World Values Survey Wave 7 in Indonesia: Sample Design.

The fieldwork for the 7th wave of World Values Survey was conducted by by SurveyMeter Indonesia (director: Wayan Suriastini). However, the University of Melbourne team also audited sample design, reviewed the pilot study, and attended in person the pre-fieldwork training week for interviewers in Yogyakarta to ensure consistency and to verify questionnaire validity.

The survey fieldwork took place from June to August 2018. The target population was private household residents of Indonesia over the age of 18. A total of 3,200 individuals were interviewed.

Provinces were selected to cover all regions of Indonesia, including those excluded from previous waves of the World Values Survey in the country. The sample design included fieldwork in 20 out of 34 provinces, including all of the most populous provinces of the country and covering all major regions/islands. This was by far the most comprehensive regional coverage of any WVS project in Indonesia.

In each province, villages and cities were grouped into urban and rural categories based on the classification of the Central Bureau of Statistics. The same number of rural and urban areas were selected at each selected province. Targeted village/cities were selected with simple random sampling from each group, i.e. the village (rural area) and cities (urban area), except for DKI Jakarta, which has no rural areas. By contrast in Papua and West Papua provinces, where the ratio of rural to urban population was highest, the number rural areas that were chosen was increased relative to the number of urban areas.

Households were listed in each selected village or city in each location by first obtaining a list of administrative areas under the village (RW). RWs were selected randomly (simple random). The list of households in each selected was obtained from the head of the RT (neighbourhood) in that selected RW. The list contained the household head name, number of households, member age (18-35, age 36-50, age 50+) by age group. Completion of the household listing at an early stage was considered as an advantage since it made sample more accurate and sample replacement was also able to be planned in advance. At the end of the listing exercise, the responsible organization obtained a sampling frame of households with members 18 years and above by age group and gender from which a sample for the village/city was drawn.

10 respondents were interviewed per PSU. When there was more than one individual of the same group and gender, random sampling was employed to select the respondent.

Interviews were conducted face-to-face with a laptop/tablet, in the respondent's household.

A CAPI (Computer Assisted Personal Interview) system was used. The data entry program was designed as soon as the translated survey instrument had been finalized which gave sufficient time to test the program to be ready for the training of editors and interviewers. This was first done internally and then, during a full pilot study in the field prior to commencing fieldwork. The programmer designed a suitable data entry program to enter data from all interviews. The use of a pre-design data entry program to be used for directly edits and enter data in the field had been proved efficient in checking errors and directly could be corrected in the field.

The study was conducted on working days and weekends (excluding public holidays and

Ramadan). Nationally-representative health and demographic surveys in Indonesia generally time the fieldwork based on the cycle of respondents' availability i.e. in order to limit response bias among non-working female respondents in the sample. Therefore, interviews were conducted during the daytime, which made it possible to have a private interview and sufficient time, but for sampling of employed respondents, fieldwork continued in the early evening or at night after work, as well as during the weekend, in order to maintain a representative sample of the target demographic.

SurveyMeter's Quality Assurance for Data Integrity

To maintain the integrity and high quality of data that were collected, SurveyMeter had put in place layers of quality assurance procedures and protocols. To summarize, the quality assurance for data integrity consisted of:

Built-in quality control: The first layer of quality control was built in through the CAPI system developed for the project. SurveyMeter operated a field-based computer data entry system as the method of data entry to support data collection of the surveys. The system had proven to be more advantageous than centralized data entry at the head office after the data collection had been completed. The time lag was minimized for the questionnaire to return to the field for further clarification. Quality control by field personnel was facilitated through the ability to audit interviews in real time. This method provided keyed and clean data more quickly to the central office for further consolidation and evaluation and flag irregularities immediately.

SurveyMeter's CAPI system provided full audio recording of all interviews, which had proven to be instrumental in maintaining data quality and integrity.

Other features built-in into the program:

- User-friendly: data entry was made as easy as possible for the user. A special menu was created for data entry to speed up the work of the fieldwork team and minimize procedural mistakes in data entry.

- Version control management. Installations and backups were named using a special standard in order to avoid mistakes in file usage or deletion. File version management was designed as a response to responsible organization's experience (there were times when the field editor used old/wrong application where skip pattern or consistency check was out of date). For installation files, log files were generated so that editor was able to track any changes in the application.

- External references. The application was merged with external reference data, such as the geocode/georeferencing of the location, panel respondent references, or other references whenever required to improve the quality of the data.

- The data entry form was visually designed to be as similar as possible to the questionnaire. This design eased the editor to check and compare the questionnaire and the data entry form such that he/she was able to speed up data entry work and reduce mistakes during the data entry process.

- Variable tree, this feature helped editor to find variables and all variables that were available in data entry.

- Consistency checks or "lookup" that were conducted directly during data entry or after all data entry was completed. During data entry, error messages appeared whenever inconsistency was detected.

- Missing value checks in order to minimize missing data.

- A recapitulation program to check consistency between variables, range and skip pattern, missingness checks and data completion. All checks were referred to the target data. The editor conducted a final check based on recapitulation output.

- A good backup system to avoid data loss during data entry in the field. The system was integrated into main menu system, so editor easily implemented this procedure.

- Web-based data transfer. To improve data transfer to the center, a password-protected internet-based data transfer was used, which checked the data during upload and notified the user if any problems were found in the data. The tool also had a file management system to improve data checking.

Quality control during data collection

A management information system during data collection facilitated the interaction between editors and interviewers. Each enumerator's completed questionnaire was edited by the CAPI editor and any errors were communicated back to him/her. The enumerators made corrections of the errors and when necessary, she/he revisited the respondent to obtain/verify the correct information. These communications were documented electronically in order to flag any fieldwork issues as they arise.

The supervisor was responsible for maintaining a consistent and high quality of data collection in the field through spot-checking the completed questionnaire (all), verification of the interviews (10%) and interview observation (5%). As noted, <u>interviews were also</u> <u>time-stamped, recorded and geocoded for central auditing including the auditing of</u> <u>supervisors</u>. The supervisor was in constant communication with the Data Manager at the central office who monitored and checked the data quality of all data sent to the central office. Supervisors were aware that they were audited in real time and held to account for the work of their team.

In addition, the Team Leader, Assistant Team Leader and Programmer were in the field in the first week of data collection for direct supervision, observation, data quality checks and fieldwork verification.

Quality control and assurance at the central office.

The Data Manager monitored and checked all data received from the field for completeness of the information. <u>The data cleaning process included provision of random audio sampling</u> <u>from the recordings to the head office in Yogyakarta,</u> for inspection and audit by independent data specialists. On this basis errors were corrected and clean data were obtained. This

involved a team of experienced data cleaners and editors and was coordinated and managed by the Data Manager. Final and additional outlier check and consistency check was made between variables in addition to build in check in data entry program.

The supervisor was responsible for on-the-spot maintenance of high data quality through spot-checking the completed questionnaire (all interviews), verification of interview to the respondent (10%) and interview observation (5%). In addition to this, all interviews were controlled centrally through automated recording, time-stamping and geocoding of interviews through data sent back to the central office via the CAPI interface.

The respondents to be re-interviewed (less than 1% of the sample) were selected based on demographic variables. Due to the introduction of CAPI surveys with central recording and auditing, the responsible organization did not experience fieldwork irregularities, beyond minor fieldworker mistakes, which were generally corrigible during the fieldwork process.

Weighting was applied when necessary.

The difference in percentage between sex and age groups in the survey data compared to census data:

	% in census data	% in your survey	Δ %
Male	50.3%	50.0%	-0.3%
Female	49,7%	50.0%	+0,3%
18-35	29.1% (44.2%)	40%	-4.2%
36-50	20.7% (31.4%)	30%	-1.4%
50+	16.0% (24,4%)	30%	+5.6%

10 respondents were interviewed per PSU. The distribution of these respondents was as follows:

Two male respondents aged 18-35 years, two female respondents aged 18-35 years, two male respondents aged 36-50 years, one female respondent aged 36-50 years, one male respondent aged 50 years and over and two female respondents aged 50 years and over.

In summary, the male respondent sampling rate was 50.0% and the female respondent rate was 50.0%; by age respondents were18-35 years, 40%, age 36-50 years, 30%, and 50+ years, 30%.

The conditions that were taken into account during the planning of fieldwork by region:

- Finding working/employed respondents during daytime fieldwork: as a result, interviews were phased in to evenings, night and weekends to ensure consistency between population and the survey samples with respect to sociodemographic identifiers.

- In remote island regions of Indonesia, such as Papua and West Papua, many of the survey target population respondents are located in highly remote areas with poor infrastructure

and limited (or unreliable/out-of-date) village/household location information. For rural survey fieldwork in these regions, local guides were used in order to identify village and distribution and household location.

- Non-compliance among urban respondents who report being too busy to participate: as a result, interviews were divided based on longer questionnaire designs into phased (repeat) visits. Non participation / refused participation rates are, generally, significantly lower in Indonesia than surveys conducted in developed countries, but sample bias was taken seriously and hence, the responsible organization sought to reduce the gap between urban and rural non- participation rates to the minimum possible.

- Fieldwork irregularities have been a feature of early surveys in Indonesia, for this reason, SurveyMeter with the World Bank / RAND had developed a strict CAPI-based system for recording and live auditing of both fieldworkers and supervisors. This was a unique comparative advantage offered by SurveyMeter, for delivering the highest possible quality of fieldwork in Indonesia.

The responsible organization needed permission from the government to do WVS-7. However, it was easy to get due to SurveyMeter's extensive experience in working with regional entities and its established reputation for work that is non-partisan / primary for the use international development and academic research.