_j , as the minimum cutoff required to be considered unconstrained in the indicator. Table 1 lists the definitions used to calculate all indicators. Person i is considered unconstrained in indicator j if their achievement, x_{ij} , is greater than or equal to the indicator cutoff (i.e., $x_{ij} \geq z_j$). Person i 's lack of constraint in indicator j is denoted as $g_{ij}^1 = 1$ if $x_{ij} \geq z_j$ and $g_{ij}^1 = 0$, otherwise.

(ii) **Aggregate constraints.** We calculate an empowerment score for each person by summing across the weighted indicators. Person i 's empowerment score, e_i , is equal to $\sum_{j=1}^d w_j \times g_{ij}^1$, where d is the number of indicators included in the empowerment score and w_j is the weight assigned to indicator i , such that $\sum_j w_j = 1$. In WEMNS, weights are distributed equally across the four domains and distributed equally across the indicators within each domain. The motivations for this choice include ease of construction, communication, and transparency. Equal weights by domain are straightforward to calculate and can be communicated easily to audiences irrespective of technical expertise. In addition, explicitly defining the weighting scheme used in WEMNS—as opposed to implicitly defining them through a data-driven approach—opens the methodology to public scrutiny and ongoing debate. The act of assigning weights to different indicators requires researchers to make value judgments about each indicator's importance for women's empowerment. The explicit weighting scheme used in WEMNS ensures these value judgments are fully transparent and standardized across contexts.

As noted above, not all indicators are suitable for use with men. Women's empowerment scores are based on all 12 indicators. Men's empowerment scores are based on 8 indicators, as shown in Table 1. Because fewer indicators are used to calculate men's empowerment scores, the indicator weights are not the same as for women.

Empowerment scores range from zero to one, with higher values reflecting greater empowerment (i.e., a greater share of weighted indicators for which a person is not constrained). Empowerment scores can be expressed equivalently in terms of disempowerment by summing the weighted indicators. Person i 's constraint in indicator j is denoted as $g_{ij}^0 = 1$ if $x_{ij} < z_j$ and $g_{ij}^0 = 0$, otherwise. Person i 's disempowerment score, c_i , is equal to $\sum_{j=1}^d w_j \times g_{ij}^0$, where d and w_j are defined as above, or simply: $c_i = 1 - e_i$. Empowerment scores cannot be calculated for persons with missing information on any of the required indicators. Such persons are, thus, excluded from the study sample (see discussion of missingness below).

(iii) **Identify who is empowered.** We identify each person as empowered or disempowered based on their empowerment score. We define an empowerment cutoff, k , as the minimum empowerment score required to be considered empowered. Person i is considered empowered if their empowerment score is greater than or equal to the empowerment cutoff (i.e., $e_i \geq k$). As WEMNS is still in the pilot phase, we are considering a range of empowerment cutoffs, between 60% and 70%. The final empowerment cutoff will be set once WEMNS has been implemented at scale.

(iv) **Compute disempowerment and empowerment headcount ratios for study sample.** The disempowerment headcount ratio, H_p , reflects the prevalence of disempowerment within the study sample. It is equal to q/n , where q is the number of women or men identified as disempowered and n is the total number of women or men, respectively, in the study sample.

The disempowerment headcount ratio can be expressed equivalently in terms of empowerment. The empowerment headcount ratio, H_e , is equal to r/n , where r is the number of women or men identified as empowered, and n is defined as before, or simply: $H_e = 1 - H_p$.

(v) **Compute intensity of disempowerment for study sample.** To identify constraints to empowerment more easily, we shift measurement focus to the situation of the disempowered. We do this by censoring empowerment scores based on whether people are empowered or disempowered. This allows us to distinguish between commonalities among the disempowered (e.g., indicators in which they are more likely to be constrained) and commonalities among the empowered, which would not otherwise be possible. In other words, censoring helps to focus attention on those who are most vulnerable and stand to benefit the most from policy interventions. Person i 's censored disempowerment score, $c_i(k)$, is equal to their disempowerment score if they are disempowered (i.e., if $c_i > 1 - k$, then $c_i(k) = c_i$) and equal to zero if they are empowered (i.e., if $c_i \leq 1 - k$, then $c_i(k) = 0$).

The intensity of disempowerment, A_p , is reflected by the average disempowerment score among disempowered women or men and calculated as:

$$A_p = \frac{\sum_{i=1}^n c_i(k)}{q}.$$

Alternatively, the intensity of disempowerment can be expressed in terms of empowerment. Person i 's censored empowerment score, $e_i(k)$, is equal to their empowerment score if they are disempowered (i.e., if $e_i < k$, then $e_i(k) = e_i$) and equal to zero if they are empowered (i.e., if $e_i \geq k$, then $e_i(k) = 0$). The average empowerment score among disempowered women or men can then be calculated as:

$$A_e = \frac{\sum_{i=1}^n e_i(k)}{q},$$

or more simply as: $A_e = 1 - A_p$.

(vi) **Compute WEMNS for the study sample.** While we present WEMNS in terms of empowerment, it is calculated more easily based on the disempowerment headcount ratio, H_p , and intensity of disempowerment, A_p , as shown here:

$$WEMNS = 1 - (H_p \times A_p),$$

or expressed equivalently as:

$$WEMNS = H_e + (H_p \times A_e).$$

Decomposing WEMNS

An advantage of the counting-based approach is that WEMNS can be decomposed to assess the contributions of specific indicators to disempowerment (or empowerment) or the contributions of different subgroups to the overall level of disempowerment within a particular sample or population. Such decompositions can be especially useful for understanding which indicators drive disempowerment and for uncovering broad patterns of disempowerment within the study sample as well as differences among key sub-groups. As such, our guidance focuses primarily on how to decompose WEMNS in terms of contributions to disempowerment; however, we also note how the formulas can be adjusted for decomposition in terms of contributions to empowerment.

(i) **Decompose WEMNS by indicator.** The contribution of each indicator to disempowerment (or more specifically: $H_p \times A_p$) is based on the censored indicator headcount ratios. The censored indicator headcount ratio, $h_j(k)$, is the proportion of women or men in the study sample who are both disempowered and simultaneously constrained in indicator j . Formally:

$$h_j(k) = \frac{\sum_{i=1}^n g_{ij}^0(k)}{n}$$

The absolute contribution to disempowerment of indicator j is calculated as $w_j \times h_j(k)$ and the relative contribution is calculated as:

$$\frac{w_j \times h_j(k)}{H_p \times A_p}.$$

If the relative contribution to disempowerment of an indicator greatly exceeds its weight, this suggests that the disempowered are disproportionately more unconstrained in this indicator compared to other indicators.

(ii) **Decompose WEMNS by subgroup.** WEMNS can also be decomposed by subgroup, although the appropriateness of such a decomposition depends on sample design and the measurement comparability of the indicator across the sub-groups. In addition, decomposition requires that the subgroups are mutually exclusive and exhaustive of the study sample.

The contribution of each subgroup to disempowerment is based on sample shares. The absolute contribution of subgroup m to disempowerment is calculated as:

$$\frac{n'}{n}(H'_p \times A'_p),$$

where n' is the sample share of the subgroup; H'_p is the disempowerment headcount ratio for the subgroup; and A'_p is the intensity of disempowerment for the subgroup. The relative contribution is calculated as:

$$\frac{\frac{n'}{n}(H'_p \times A'_p)}{H_p \times A_p}.$$

If the relative contribution to disempowerment of a subgroup greatly exceeds its sample share, this suggests that the group bears a disproportionate share of disempowerment.

As above, it is also possible to calculate the contribution of each subgroup to empowerment. This can be done using the same formulas as above can be used, substituting $1 - (H_p \times A_p)$ in place of $H_p \times A_p$.

RESULTS

Respondent characteristics

Table 2 presents basic demographic information for study sample respondents for whom complete data on all indicators are available. Because our samples are not nationally representative, the results should not be taken as representative of empowerment in the respective countries, but as illustrative of our methodology and its potential applications. Thus, when we refer to each country setting, this should be interpreted as the study sample or sub-sample from each country.

The respondents were between the ages of 18 and 64. Most respondents were married and had either never attended school or had completed only primary school. Respondents' employment status was heterogenous across country settings, likely due to differences in sampling strategies. The sampling strategy also may explain why respondents were equally split between rural and urban areas.

Table 2. Respondent characteristics by pilot study setting and gender

	Bangladesh		Malawi		Nepal		Guatemala	
	Women	Men	Women	Men	Women	Men	Women	Men
<i>Age group, %</i>								
18-24	21.9	13.6	31.2	27.9	19.5	17.8	12.3	13.7
25-34	31.3	27.4	32.5	30.3	28.5	24.6	28.0	24.2
35-44	24.3	26.1	20.9	22.5	27.0	25.3	24.8	24.4
45-54	16.9	17.9	11.4	12.6	17.7	18.4	19.1	19.1
55-64	5.8	15.0	4.1	6.7	7.3	14.0	15.8	18.6
<i>Education level, %</i>								
None	41.6	52.6	54.1	43.6	67.4	51.5	39.7	24.7
Primary completed	43.4	29.1	29.1	30.3	23.3	35.8	35.5	40.3
Secondary or higher completed	15.0	18.3	16.8	26.1	9.3	12.7	24.8	35.0
<i>Marital status, %</i>								
Married	93.3	90.6	85.0	78.6	88.9	83.0	86.2	84.0
<i>Residence, %</i>								
Rural	50.0	50.0	49.9	50.0	48.2	48.35	51.9	51.7
Urban	50.0	50.0	50.1	50.0	51.9	51.9	48.1	48.3
<i>Employment status, %</i>								
Employed on-farm	1.0	29.1	44.1	34.9	45.1	49.9	1.6	42.7
Employed off-farm	11.1	65.6	24.0	49.9	8.5	40.3	17.4	49.7
Unemployed	3.3	5.1	31.6	15.3	3.6	9.7	2.7	7.2
Cares for household	84.6	0.1	0.3 ^a	0.0	42.8	0.1	78.3	0.5
Observations	800	800	791	792	810	812	811	811

Source: Authors' calculations.

Note: ^a Household rosters differed slightly across country. For Malawi, a response category for "cares for the household" was not included. We recovered a few reports of "cares for the household" based on "other (please specify)" responses, but it is likely that the share of unemployed women (and men) would be lower if we had explicitly included a response category for "cares for the household."

Aggregate results

Table 3 presents the aggregate results for women and men respondents in each study setting, based on a 70 percent empowerment cutoff; corresponding results for 60 and 65 percent empowerment cutoffs are presented in Table 14 in Appendix III. Comparing WEMNS scores, for which higher values reflect greater empowerment, women are less empowered than men across all four countries, albeit with some heterogeneity. WEMNS scores range between 0.44 and 0.57 for women and between 0.49 and 0.70 for men.

Table 3. WEMNS results by pilot study setting and gender

	Bangladesh		Malawi		Nepal		Guatemala	
	Women	Men	Women	Men	Women	Men	Women	Men
WEMNS score	0.50	0.68	0.57	0.70	0.44	0.47	0.43	0.49
Disempowerment headcount ratio (H_p)	0.90	0.69	0.81	0.62	0.93	0.93	0.92	0.90
Mean disempowerment score among disempowered (A_p)	0.55	0.47	0.53	0.47	0.61	0.57	0.62	0.57
Observations	797	797	787	773	809	802	796	779

Source: Authors' calculations.

Note: Based on an empowerment cutoff of 70% or higher. $WEMNS = 1 - (H_p \times A_p)$

We also observe heterogeneity when comparing the pervasiveness and intensity of disempowerment across study settings. The disempowerment headcount ratio, which reflects the pervasiveness of disempowerment, indicates that 81 percent of women respondents in Malawi are disempowered, relative to 90 to 93 percent of women respondents in Bangladesh, Guatemala, and Nepal. In comparison, 62 to 69 percent of men respondents in Malawi and Bangladesh are disempowered, relative to 90 to 93 percent in Guatemala and Nepal. The mean disempowerment score among the disempowered, which reflects the intensity of disempowerment, is 0.06 – 0.10 higher for women and men respondents in Nepal and Guatemala compared to those in Bangladesh and Malawi.

Differences are also evident when comparing the distributions of women's and men's empowerment scores across pilot study settings (Figure 1). Empowerment scores are relatively higher for women and men respondents in Bangladesh and Malawi, compared to those in Nepal and Guatemala, as evident by the rightward shift of the distributions in Figure 1.

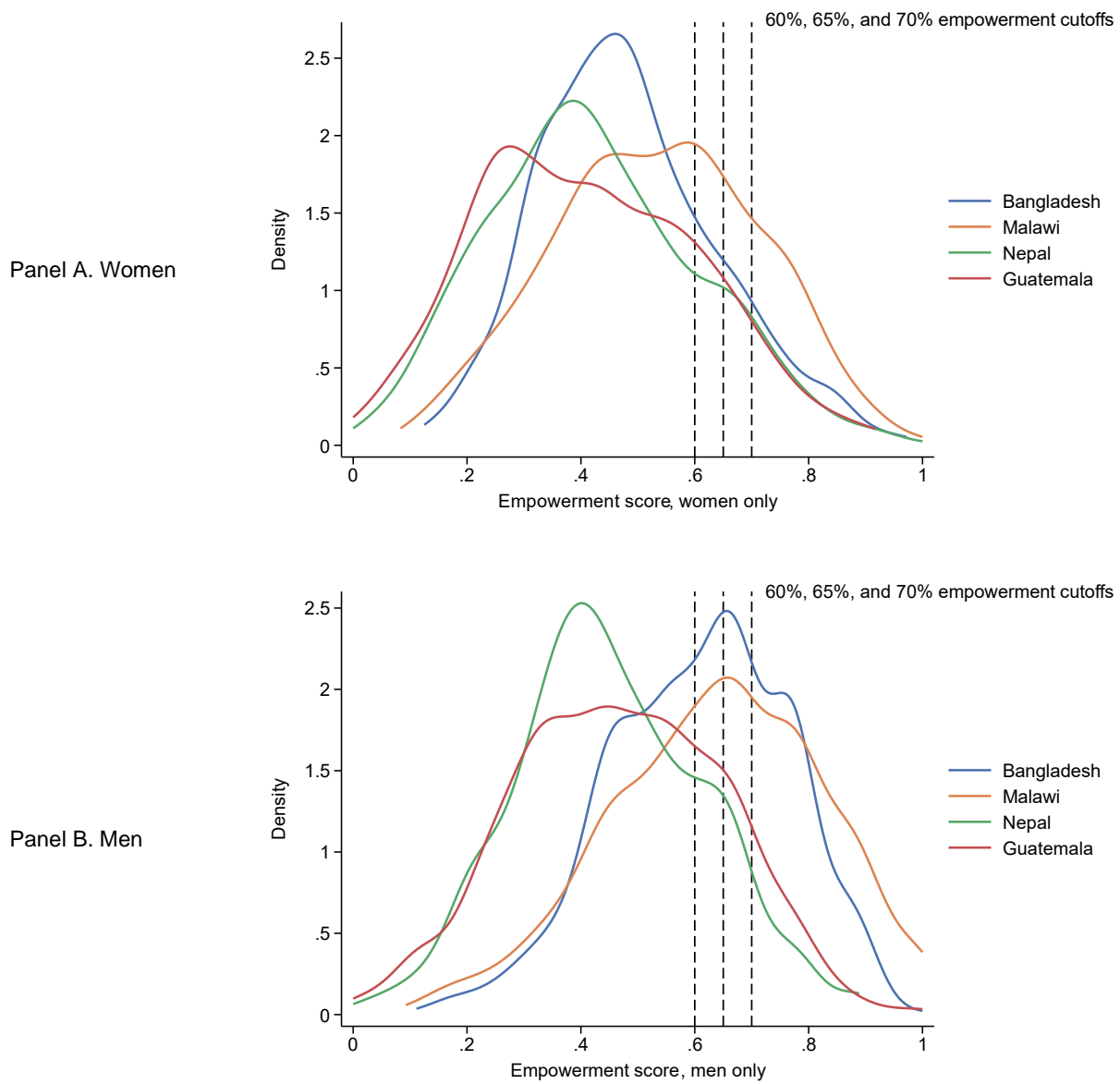


Figure 1. Kernel density estimates of empowerment scores, by pilot study setting and gender

Source: Authors' calculations.

Indicator-level results

While the aggregate results allow us to understand broad patterns or trends in women's and men's empowerment, the uncensored and censored indicator headcount ratios, shown in Table 4, enable a better understanding of how women and men respondents fare across the indicators. The uncensored headcount ratios reflect the share of respondents who are constrained in each indicator (or sub-indicator). The

censored headcount ratios, shown in parentheses in Table 4, reflect the share of respondents who are both disempowered and constrained in each indicator (or sub-indicator).

Table 4. Uncensored and censored indicator headcount ratios (% constrained in each indicator), by pilot study setting and gender

Indicator	Sub-indicator	Bangladesh		Malawi		Nepal		Guatemala	
		Women	Men	Women	Men	Women	Men	Women	Men
Endorsement of women's freedom in livelihood choices*		68.0 (66.8)	-	45.0 (42.1)	-	56.4 (53.4)	-	53.6 (53.4)	-
Endorsement of women's freedom in family formation choices*		77.8 (75.4)	-	58.1 (53.0)	-	69.2 (67.2)	-	57.7 (57.3)	-
Rejection of women's subjection to sexual harassment*		45.5 (44.9)	-	29.2 (24.8)	-	19.0 (18.7)	-	32.4 (32.0)	-
Influence over time allocation		30.5 (30.1)	33.5 (33.0)	65.3 (59.0)	50.3 (41.3)	63.4 (61.8)	77.4 (76.7)	52.4 (51.4)	79.6 (73.4)
Influence in spending decisions		66.4 (64.5)	42.4 (42.0)	76.1 (68.5)	53.2 (43.6)	80.5 (79.2)	69.1 (68.6)	74.7 (73.9)	75.4 (69.8)
Influence in own health decisions		26.1 (25.6)	11.3 (11.0)	31.6 (31.0)	13.8 (12.9)	40.0 (40.0)	23.9 (23.8)	44.7 (44.6)	32.1 (30.8)
Participation in organizations		19.2 (18.9)	30.4 (29.6)	28.1 (27.1)	17.1 (16.7)	49.9 (49.9)	63.6 (63.2)	63.2 (58.8)	48.4 (44.0)
Leadership in organizations		85.7 (80.9)	82.9 (78.4)	61.4 (56.3)	46.8 (40.0)	86.3 (82.9)	88.9 (87.5)	83.4 (77.9)	69.4 (62.8)
Confidence in women's/men's community engagement		87.8 (83.2)	87.7 (82.7)	39.1 (35.7)	41.9 (37.6)	95.2 (89.1)	94.0 (92.0)	76.8 (73.7)	63.9 (61.5)
Use of information communication technologies (ICT)		22.3 (21.7)	7.8 (7.3)	38.5 (36.3)	23.9 (20.7)	43.5 (42.6)	13.2 (13.1)	33.8 (33.5)	22.1 (21.6)
Use of financial services and access to credit	Use of financial services	57.7 (52.7)	27.2 (24.8)	55.0 (51.3)	37.1 (30.8)	71.6 (69.2)	51.6 (51.1)	76.3 (72.4)	57.0 (52.1)
	Access to credit	2.6 (2.4)	3.3 (3.3)	14.6 (13.6)	19.8 (16.7)	15.1 (15.0)	9.6 (9.6)	59.4 (58.4)	38.3 (33.6)
Secure property rights	Documented property rights	86.6 (79.7)	55.5 (51.6)	84.2 (71.5)	76.2 (58.1)	77.9 (75.2)	56.5 (56.2)	73.0 (70.2)	47.8 (45.3)
	Secure transfer rights	62.1 (58.7)	32.4 (30.4)	41.8 (36.8)	44.2 (36.9)	53.8 (52.2)	38.3 (38.0)	57.7 (55.4)	28.6 (27.0)
	Secure tenure rights	33.4 (31.157 .7)	32.4 (23.8)	41.8 (35.1)	44.2 (30.9)	53.8 (51.7)	38.3 (35.3)	57.7 (55.4)	28.6 (27.2)
Observations		797	797	787	773	809	802	796	779

Source: Authors' calculations.

Note: Censored inadequacy headcount ratios in parentheses. * Indicators not calculated for men. Censored inadequacy headcounts are based on an empowerment cutoff = 70%.

A higher share of women compared to men are inadequate in nearly all indicators (or sub-indicators) across the four countries. Because most women (78–92 percent) are disempowered, the uncensored and censored inadequacy headcount ratios for women are broadly similar. This is true, as well, for men in Nepal and Guatemala, where 85 to 87 percent are disempowered. In Bangladesh and Malawi, where men are less likely to be disempowered, we find the uncensored and censored inadequacy headcount ratios are similar for only some indicators (e.g., participation in organizations and influence in own health

decisions). The more similar are the uncensored and censored inadequacy headcount ratios for an indicator, the more likely it is that disempowered women (or men) are constrained in the indicator.

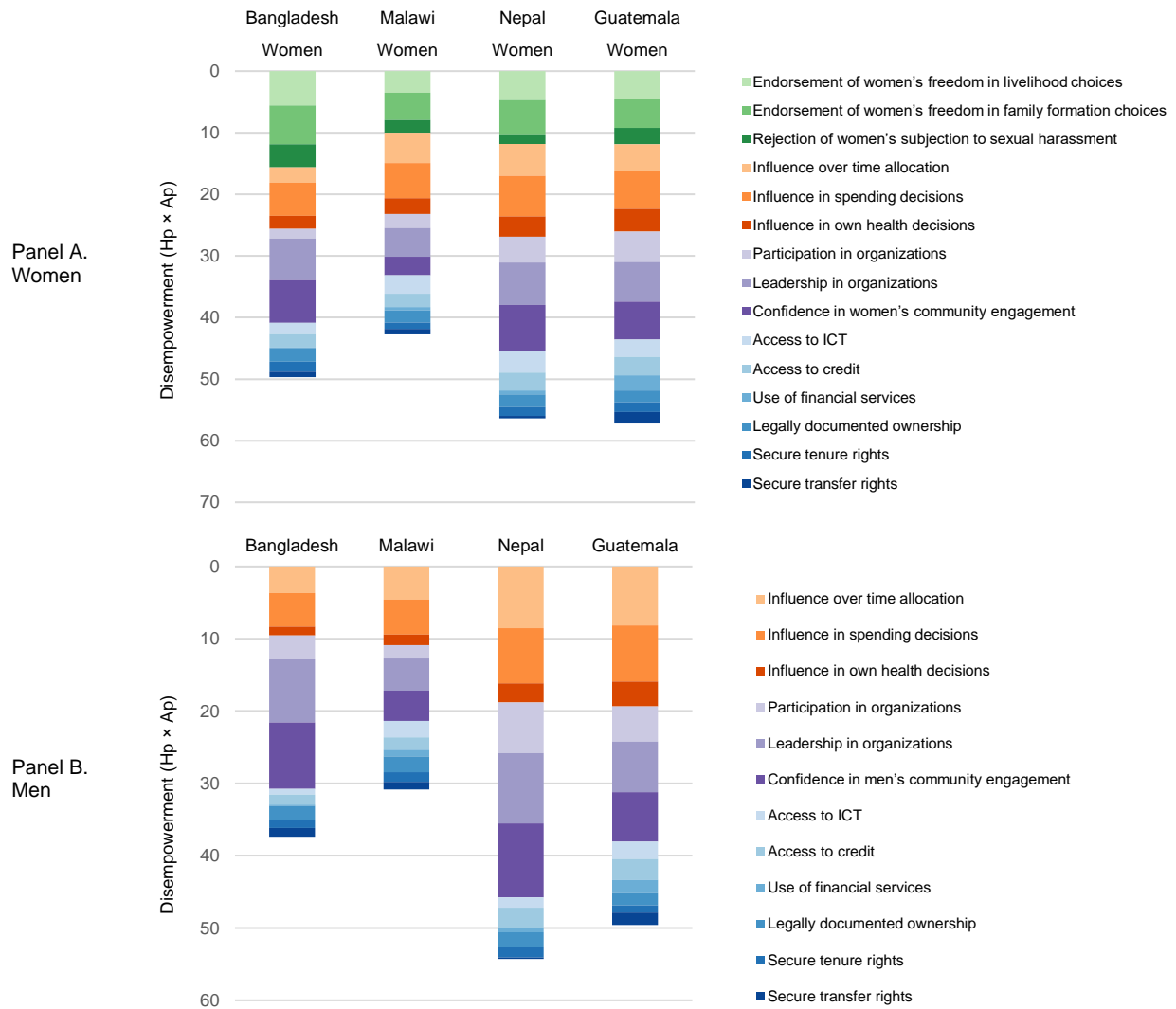


Figure 2. Contribution of each indicator to disempowerment, by pilot study setting and gender

Source: Authors' calculations.

Note: Based on an empowerment cutoff of 70% or higher. The green, orange, purple, and blue shadings used in the figure correspond to the intrinsic agency, instrumental agency, collective agency, and agency-enabling resources domains, respectively.

Based on the censored headcount ratios, we can decompose how much each indicator contributes to disempowerment within the population, as shown in Figure 2. This type of analysis is particularly useful as a diagnostic tool to identify the indicators with the most potential for improving the empowerment of

women (or men) within the population. In the pilot sample, for instance, confidence in women's community engagement, leadership in organizations, and influence in spending decisions are among the indicators that contribute the most to women's disempowerment. The decomposition, thus, suggests that these are potentially high-impact areas for policy interventions among the study sample.

This type of decomposition is, in principle, sensitive to the choice of empowerment cutoff. Figure 2, for instance, is based on an empowerment cutoff of 70 percent. In this instance, however, the top three contributors to women's and men's disempowerment in each of the study settings are the same regardless of whether we define the empowerment cutoff as 60, 65, or 70 percent (see Figure 4 and Figure 5 in Appendix III).

Decomposition of WEMNS by subgroup

WEMNS is decomposable at any level for which the underlying data are representative. Thus, when implementing WEMNS, it would be important to ensure that the sample sizes are sufficiently large and representative for relevant subgroups. In

Table 5, we decompose the aggregate WEMNS results (analogous to Table 3) by rural/urban status. In this example, WEMNS scores are lower for women in rural areas vis-à-vis urban areas in Malawi and Guatemala and broadly similar in Bangladesh and Nepal. Scores for men in rural areas are higher in Bangladesh, lower in Malawi, and broadly similar in Nepal and Guatemala. The disempowerment headcount ratios and mean disempowerment scores among disempowered respondents show similar, if not entirely consistent, variation across country. These results suggest that patterns of disempowerment across rural and urban areas are not the same for women and men, nor are they the same across the pilot study settings. While all decompositions presented here are illustrative, it should additionally be noted that the psychometric analysis conducted on the WEMNS item sets did not test for measurement invariance by rural/urban location.

Table 5. Decomposition of the WEMNS results, by pilot study settings, gender, and rural/urban residence

		Bangladesh		Malawi		Nepal		Guatemala	
		Women	Men	Women	Men	Women	Men	Women	Men
Rural	<i>WEMNS score</i>	0.51	0.73	0.52	0.67	0.45	0.47	0.38	0.49
	Disempowerment headcount ratio (H_p)	0.89	0.61	0.88	0.67	0.92	0.93	0.96	0.89
	Mean disempowerment score among disempowered (A_p)	0.56	0.45	0.55	0.49	0.60	0.57	0.65	0.57
	Observations	398	399	394	387	390	386	413	409
Urban	<i>WEMNS score</i>	0.50	0.62	0.63	0.74	0.43	0.47	0.48	0.49
	Disempowerment headcount ratio (H_p)	0.92	0.77	0.74	0.56	0.93	0.93	0.88	0.90
	Mean disempowerment score among disempowered (A_p)	0.55	0.49	0.50	0.46	0.61	0.57	0.59	0.57
	Observations	399	398	393	386	419	416	383	409

Source: Authors' calculations.

Note: Based on an empowerment cutoff of 70% or higher. $WEMNS = 1 - (H_p \times A_p)$

Robustness analysis

To analyze the robustness of our measures, we consider the non-response rates, the associations between indicators, and the associations of the indicators with other outcomes.

Missingness and non-response rates

To calculate WEMNS, responses are required for every indicator for each respondent. Non-response, resulting in a missing indicator, occurs when a respondent does not answer at least one of the required survey questions required to calculate an indicator.¹¹ For example, to be considered unconstrained in endorsement of women's freedom in livelihood choices, a respondent must have recorded responses to each of the five statements used to calculate the indicator. If responses were missing for any of the five statements, the indicator would be missing. As noted above, we deliberately did not select items with high levels of missingness for use in constructing WEMNS indicators.

As a result, missingness among indicators in the study sample, shown in Table 6, was extremely low. For six indicators, no respondents have non-responses. For the remaining nine indicators, the proportion of respondents with non-responses is, generally, less than 1 percent. The indicator with the highest proportion of missingness is influence over time allocation, likely due to outlier cases in which

¹¹ Note that for leadership in organizations, if a respondent reported not being a participant in that organization, their leadership in that organization was automatically recorded as 0.

respondents did not spend any time (during the 7 preceding days) on one or more of the six paid and unpaid work categories asked about in the module. Overall, 98.6 percent of respondents in the pilot sample answered all the survey questions needed to compute WEMNS. The consistency of these findings across country settings and the results of cognitive interviewing (Appendix II) suggest that the items used to construct WEMNS are readily answerable by respondents.

Table 6. Non-response rates (% missingness) for WEMNS indicators, by pilot study setting and gender

Indicator	Sub-indicator	Bangladesh		Malawi		Nepal		Guatemala	
		Women	Men	Women	Men	Women	Men	Women	Men
Endorsement of women's freedom in livelihood choices		0.00	-	0.00	-	0.00	-	0.00	-
Endorsement of women's freedom in family formation choices		0.00	-	0.00	-	0.00	-	0.00	-
Rejection of women's subjection to sexual harassment		0.00	-	0.00	-	0.00	-	0.00	-
Influence over time allocation		0.37	0.37	0.50	2.27	0.00	1.10	1.10	3.58
Influence in spending decisions		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Influence in own health decisions		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Participation in organizations		0.00	0.00	0.00	0.00	0.00	0.25	3.70	0.00
Leadership in organizations		0.00	0.00	0.00	0.00	0.00	0.25	3.70	0.00
Confidence in women's/men's community engagement		0.00	0.00	0.00	0.00	0.00	0.00	3.70	0.00
Use of ICT		0.00	0.00	0.00	0.13	0.00	0.25	0.12	0.00
Use of financial services and access to credit	Use of financial services	0.00	0.00	0.00	0.13	0.00	0.25	0.12	0.00
	Access to credit	0.00	0.00	0.00	0.00	0.12	0.37	0.37	0.25
Secure property rights	Documented property rights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Secure tenure rights	0.00	0.00	0.00	0.00	0.00	0.25	0.12	0.37
	Secure transfer rights	0.00	0.00	0.00	0.00	0.00	0.25	0.12	0.37
Observations		800	800	791	792	810	812	811	811

Associations among WEMNS indicators

Next, we consider pairwise associations between WEMNS indicators using Cramer's V. A strong pairwise correlation suggests redundancy, which could contribute to unnecessarily long interview times. It could also result in a greater than intended implicit weight being assigned to indicator combinations, which would need to be theoretically justified.

As shown in Table 7, most of the 15 indicators show low to negligible correlation with each other ($V < 0.30$). We observe moderate correlation ($V = 0.30$ – 0.50) within the intrinsic agency domain between endorsement of women's freedom in livelihood and family formation choices ($V = 0.41$); within the instrumental agency domain between influence in spending and own health decisions ($V = 0.32$); within

the collective agency domain between participation and leadership in organizations ($V = 0.46$); and within the agency-enabling resources domain between use of ICT and access to credit ($V = 0.31$) and between documented property rights, secure tenure rights, and secure transfer rights ($V = 0.31-0.50$). Correlations in this range are not unexpected in our experience and are tolerable from an index-building perspective, given the related nature of the indicators within each domain and the similar item structure in many cases. In all cases, the conceptual importance of maintaining them as distinct indicators outweighs arguments for combining them based on these correlations.

Correlation among indicators across domains is potentially more problematic for maintaining a balanced index. Three indicators exhibit moderate correlation across domains: endorsement of women's freedom in livelihood and family formation choices within the intrinsic agency domain and confidence in women's/men's community engagement ($V = 0.30-0.32$) within the collective agency domain, respectively. These correlations may be due to similarities in the phrasing of the survey items and responses and the sequence in which the questions were asked. All three item sets utilize the same question and response structure. Respondents are read a statement and asked to indicate via an optional pictorial response card whether they fully disagree, partly disagree, partly agree, or fully agree with the statement. Moreover, the item sets are asked nearly one after the other, separated only by an item set asking about confidence in women's/men's community engagement that, again, utilizes the same question and response structure. Additional analysis and cognitive testing of these items is planned to investigate these issues. Nevertheless, given the conceptual significance of the indicators we feel justified in maintaining them in the index.

Table 7. Pairwise associations (Cramer's V) among WEMNS indicators, total sample

	Intrinsic Agency			Instrumental Agency			Collective Agency	
	Endorsement of women's freedom in livelihood choices	Endorsement of women's freedom in family formation choices	Rejection of women's subjection to sexual harassment	Influence over time allocation	Influence in spending decisions	Influence in own health decisions	Participation in organizations	Leadership in organizations
Endorsement of women's freedom in livelihood choices	1.00							
Endorsement of women's freedom in family formation choices	0.41	1.00						
Rejection of women's subjection to sexual harassment	0.14	0.13	1.00					
Influence over time allocation	0.04	0.05	-0.03	1.00				
Influence in spending decisions	0.03	0.03	0.01	0.26	1.00			
Influence in own health decisions	0.05	0.05	-0.01	0.20	0.32	1.00		
Participation in organizations	-0.01	0.02	-0.09	0.05	0.09	0.11	1.00	
Leadership in organizations	0.06	0.11	-0.01	-0.02	0.04	0.06	0.46	1.00
Confidence in women's/men's community engagement	0.30	0.32	0.10	0.05	0.07	0.07	0.10	0.18
Use of information communication technologies (ICT)	-0.01	0.01	-0.02	0.08	0.12	0.15	0.07	0.05
Use of financial services	-0.02	-0.03	-0.02	0.06	0.11	0.14	0.21	0.04
Access to credit	0.00	0.02	-0.03	0.06	0.14	0.17	0.16	0.11
Documented property rights	0.00	-0.03	0.01	0.01	0.13	0.07	0.03	0.07
Secure tenure rights	0.00	-0.05	0.05	0.01	0.08	0.06	0.04	0.03
Secure transfer rights	0.01	-0.02	0.01	-0.01	0.13	0.07	0.07	0.10
	Collective Agency			Enabling Resources				
	Confidence in women's/men's community engagement	Use of information communication technologies	Use of financial services	Access to credit	Documented property rights	Secure tenure rights	Secure transfer rights	
Confidence in women's/men's community engagement	1.00							
Use of information communication technologies (ICT)	-0.01	1.00						
Use of financial services	-0.01	0.14	1.00					
Access to credit	0.06	0.31	0.23	1.00				
Documented property rights	-0.01	0.11	0.02	0.13	1.00			
Secure tenure rights	-0.04	0.05	0.18	0.04	0.32	1.00		
Secure transfer rights	0.05	0.04	0.09	0.09	0.50	0.31	1.00	

Source: Authors' calculations.

Pairwise associations of WEMNS with other outcomes

As a final robustness check, we assessed the association between empowerment scores and measures of food security, agency, subjective well-being, and attitudes toward intimate partner violence from the World Values Survey (WVS), which were purposefully included in the WEMNS pilot questionnaire for testing construct validity (Westen and Rosenthal 2003). The theoretical relationship between these outcomes and empowerment is not always clear. For example, subjective well-being measures, like happiness and life satisfaction, are known to be sensitive to weather and environmental conditions, which could confound any correlation with empowerment (Zhang, Zhang, and Chen 2017; Zapata 2022). However, for the most part, we expect that these outcomes should be positively correlated or, at minimum, not negatively correlated with empowerment. Pairwise (Pearson’s) correlations between women’s and men’s empowerment scores and the WVS outcomes are shown in Table 8.

Table 8. Pearson's correlation between empowerment scores and other outcomes by pilot study setting and gender

		Food security	Free choice	Happiness	Self-Rated Health	Life satisfaction	Financial satisfaction	Rejection of intimate partner violence
Bangladesh	Men	0.06	0.12	0.06	0.06	0.05	0.11	0.09
	Women	0.05	0.19	0.06	0.13	0.10	0.10	0.09
Malawi	Men	0.10	0.11	0.10	0.08	0.10	0.13	0.55
	Women	0.21	0.22	0.10	0.11	0.13	0.12	0.16
Nepal	Men	0.03	0.10	0.07	0.07	0.08	0.08	0.05
	Women	0.13	0.24	0.08	0.07	0.03	0.12	0.04
Guatemala	Men	0.16	0.27	0.14	0.20	0.24	0.20	0.08
	Women	0.14	0.31	0.08	0.05	0.25	0.22	0.12

Source: Authors’ calculations.

Notes: Food security was measured as the frequency the respondent reported going without enough food in the last 12 months (often=1, sometimes=2, rarely=3, never=4). Free choice was measured as the extent of free choice and control respondents feel like they have over their lives (none=1, a little=2, some=3, a great deal=4). Overall happiness, self-rated health, life satisfaction, and financial satisfaction were measured on similar 4-point scales (not at all=1 to very happy=4; very poor=1 to very good=5; completely dissatisfied=1 to completely satisfied=4). Rejection of intimate partner violence was measured as the extent to which the respondent thought it was not justified for a man to beat his wife (never=1, rarely=2, sometimes=3, always=4).

In general, the association between empowerment scores and WVS outcomes is very low. While there are exceptions, such as respondents’ extent of free choice, life satisfaction, and rejection of intimate partner violence in some study settings, most correlations are below 0.20. Notably, however, none of the correlations are negative. Thus, while these results do not suggest the existence of a positive association

between WEMNS and these outcomes in the study sample, they do, at least, support the notion that WEMNS is not negatively associated with these outcomes. These results corroborate those found using factor-derived scores (Yount et al., 2023).

Estimated interview times

For WEMNS to be widely adopted, short interview duration is important. The CAPI program for the pilot questionnaire included embedded time stamps for the start and end times for each module. Given that some modules of the pilot questionnaire included items that were not ultimately adopted for use in the construction of indicators, simply summing the module-specific interview times would lead to an inflated estimate of the time required to collect only those items strictly necessary for calculating the WEMNS indicators. To account for this, we divided the module-specific interview times by the number of items included in the module to create item-specific time estimates. Using these, we calculate an adjusted estimate of the time required to collect WEMNS, based on the minimum number of required items per module.

Kernel density estimates for the adjusted interview times and median values for each study setting are shown in Figure 3 and Table 9, respectively. All time estimates are generally within (or close to) the target interview range of 15–20 minutes set by stakeholders at the outset of the effort to develop WEMNS. Median adjusted interview times range between 13.3 to 21.4 minutes across the study settings, with an overall median of 16 minutes in the pooled sample. Interview times are, on average, estimated to be approximately three minutes shorter for men than for women due to the difference in the number of questions asked. The nine indicators used to calculate WEMNS for men require 42-53 items; the full set of twelve indicators used to calculate WEMNS for women requires 56-67 items.¹² Many of these items have identical roots; for example, the same questions are asked of multiple activity categories, which may reduce the time needed for individuals to comprehend the questions they are asked.

¹² The number of required questions is specified in terms of a range due to skip patterns in the questionnaire.

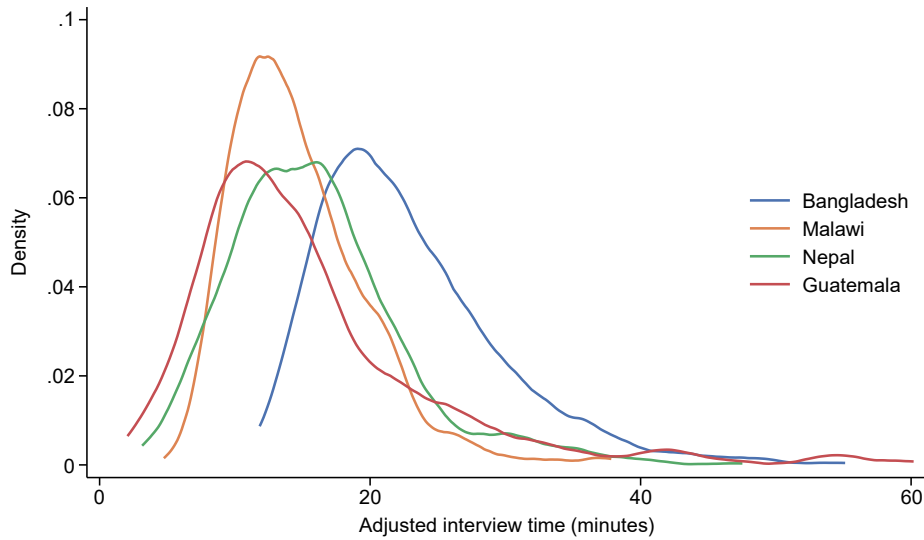


Figure 3. Kernel density estimates for the adjusted overall interview times by study setting

Note: Includes both women and men respondents.

Source: Authors' calculations.

Table 9. Median adjusted interview time (in minutes) by pilot study setting and gender, interquartile range in parentheses

	Bangladesh	Malawi	Nepal	Guatemala	Total
Women	23.7 (19.9–28.8)	14.9 (12.2–18.1)	17.3 (13.3–21.6)	14.1 (10.6–19.7)	17.5 (13.0–23.0)
Men	19.5 (16.5–23.4)	12.7 (10.1–16.5)	13.3 (10.6–16.5)	12.4 (8.7–17.3)	14.8 (10.9–19.1)
Total	21.4 (17.9–26.4)	13.7 (11.1–17.4)	15.3 (11.5–19.1)	13.3 (9.6–18.3)	16.1 (11.9–21.1)
Observations	1600	1583	1622	1622	6427

Note: Interquartile range in parentheses

Source: Authors' calculations.

CONCLUSION

Despite the growing commitment to women's empowerment, its multidimensional nature has hampered the development of measures that can be used in large sample surveys to assess the status of empowerment in different countries and sub-groups, to identify specific domains that need attention, and to track progress over time. To address these challenges, WEMNS provides a headline measure of empowerment that can also be broken down and analyzed to inform policy. Unlike other existing measures, WEMNS is designed to be applicable across livelihood strategies, in rural and urban areas, and across adult women and men of all life stages (not just those of reproductive age nor the primary couple in the household). Although there are four additional indicators for women that are not applicable to men, the measure permits assessing men's empowerment as well as women's and allows for comparisons by sex across specific indicators, and thereby provides the basis for assessing gender (in)equality in empowerment. The domains and indicators of WEMNS are aligned with Kabeer's (1999) framework for assessing empowerment, covering intrinsic, instrumental, and collective agency, as well as agency-enabling resources. Particular indicators are also aligned and consistent with SDG indicators, especially those related to property rights and information-related resources. The iterative process of developing WEMNS, building from initial key informant interviews and assessment of existing survey modules through a process of iterative refinement of the instrument, drawing on cognitive interviewing, two rounds of phone surveys, psychometric testing, and the expertise of national partners with extensive experience in fielding large-scale surveys in four countries in Asia, Africa, and Latin America has led to an instrument that has the potential to be used in a wide variety of contexts. Results from face-to-face piloting of WEMNS in Bangladesh, Malawi, Nepal, and Guatemala indicate that the measure can distinguish different levels of empowerment across the countries and between women and men.

The WEMNS survey module also has several major strengths and advantages over other women's empowerment modules. Relatively few questions are needed to calculate the WEMNS index (56-67 items required for women; 42-53 items required for men), and as a result, the median time required to

administer the WEMNS module ranges from 13–21 minutes per individual. The relative brevity of WEMNS suggests it is a reasonable length for integration into multi-topic surveys based on feedback from survey practitioners, without greatly increasing respondent burden. The addition of the WEMNS module into major survey platforms would increase the value of the WEMNS data, by allowing analysis of the associations between (women’s) empowerment and welfare outcomes such as health, and nutritional status, and economic outcomes like income and productivity. Including WEMNS in survey platforms that collect longitudinal data also makes it possible to track progress in women’s empowerment over time. This makes it particularly suitable for the 50x2030 Initiative and other efforts to build nationally comparable data. It also makes it easier to add women’s empowerment data to other research and impact assessment efforts. While building comparable data from nationally representative data on WEMNS and other variables of interest would be the best way to demonstrate the value of WEMNS, the data collected from this pilot study provides a wealth of information that can be explored to establish relationships between women’s (and men’s) empowerment and other development outcomes.

Our process—sampling households with an adult woman and man, and then two randomly selected adults of both sexes—may yield different results from surveys like the World Value Surveys, where women and men are randomly selected, and only one is surveyed from each household. We interviewed two individuals from each household not only to permit intrahousehold comparisons but also to minimize survey costs. However, the interpretation of comparisons between women and men within the sample household is different from that in the WEAI or pro-WEAI, where only the primary man and woman are interviewed. WEMNS thus allows us to examine empowerment across the life cycle, not only for those in their reproductive years, and across all types of household members (not just the primary woman or man). Future applications can create more representative samples by also including households with only an adult woman or man.

Despite these important contributions, some limitations to the WEMNS module and counting-based methodology are notable. First, the pilot study samples are not nationally representative. Second, some

indicators are only applicable to women and not men, implying that women's and men's empowerment scores are not directly comparable due to differences in composition and weights.

The iterative process of piloting the WEMNS module provides a promising foundation for advancing the research and policy agenda on women's empowerment. To fully reap the returns of this initial investment in instrument development, we recommend continued research and testing of WEMNS along two complementary, and mutually reinforcing, tracks. The first involves refining the instrument, and the second, demonstrating its utility in advancing the policy agenda.

With respect to refining the instrument, because the current version of WEMNS was tested using samples that are not nationally representative, we recommend conducting further sensitivity analyses around the cutoffs used in this study using representative samples. This step would enable us to assess minimum thresholds for being considered unconstrained in other socioeconomic settings and "at scale." The WEMNS module also could be administered in cognitive interviews and sample surveys in other settings to assess its face validity and psychometric properties more widely, as well as in repeated cross-sectional studies and/or longitudinal study to assess its measurement invariance over time and age, a key criterion for its use to monitor change over time (Yount et al. 2023). As others begin to use WEMNS, users can document how the instrument functions in different settings, so that further modifications can be guided by user feedback, similar to our experience in developing and scaling up WEAI and pro-WEAI.

Our recommendations for the other track on advancing the policy agenda for women's empowerment draw on our experience in scaling up WEAI since 2012. Like WEMNS, WEAI was initially released as a pilot instrument used in population-based surveys of the Feed the Future Initiative. Extensive analysis of the initial round of WEAI data in relation to other development outcomes related to household food security, diets, and nutrition demonstrated both the utility of collecting empowerment data within a population-based multitopic survey and its ability to influence national policy and project design (Malapit et al. 2014; Quisumbing et al. 2021). Use of WEAI as a diagnostic has influenced the design of nutrition-sensitive agricultural programs in Bangladesh and donor programming elsewhere, owing to the ability to diagnose areas in which disempowerment is greatest. The decomposability of WEMNS lends itself to

similar uses. We also propose exploring the relationship of WEMNS, both the overall index and its component indicators, to other indicators collected in the validation module. Once WEMNS is piloted in several nationally representative multitopic household surveys, its relationship to other development outcomes in different settings can be analyzed, building the evidence base on women's empowerment and other development goals. Other research and impact assessment studies using smaller samples can then compare their results with nationally representative data to contextualize patterns of women's and men's empowerment in their sites, relative to the country as a whole, or to other rural or urban areas.

As our previous experience suggests, the best way to realize the value of investing in a new metric of women's empowerment is to use it and to do so at scale. As a new metric, WEMNS inevitably will be imperfect, but waiting for a perfect instrument would be a missed opportunity to influence the policy agenda on women's empowerment. Transparency about modifications and adjustments to WEMNS resulting from user feedback can contribute to an improved module and greater take up on a larger scale. Continued support for this endeavor by governments, donors, and policymakers will be important.

REFERENCES

- Abubakar, Naima Hafiz, and Salihu Ibrahim Dasuki. 2018. "Empowerment in Their Hands: Use of WhatsApp by Women in Nigeria." *Gender, Technology and Development* 22 (2): 164–83. <https://doi.org/10.1080/09718524.2018.1509490>.
- Alkire, Sabina, and James Foster. 2011. "Counting and Multidimensional Poverty Measurement." *Journal of Public Economics* 95 (7–8): 476–87. <https://doi.org/10.1016/j.jpubeco.2010.11.006>.
- Alkire, Sabina, James E Foster, Suman Seth, Maria Emma Santos, Jose M Roche, and Paola Ballon. 2015a. "Normative Choices in Measurement Design." In *Multidimensional Poverty Measurement and Analysis*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199689491.003.0007>.
- . 2015b. "Robustness Analysis and Statistical Inference." In *Multidimensional Poverty Measurement and Analysis*. Oxford: Oxford University Press.
- Alkire, Sabina, Ruth Meinzen-Dick, Amber Peterman, Agnes R. Quisumbing, Greg Seymour, and Ana Vaz. 2013. "The Women's Empowerment in Agriculture Index." *World Development* 52 (December): 71–91. <https://doi.org/10.1016/j.worlddev.2013.06.007>.
- Atkinson, Anthony B. 2019. *Measuring Poverty around the World*. Princeton and Oxford: Princeton University Press.
- Boateng, Godfred O., Torsten B. Neilands, Edward A. Frongillo, Hugo R. Melgar-Quiñonez, and Sera L. Young. 2018. "Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer." *Frontiers in Public Health* 6. <https://doi.org/10.3389/fpubh.2018.00149>.
- Cronbach, Lee J., and Paul E. Meehl. 1955. "Construct Validity in Psychological Tests." *Psychological Bulletin* 52 (4): 281–302. <https://doi.org/10.1037/h0040957>.
- Eissler, Sarah, Jessica Heckert, Emily Myers, Greg Seymour, Sheela Sinharoy, and Kathryn Yount. 2022. "Measuring Women's Empowerment: Gender and Time-Use Agency in Benin, Malawi and Nigeria." *Development and Change* 53 (5): 1010–34. <https://doi.org/10.1111/dech.12725>.
- Elias, Marlène, Steven M. Cole, Agnes Quisumbing, Ana Maria Paez Valencia, Ruth Meinzen-Dick, and Jennifer Twyman. 2021. "Assessing Women's Empowerment in Agricultural Research." In *Advancing Gender Equality through Agricultural and Environmental Research: Past, Present, and Future*, edited by Rhiannon Pyburn and Anouka van Eerdewijk, 329–64. Washington, DC: International Food Policy Research Institute.
- Ewerling, Fernanda, John W. Lynch, Cesar G. Victora, Anouka van Eerdewijk, Marcelo Tyszler, and Aluisio J.D. Barros. 2017. "The SWPER Index for Women's Empowerment in Africa: Development and Validation of an Index Based on Survey Data." *The Lancet Global Health* 5 (September): e916–23. [https://doi.org/10.1016/S2214-109X\(17\)30292-9](https://doi.org/10.1016/S2214-109X(17)30292-9).
- Freire, Paulo. 1970. *Pedagogy of the Oppressed*. Translated by Myra Bergman Ramos. A Continuum Book. New York: Herder and Herder.

- Garikipati, Supriya, Isabelle Agier, Isabelle Guérin, and Ariane Szafarz. 2017. “The Cost of Empowerment: Multiple Sources of Women’s Debt in Rural India.” *The Journal of Development Studies* 53 (5): 700–722. <https://doi.org/10.1080/00220388.2016.1205734>.
- Guérin, Isabelle, Bert D’Espallier, and Govindan Venkatasubramanian. 2013. “Debt in Rural South India: Fragmentation, Social Regulation and Discrimination.” *The Journal of Development Studies* 49 (9): 1155–71. <https://doi.org/10.1080/00220388.2012.720365>.
- Hannan, Anika, Jessica Heckert, Laurie James-Hawkins, and Kathryn M. Yount. 2019. “Cognitive Interviewing to Improve Women’s Empowerment Questions in Surveys: Application to the Health and Nutrition and Intrahousehold Relationships Modules for the Project-level Women’s Empowerment in Agriculture Index.” *Maternal & Child Nutrition*, no. December 2018, 1–13. <https://doi.org/10.1111/mcn.12871>.
- Heckert, Jessica, Monica Dennise Dardon Sagastume, and Juan Pablo Pira Martinez. 2023. “Answering the Call for New Approaches in Phone Surveys: Reaching Indigenous-Language Speakers in Guatemala.” *Engendering Data* (blog). 2023. <https://gender.cgiar.org/news/answering-call-new-approaches-phone-surveys-reaching-indigenous-language-speakers-guatemala>.
- Heckert, Jessica, and Madeleine Short Fabric. 2013. “Improving Data Concerning Women’s Empowerment in Sub-Saharan Africa.” *Studies in Family Planning* 44 (3): 319–44. <https://doi.org/10.1111/j.1728-4465.2013.00360.x>.
- Heckert, Jessica, Elena M. Martinez, Greg Seymour, Audrey Pereira, Shalini Roy, Sunny S. Kim, Hazel Malapit, and Gender Assets and Agriculture Project Phase 2 (GAAP2) Health and Nutrition Study Team. 2023. “Development and Validation of a Health and Nutrition Module for the Project-Level Women’s Empowerment in Agriculture Index (pro-WEAI+HN).” *Maternal & Child Nutrition* 19 (2): e13464. <https://doi.org/10.1111/mcn.13464>.
- Huq, M., and M. Sultan. 1991. “‘Informality’ in Development: The Poor as Entrepreneurs in Bangladesh.” In *The Silent Revolution*. San Francisco: International Center for Economic Growth.
- Ibrahim, Solava, and Sabina Alkire. 2007. “Agency and Empowerment: A Proposal for Internationally Comparable Indicators.” *Oxford Development Studies* 35 (4): 379–403. <https://doi.org/10.1080/13600810701701897>.
- Kabeer, Naila. 1994. *Reversed Realities: Gender Hierarchies in Development Thought*. London: Verso.
- . 1999. “Resources, Agency, Achievements: Reflections on the Measurement of Women’s Empowerment.” *Development and Change* 30 (3): 435–64. <https://doi.org/10.1111/1467-7660.00125>.
- Malapit, Hazel J., Jessica Heckert, Patrice Ygué Adegbola, Geraud Fabrice Crinot, Sarah Eissler, Simone Faas, Geoffroy Gantoli, et al. 2023. “Measuring Empowerment across the Value Chain: The Evolution of the Project-Level Women’s Empowerment Index for Market Inclusion pro-WEAI+MI.” 0 ed. Washington, DC: International Food Policy Research Institute. <https://doi.org/10.2499/p15738coll2.136606>.
- Malapit, Hazel J., Kathryn Sproule, Chiara Kovarik, Ruth S. Meinzen-Dick, Agnes R. Quisumbing, Farzana Ramzan, Emily Hogue, and Sabina Alkire. 2014. “Measuring Progress toward Empowerment: Women’s Empowerment in Agriculture Index: Baseline Report.” Washington,

- DC: International Food Policy Research Institute. <https://www.ifpri.org/publication/measuring-progress-toward-empowerment-womens-empowerment-agriculture-index-baseline>.
- Malapit, Hazel, Agnes Quisumbing, Ruth Meinzen-Dick, Greg Seymour, Elena M. Martinez, Jessica Heckert, Deborah Rubin, Ana Vaz, and Kathryn M. Yount. 2019. "Development of the Project-Level Women's Empowerment in Agriculture Index (pro-WEAI)." *World Development* 122 (October):675–92. <https://doi.org/10.1016/j.worlddev.2019.06.018>.
- Meinzen-Dick, Ruth Suseela, Deborah Rubin, Marlène Elias, Annet Abenakyo Mulema, Emily Myers, and Mulema. 2019. "Women's Empowerment in Agriculture: Lessons from Qualitative Research." IFPRI Discussion Paper 1797. International Food Policy Research Institute, Washington, DC, DC.
- Miedema, Stephanie Spaid, Regine Haardörfer, Amy Webb Girard, and Kathryn M. Yount. 2018. "Women's Empowerment in East Africa: Development of a Cross-Country Comparable Measure." *World Development* 110:453–64. <https://doi.org/10.1016/j.worlddev.2018.05.031>.
- Nepa School of Social Sciences and Humanities. 2017. "Empowerment."
- Putnick, Diane L., and Marc H. Bornstein. 2016. "Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research." *Developmental Review* 41 (September):71–90. <https://doi.org/10.1016/j.dr.2016.06.004>.
- Quisumbing, Agnes, Steven Cole, Marlène Elias, Simone Faas, Alessandra Galiè, Hazel Malapit, Ruth Meinzen-Dick, Emily Myers, Greg Seymour, and Jennifer Twyman. 2023. "Measuring Women's Empowerment in Agriculture: Innovations and Evidence." *Global Food Security* 38 (September):100707. <https://doi.org/10.1016/j.gfs.2023.100707>.
- Quisumbing, Agnes R., Ruth Meinzen-Dick, Hazel J. Malapit, Greg Seymour, Jessica Heckert, Cheryl Doss, Nancy Johnson, et al. 2022. "Can Agricultural Development Projects Empower Women? A Synthesis of Mixed Methods Evaluations Using Pro-WEAI in the Gender, Agriculture, and Assets Project (Phase 2) Portfolio." IFPRI Discussion Paper 2137. Washington, DC: International Food Policy Research Institute.
- Quisumbing, Agnes R., Kathryn Sproule, Elena M. Martinez, and Hazel Malapit. 2021. "Do Tradeoffs among Dimensions of Women's Empowerment and Nutrition Outcomes Exist? Evidence from Six Countries in Africa and Asia." *Food Policy* 100 (December 2020): 102001. <https://doi.org/10.1016/j.foodpol.2020.102001>.
- Rettig, Erica M., Stephen E. Fick, and Robert J. Hijmans. 2020. "The Female Empowerment Index (FEMI): Spatial and Temporal Variation in Women's Empowerment in Nigeria." *Heliyon* 6 (5): e03829. <https://doi.org/10.1016/J.HELIYON.2020.E03829>.
- Rettig, Erica M., and Robert J. Hijmans. 2022. "Increased Women's Empowerment and Regional Inequality in Sub-Saharan Africa between 1995 and 2015." *PLOS ONE* 17 (9): e0272909. <https://doi.org/10.1371/journal.pone.0272909>.
- Rowlands, Jo. 1995. "Empowerment Examined." *Development in Practice* 5 (2): 101–7. <https://doi.org/10.1080/0961452951000157074>.

- Seymour, Greg, Hazel Malapit, and Agnes Quisumbing. 2020. "Measuring Time Use in Developing Country Agriculture: Evidence from Bangladesh and Uganda." *Feminist Economics* 26 (3): 169–99. <https://doi.org/10.1080/13545701.2020.1749867>.
- Sinharoy, Sheela, Yuk Fai Cheong, Greg Seymour, Jessica Heckert, Erin Johnson, and Kathryn M. Yount. 2023. "The Time-Use Agency Scale: Development and Validation of a Measure for Ghana and Beyond." *Feminist Economics*. <https://doi.org/10.1080/13545701.2023.2262476>.
- Sinharoy, Sheela, Erin Johnson, Agnes R. Quisumbing, Hazel Jean Malapit, Talip Kilic, Md Zahidul Hassan, Shelton S. E. Kanyanda, et al. 2023. "The Women's Empowerment Metric for National Statistical Systems (WEMNS) Module: Development and Psychometric Assessment of a Preliminary Phone Survey Instrument." SSRN Scholarly Paper. Rochester, NY. <https://doi.org/10.2139/ssrn.4528648>.
- Westen, Drew, and Robert Rosenthal. 2003. "Quantifying Construct Validity: Two Simple Measures." *Journal of Personality and Social Psychology* 84 (3): 608–18. <https://doi.org/10.1037/0022-3514.84.3.608>.
- Willis, Gordon B. 2004. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks: SAGE Publications Ltd.
- Willis, Gordon B., and K Miller. 2011a. "Cross-cultural Cognitive Interviewing: Seeking Comparability and Enhancing Understanding." *Field Methods* 23 (4): 331–41.
- Willis, Gordon B., and Kristen Miller. 2011b. "Cross-Cultural Cognitive Interviewing." Edited by Gordon B. Willis and Kristen Miller. *Field Methods* 23 (4): 331–41. <https://doi.org/10.1177/1525822X11416092>.
- Yount, Kathryn M., Zara Khan, Stephanie Miedema, Yuk Fai Cheong, and Ruchira T. Naved. 2020. "The Women's Agency Scale 61 (Was-61): A Comprehensive Measure of Women's Intrinsic, Instrumental, and Collective Agency." *SSRN Electronic Journal* 61:1–48. <https://doi.org/10.2139/ssrn.3670180>.
- Yount, Kathryn M., Agnes R. Quisumbing, Ruth S. Meinzen-Dick, Hazel Jean Malapit, Md Zahidul Hassan, Shelton S. E. Kanyanda, Sudhindra Sharma, et al. 2023. "Women's Empowerment Metric for National Statistical Systems (WEMNS): Development and Psychometric Properties of a Face-to-Face Survey Module." SSRN Scholarly Paper. Rochester, NY. <https://doi.org/10.2139/ssrn.4657841>.
- Zapata, Oscar. 2022. "Happiness in the Tropics: Climate Variables and Subjective Wellbeing." *Environment and Development Economics* 27 (3): 250–71. <https://doi.org/10.1017/S1355770X21000267>.
- Zhang, Xin, Xiaobo Zhang, and Xi Chen. 2017. "Happiness in the Air: How Does a Dirty Sky Affect Mental Health and Subjective Well-Being?" *Journal of Environmental Economics and Management* 85 (September):81–94. <https://doi.org/10.1016/j.jeem.2017.04.001>.

APPENDIX I

Table 10. Item sets, Sustainable Development Goal linkages, and sources of item sets included in WMENS

Item set	SDG Target linked to the item set	Content	Current module informed by previous work by
Influence over decisions on time spent on both paid and unpaid activities	5.4 (value of unpaid work)	Awareness of inequalities on how women and men allocate their time. Activities respondent spent time on in past seven days. Ability of the respondent to influence how they use their time.	Sinharoy et al. 2023; Eissler et al. 2022 Sinharoy et al. 2023; Eissler et al. 2022
Participation and leadership in community organizations	5.5 (full participation in leadership)	Respondent's confidence to participate in community groups. Whether respondent participated in groups Whether respondent acted as a leader in groups Attitudes about women's participation in groups Attitudes about men's participation in groups	Yount et al. 2020
Life transitions and awareness of rights	5.a (reforms to give women equal rights to economic resources) 5.3 (eliminate harmful traditional practices)	Attitudes on women's freedom to choose her own preferences in livelihoods and family formation.	
Access to services and ownership of resources	1.4 (equal rights to economic resources and basic services) 9.3 (access to financial services and credit) 16.9 (legal identity for all) 5.b (use of technologies to empower women)	Ownership and use of different financial services. Access to different forms of credit. Possession of government-issued ID. Ownership of a mobile phone Access to key services (dependent care, healthcare, and transportation)	
Property ownership	1.4 (equal rights to economic resources and basic services)	Ownership and security of tenure of land and other property	
Decision-making	1.4 (equal rights to control and ownership over land and property) 5.6.1 (makes decisions on reproductive healthcare)	Control over a source of income and influence in own health decisions.	Demographic and Health Survey Program
Access and use of information and communication technologies	5.b (use of technologies to empower women)	Use of key information and communication technology services	
Perceptions of sexual harassment in the community	10.3 (eliminate discrimination)	The acceptability of different forms of sexual harassment.	Heckert et al. in development

APPENDIX II. DESCRIPTION OF COGNITIVE INTERVIEWING IN BANGLADESH, GUATEMALA, MALAWI, AND NEPAL

Cognitive interviewing is a qualitative approach used to identify discrepancies between how researchers ask questions and how respondents interpret the question, recall the information requested, or answer (Willis and Miller 2011a). Cognitive interviewing examines four cognitive processes (Willis 2004):

- Comprehension: a respondent's understanding of a question's content, including key terms;
- Retrieval: recall of the information in question, e.g., a time period stated in the question;
- Judgement: discomfort a respondent may feel while answering, e.g., social desirability bias; and
- Response: the ease of replying to a question in its suggested format, specifically being able to choose from pre-specified response options.

One goal of cognitive interviewing, and our primary purpose, is to use the results to revise survey items that elicit more accurate answers (Willis and Miller 2011a). Additionally, if the sample is designed with respondents from different groups (e.g., genders, language groups), cognitive testing may reveal if respondents from these groups interpret survey items differently and can help maintain validity across these populations (Hannan et al. 2019; Willis and Miller 2011a).

We cognitively tested all eight item sets in all four pilot countries. In Bangladesh, Malawi, and Nepal, cognitive interviewing was conducted before the first wave of phone surveys, and revisions informed the version of the questionnaire that was piloted in the first wave of phone surveys. In Guatemala, cognitive interviewing was conducted following the phone surveys, and revisions informed the in-person survey.

Due to limitations on in-person data collection during the COVID-19 pandemic, cognitive interviews were conducted over the phone. Interviewers were trained virtually on the cognitive interviewing instrument and related research protocols. For Bangladesh, Malawi, and Nepal, trainings occurred in June 2021 and data collection began immediately afterward in July 2021. For Bangladesh, we randomly selected households from the endline Feed the Future Initiative survey. The Bangladesh cognitive interviewing sample excluded households that had participated in the Bangladesh Integrated Household

Survey (BIHS), as this survey included questions about women’s empowerment, and we did not want to include respondents with prior exposure to empowerment questions in the cognitive interviewing. For the Malawi cognitive interviewing sample, we used the Fifth Integrated Household Survey as the sampling frame from which to draw respondents. We randomly selected households from this sampling frame, including an equal number of households from rural and urban settings. For the Nepal cognitive interviewing sample, we used a household census that had been completed in two districts (Kapilvastu and Rupandehi) in April-May 2019 as the sampling frame from which to draw respondents. We randomly selected households from this sampling frame. In these three countries, we then selected one woman or one man from each household, between the ages of 18 and 64. We interviewed 20 women and 20 men in Bangladesh, 24 women and 24 men in Malawi, and 20 women and 20 men in Nepal (Table 11).

For Guatemala, trainings occurred in September 2022, and data collection occurred in October 2022. Cognitive interviewing respondents were selected from those that had participated in the earlier WEMNS phone survey (see Heckert et al. 2023) for information about the Guatemala phone survey sampling) and spoke either Spanish, K’iche’, or Q’eqchi’, to account for the linguistic diversity in Guatemala.¹³ We interviewed 26 women and 26 men in Guatemala (Table 11).

Table 11. Women and men sampled for cognitive interviewing by country

Country	Women	Men	Total
Bangladesh	20	20	40
Guatemala	26	26	52
Malawi	24	24	48
Nepal	20	20	40
Total	90	90	180

In three countries, cognitive testing was carried out in one language: Bangla in Bangladesh, Chichewa in Malawi, and Nepali in Nepal. In Guatemala, interviews were conducted in Spanish, K’iche’, and Q’eqchi’. The interviews were structured so that after a module collecting basic respondent characteristics, each item set was asked of the respondent. Each item set was followed by scripted, but open-ended questions to understand how respondents interpreted and responded to specific survey items.

¹³ Only one person in each household had participated in the phone survey, and therefore only one person per household was eligible to participate in the cognitive interviewing in Guatemala.

Interviewers took detailed notes on paper forms to record responses. Data were transcribed and translated into English and entered into Excel. The open-ended questions were analyzed using inductive codes, developed by identifying emerging themes directly from the data. Fieldnotes were reviewed for additional information to contextualize the findings as available and needed.

Overall, most survey items were well understood and easily answered in all four countries. Table 12 shows the types of cognitive error found in each module by country.¹⁴ Comprehension was the most common type of error and was found in all eight of the modules tested. Response errors were found in five modules, while judgement errors were found in one. Finally, no recall errors were identified in any of the item sets. Ultimately, the research team discussed the findings from all four countries and revised the survey items to reword items more clearly, amend response options, add questions, and eliminate questions (Table 12). Table 13 provides select examples (from four item sets) of the revisions made to items based on the results of the cognitive interviewing.

¹⁴ Fewer errors were identified in Guatemala possibly due to the WEMNS item sets having undergone more revisions by the time cognitive interviewing was conducted in there.

Table 12. Key cognitive interviewing findings by country

Item sets	Errors identified				Revisions
	Bangladesh	Guatemala	Malawi	Nepal	
Influence over decisions on time spent on both paid and unpaid activities	Comprehension Response	Comprehension	Comprehension Response	Response	Reword questions Revise response options
Participation and leadership in community organizations	Comprehension		Comprehension	Response	Reword questions Revise response options Add question Eliminate questions
Life transitions and awareness of rights	Comprehension		Comprehension		Reword questions Revise response options Add questions Eliminate questions
Access to services and ownership of resources	Comprehension Response		Comprehension Response Judgement	Response	Reword questions Revise response options Eliminate questions
Property ownership	Comprehension Response	Comprehension	Comprehension Response		Reword questions Revise response options
Decision-making and control over income			Comprehension Response		Reword questions
Access and use of information and communication technologies	Comprehension		Comprehension		Eliminate questions Revise response options
Perceptions of sexual harassment in the community	Comprehension Judgement		Comprehension Judgement	Judgement	Reword questions Revise response options Eliminate questions

Table 13. Select survey items cognitively tested and revised survey items for WEMNS questionnaire

Item set	Type of revision	Survey item cognitively tested	Revised survey item	Justification for revision
Influence over decisions on time spent on both paid and unpaid activities	Reword question Revise response option	You are satisfied with how much time you spend caring for household members, such as children or elderly family members. 1 = Not at all satisfied 2 = Somewhat satisfied 3 = Very satisfied	During the last 7 days, did you spend any time on [ACTIVITY], even if for one hour? Caring for household members, such as children or older family members. 1 = Yes 2 = No	Re-worded to clarify intent to respondents Revised response option to reflect re-worded question
Participation and leadership in community organizations	Add question	[none]	In the last 12 months, have you participated in [ORGANIZATION]? Groups that provide local services, such as resource user groups, community health and education service groups, or mutual aid groups 1 = Yes 2 = No	Added questions based on cognitive interview findings that identified additional community groups salient to respondents
Life transitions and awareness of rights	Reword question Revise response option	Now I'm going to read some statements to you. Please tell me whether you do not agree, somewhat agree, or strongly agree with each statement. Every woman should be able to finish secondary school and study further if she chooses. 1 = Do not agree 2 = Somewhat agree 3 = Strongly agree	Now I'm going to read several statements of <u>PERCEPTIONS OF WOMEN'S RIGHTS</u> . Please listen to each statement and indicate whether you fully disagree, partly disagree, partly agree, or fully agree with each statement. Every woman should be free to choose whether to complete secondary school. 0 = Fully disagree 1 = Partly disagree 2 = Partly agree 3 = Fully agree	Reworded question to more accurately introduce the next survey topic (this rewording was not based on cognitive interview findings) Reworded question to clarify its intent to respondents Revised response options to better account for varying levels of agreement by respondents
Perceptions of sexual harassment in the community	Eliminate question	In your community, do most people think it is acceptable for a man to make a woman afraid that she will be treated poorly, for example threaten her or not give her work, if she does not cooperate sexually? 1 = Yes 2 = No	[none]	Eliminated question as respondents found the phrasing confusing and difficult to differentiate between the meaning of this survey item and others in the same item set

APPENDIX III

Table 14. WEMNS headline results by gender and country using alternative empowerment cutoffs

Empowerment cutoff		Bangladesh		Malawi		Nepal		Guatemala	
		Women	Men	Women	Men	Women	Men	Women	Men
60%	WEMNS score	0.55	0.76	0.63	0.78	0.47	0.51	0.46	0.55
	Disempowerment headcount ratio (H_p)	0.78	0.45	0.65	0.41	0.82	0.80	0.82	0.73
	Mean disempowerment score among disempowered (A_p)	0.58	0.53	0.57	0.55	0.64	0.61	0.65	0.62
65%	WEMNS score	0.53	0.72	0.60	0.74	0.46	0.49	0.44	0.52
	Disempowerment headcount ratio (H_p)	0.84	0.57	0.72	0.51	0.86	0.86	0.87	0.81
	Mean disempowerment score among disempowered (A_p)	0.57	0.50	0.55	0.51	0.63	0.59	0.64	0.59
Number of observations		797	797	787	773	809	802	796	779

Note: WEMNS = $1 - (H_p \times A_p)$

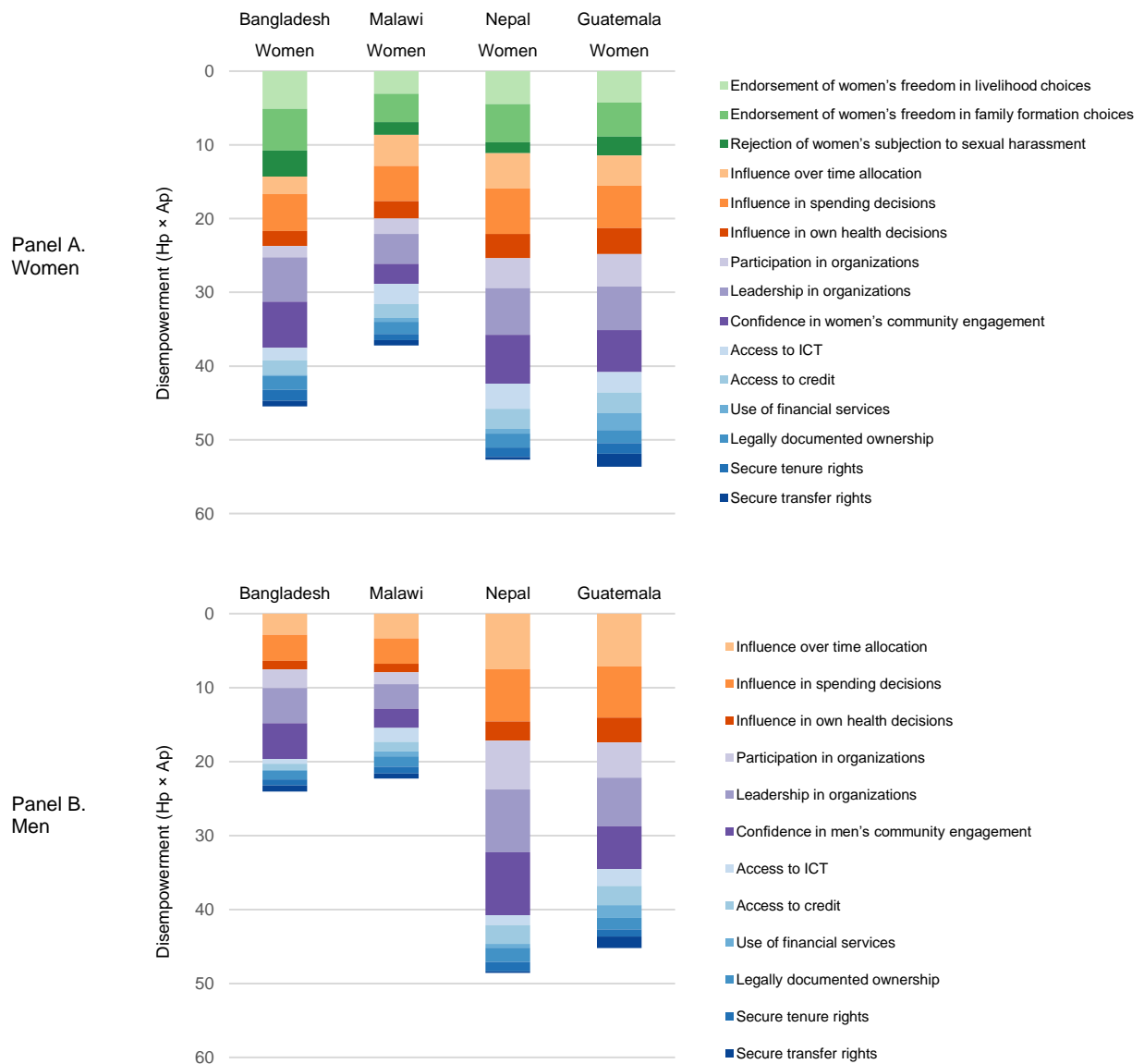


Figure 4. Contribution of each indicator to disempowerment by pilot study setting and gender, based on 60% empowerment cutoff

Source: Authors' calculations.

Note: The green, orange, purple, and blue shadings used in the figure correspond to the intrinsic agency, instrumental agency, collective agency, and agency-enabling resources domains, respectively.

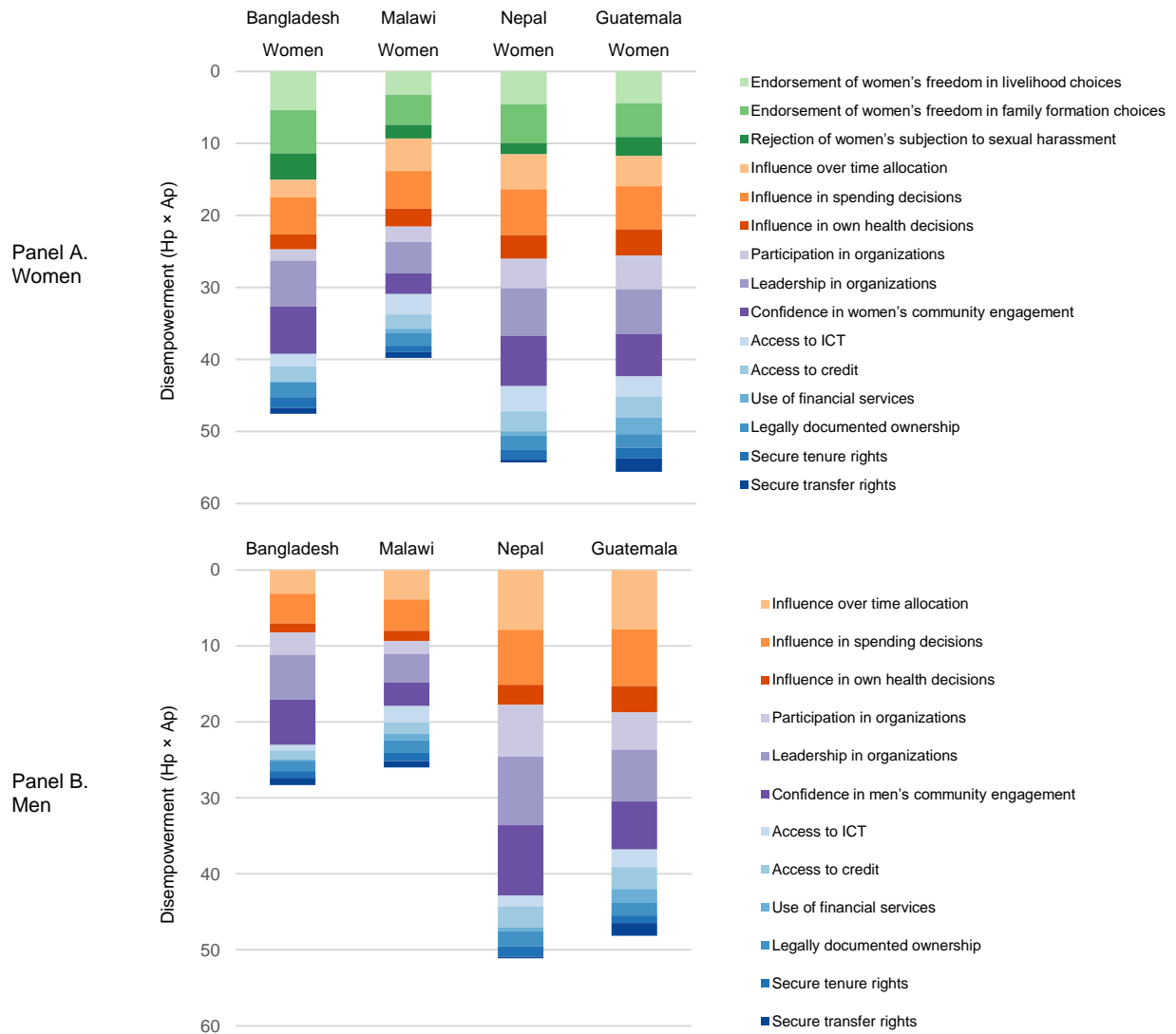


Figure 5. Contribution of each indicator to disempowerment by pilot study setting and gender, based on 65% empowerment cutoff

Source: Authors' calculations.

Note: The green, orange, purple, and blue shadings used in the figure correspond to the intrinsic agency, instrumental agency, collective agency, and agency-enabling resources domains, respectively.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available [here](#)

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org