

General Population Survey Methodology Report

October 31, 2018

Contents

1. Study Objectives	5
2. Study Overview.....	5
3. Sample Design and Selection	5
3.1. Sample frame	5
3.2. Sample Design	8
3.3. Calculating the Sample Size.....	8
3.4. Sample Allocation.....	9
3.5. Weights.....	12
3.6. Relative Weight	14
3.7. Sketches from the Department of Statistics (DOS):	14
Replacement Processes.....	14
4. Respondent selection	14
5. Response Rates	15
6. Questionnaire Review	15
7.1. Translation Process:.....	16
7.2. Scripting Process	17
8. Interviewers and Training.....	17
Day 1 – Topics Covered:	17
Day 2 – Topics Covered:	17
Day 3 – Topics Covered:	18
Day 4 – Topics Covered:	18
9. Field Team Composition	18
10. Quality Control Measures	18
10.1. Through field supervisors:	19
10.2. Through the QC team:.....	19
10.3. Through the back-check team:	19
10.4. Through the data analyst:.....	19
11. Coding and Data Entry	20

11.1.	Coding	20
11.2.	Data Cleaning and Processing.....	20
12.	Project Schedule.....	21
13.	Study Challenges	22
Appendix: Post-Research Data Processing and Quality Control Procedure Checklist.....		24
Data Quality/Cleaning Checklist		24
Sight Checks.....		24
Data Cleaning		24
Perform Logic Checks (marginal/crosstabs).....		24
Check Para/Meta Data		24
Interviewer Checks.....		24
Duplicates.....		25

1. Study Objectives

The USAID Monitoring and Evaluation Support Project (MESP) General Population Survey is one of Jordan's largest national surveys of attitudes and opinions. The survey consisted of 11,963 face-to-face interviews with adults across the Kingdom. The survey gathered critical USAID mission indicator data. Specifically, the survey gathered data on the opinions of Jordanians and residents about the following:

- General outlook
- Satisfaction with public services
- Education
- Civic participation
- Women's role in society
- Employment
- Entrepreneurship

2. Study Overview

The original sample size of the survey was 12,000 interviews distributed across the 12 governorates in the Kingdom. A team of 100 enumerators and 25 supervisors collected data for the survey. In total, the survey had a response rate of 76%, which is significantly higher than the global average for surveys. The average interview duration was 35 minutes. The sample was proportionally divided to reflect the population of each governorate. Sample weights were also developed to make the data representative of the population in Jordan.

Preparation for the survey began in February 2018 and fieldwork started on June 26, 2018 and was completed on September 16, 2018.

3. Sample Design and Selection

3.1. Sample frame

The sample frame of this survey is based on the latest Population and Housing Census of 2015. The sample frame was provided by the Department of Statistics. Administratively, Jordan is divided into 12 governorates, each governorate is divided into districts, each district is divided into sub-districts, and each sub-district is divided into municipalities. However, the DOS frame does not go down to the municipality level; it stops at the sub-district level. For the purpose of this survey, we need to add another layer: municipalities. Using data from the Ministry of Municipal Affairs, we matched the sub-districts to municipalities. The frame of municipalities became our secondary frame, and we used it to request a random sample of clusters from the DOS. Therefore, the final frame contained 100 municipalities, as shown in the table below.

Table 1: Description of DOS Frame Including Administrative Structure

Region	Governorate	District	Municipality
Central	Amman	Hosba'n	Hosba'n
Central	Amman	Jizah	Jizah
Central	Amman	Muaqqar	Muaqqar

Central	Amman	Na'oor	Na'oor
Central	Amman	Sahab	Sahab
Central	Amman	Um Alrasas	Um Alrasas
Central	Amman	NA	Um Elbasatien
Central	Amman	Na'oor	Amireah
Central	Balqa	Ain Albasha	Ain Albasha
Central	Balqa	Al Ardha	Al Ardha
Central	Balqa	Alshoneh Alwasta	Alshoneh Alwasta
Central	Balqa	Dair Alla	Dair Alla
Central	Balqa	Fuhais	Fuhais
Central	Balqa	M'addi	M'addi
Central	Balqa	NA	Salt Kubrah
Central	Balqa	Mahes	Mahes
Central	Balqa	Alshoneh Aljnobe	Swaimah
Central	Madaba	Dieban Jadedah	Dieban Jadedah
Central	Madaba	Qasibah Madaba	Madaba Alkubrah
Central	Madaba	Jabal bani Hamedah	Jabal bani Hamedah
Central	Madaba	NA	Leb & Mlaih
Central	Zarqa	Al Hashemiyah	Al Hashemiyah
Central	Zarqa	Dhlail	Dhlail
Central	Zarqa	NA	Russeifa
Central	Zarqa	Qasibah Zarqa	Zarqa
Central	Zarqa	NA	Azraq
Central	Zarqa	Zarqa	Bierain
Central	Zarqa	Zarqa	El-Hallabat
North	Ajlun	Alauion	Alauion
North	Ajlun	Janed	Janed
North	Ajlun	Kufrangeh Aljadedah	Kufrangeh Aljadedah
North	Ajlun	Ajlun	Ajlun Alkubrah
North	Ajlun	Shafa	Shafa
North	Irbid	Al Seru	Al Seru
North	Irbid	Alkfarat	Alkfarat
North	Irbid	Brkash	Brkash
North	Irbid	Dair Abi Sa'id Jadedah	Dair Abi Sa'id Jadedah
North	Irbid	Garb Irbid	Garb Irbid
North	Irbid	Khaled Abn Alwaleed	Khaled Abn Alwaleed
North	Irbid	Mazar Jadedah	Mazar Jadedah
North	Irbid	Mo'ath Abn Jabal	Mo'ath Abn Jabal
North	Irbid	Irbid	Irbid Alkubrah
North	Irbid	Rabeat Al Koorah	Rabeat Al Koorah
North	Irbid	Ramtha Jadedah	Ramtha Jadedah
North	Irbid	Sahel Horan	Sahel Horan
North	Irbid	Sharhabeel Abn Hasnah	Sharhabeel Abn Hasnah
North	Irbid	Alshoneh Alshmalieh	Tabaqat Fahl
North	Irbid	Taybeh Jadedah	Taybeh Jadedah
North	Irbid	Wastiyyah	Wastiyyah
North	Irbid	Bani Kenana	Al Shoaleh
North	Irbid	Bani Kenana	Alyarmook Aljadedah

North	Jarash	Alm'arad	Alm'arad
North	Jarash	Alnasim	Alnasim
North	Jarash	Qasibah Jarash	Jarash Alkubrah
North	Jarash	Bab Amman	Bab Amman
North	Jarash	Borma	Borma
North	Mafraq	Alrwashed Aljadedah	Ruwashid Aljadedah
North	Mafraq	Bal'ama Aljadedah	Bal'ama Aljadedah
North	Mafraq	NA	Mafraq Alkubrah
North	Mafraq	Rhab Aljadedah	Rhab Aljadedah
North	Mafraq	Sabha & Defianeh	Sabha & Defianeh
North	Mafraq	Um aljmal Aljadedah	Um aljmal Aljadedah
North	Mafraq	Albadiuh alshamaliuh algharbiuh	Alza'tary & Almansheah
North	Mafraq	Albadiuh alshamaliuh	Aum Qutain & Makfieah
North	Mafraq	Albadiuh alshamaliuh	Dair Alkahf Aljadedah
North	Mafraq	NA	Husha Aljadedah
North	Mafraq	Khaldiyyah Jadedah	Khaldiyyah
North	Mafraq	Mafraq	Manshiat Bane Hasan
North	Mafraq	NA	Prince Alhusain Ben Abdollah
North	Mafraq	NA	Salhiah & Naifeh
North	Mafraq	NA	Asafawi
North	Mafraq	Albadiuh alshamaliuh	Bani Hashem
North	Mafraq	NA	Basleah
North	Mafraq	Albadiuh alshamaliuh algharbiuh	Serhan
South	Aqaba	QuairahAljadedah	QuairahAljadedah
South	Aqaba	AlQweira	Hud Aldisah
South	Aqaba	Wadi Araba	Wadi Araba
South	Aqaba	Wadi Araba	Krekrah & Finan
South	Aqaba	Wadi Araba	Qatar & Rahmah
South	Karak	Abdulah Bin Ruaha	Abdulah Bin Ruaha
South	Karak	Ghawr Safi & Almazra'a	Ghawr Safi & Almazra'a
South	Karak	Hazman	Hazman
South	Karak	Mu'ab Aljadedah	Mu'ab Aljadedah
South	Karak	Mu'ata & Almazar	Mu'ata & Almazar
South	Karak	NA	Karak Alkubrah
South	Karak	Qatranah	Qatranah
South	Karak	Shehan	Shehan
South	Karak	NA	Talal Aljadedah
South	Karak	Qatranah	Sultani
South	Maan	Husanieh Aljadedah	Husanieh Aljadedah
South	Maan	NA	Ma'an
South	Maan	Shobak Aljadedah	Shobak Aljadedah
South	Maan	Maan	Al Jafer
South	Maan	NA	Alashari
South	Maan	Maan	Alsharah
South	Maan	Maan	Iel Jadedah
South	Tafela	Al Hasa	Al Hasa

South	Tafela	Alhareth Bin Aumair	Alhareth Bin Aumair
South	Tafela	Qasibah	Al Tafeilah Alkubrah
South	Tafela	AlBasira	Qadesiah

Using the above frame, DOS randomly selected clusters from each municipality. Each municipality was therefore considered an independent stratum and the clusters were the primary sampling unit (PSU).

3.2. Sample Design

A stratified multistage cluster design was used to develop the sample for this survey. The design followed these stages:

- Stage 1: In the first stage, the primary sampling units were drawn systematically from each municipality by using probability proportional to size (PPS) to account for the differences in cluster populations. The number of households was the size parameter.
- Stage 2: In the second stage, the buildings in the clusters were numbered and ordered. From each cluster, a minimum of 10 buildings were randomly selected. The randomization ensured that the selected buildings were widely distributed in all geographical segments of the selected cluster. This also served to reduce sampling error as a result of intra-class correlation.
- Stage 3: In the third stage, one household was randomly selected from each building drawn in the second stage. If the building was a stand-alone home, it was selected without randomization.

3.3. Calculating the Sample Size

The sample size was calculated to achieve the confidence level required for this survey so that the main variables were not less than 98% at regional level, 95% at governorate level, and 95% at stratum level. The maximum accepted error is 5% at stratum level.

The following equation were used to estimate the sample size

$$n = n^0 \frac{1}{\left(1 + \frac{n^0}{N}\right)}$$

$$\text{where } n^0 = t^2 * p * q / e^2$$

n: sample size

t: desired level of confidence

P: percent of persons who have a character of interest (set at .5)

q: percent of persons who do not have a character of interest (set at .5)

N: total number of households in the stratum

e: tolerated margin of error

3.4. Sample Allocation

The sample size was calculated with a minimum confidence level (CL) of 90% and maximum accepted error of 10% at the stratum (municipality) level to estimate the minimum sample size in each stratum. A proportional allocation method using the 12,000 household sample was then used to allocate the sample.

Mindset used two methods: first, a minimum number of number of households was set to meet the required CL, and second, a proportional allocation was used to distribute the remaining sample among the other strata to maximize the accuracy of the results.

The table below shows the CL and margin of error per each municipality.

Table 2: Suggested Sample Allocation among Different Municipalities

District	Sub-District	Municipality	Sample Size	Confidence Level	Margin of error
Region: Central			4570	98.00%	5.00%
Governorate: Amman			1510	98.00%	5.00%
Amman capital	Amman capital	Amman capital	1000	98.00%	5.00%
Amireah	Na'oor	Amireah	120	90.00%	5.00%
Jizah	Jizah	Jizah	130	90.00%	5.00%
Na'oor	Na'oor	Na'oor	140	90.00%	5.00%
Um Al-Rasas	Um Al-Rasas	Um Al-Rasas	120	90.00%	5.00%
Governorate: Balqa			950	98.00%	5.00%
Ain Albasha	Ain Albasha	Ain Albasha	410	90.00%	5.00%
Alshoneh Alwasta	NA	Alshoneh Alwasta	180	90.00%	5.00%
Fuhais	Fuhais	Fuhais	120	90.00%	5.00%
Mahes	Mahes	Mahes	120	90.00%	5.00%
Alshoneh Aljnobe	Alshoneh Aljnobe	Swaimah	120	90.00%	5.00%
Governorate: Madaba			600	95.00%	5.00%
Dieban	Jabal bne Hamedah	Jabal bne Hamedah	140	90.00%	5.00%
NA	Qasibah	Madaba Alkubrah	460	90.00%	5.00%
Governorate: Zarqa			1510	98.00%	5.00%
Al Hashemiyah	Al Hashemiyah	Al Hashemiyah	280	90.00%	5.00%
Zarqa	Bierain	Bierain	120	90.00%	5.00%
Zarqa	El-Hallabat	El-Hallabat	120	90.00%	5.00%
Qasibah Zarqa	Zarqa	Zarqa	990	98.00%	5.00%
Region: North			4430	98.00%	5.00%
Governorate: Ajloun			600	95.00%	5.00%
Ajlun	Ajlun	Ajlun Alkubrah	300	90.00%	5.00%
Sakhra	Sakhra	Janed	170	90.00%	5.00%

Ajlun	Ajlun	Shafa	130	90.00%	5.00%
Governorate: Irbid			1990	98.00%	5.00%
Bani Kenana	Al Shoaleh	Al Shoaleh	120	90.00%	5.00%
Bani Kenana	Alyarmook Aljadedah	Alyarmook Aljadedah	120	90.00%	5.00%
Dair Abi Sa'id	Dair Abi Sa'id	Dair Abi Sa'id Jadedah	120	90.00%	5.00%
Irbid	Qasibah	Irbid Alkubrah	950	98.00%	5.00%
Mazar Alshмали	Mazar Alshмали	Mazar Jadedah	120	90.00%	5.00%
Dair Abi Sa'id	Dair Abi Sa'id	Rabeat Al Koorah	120	90.00%	5.00%
Ramtha	Ramtha	Sahel Horan	200	90.00%	5.00%
Alshoneh Alshmalieh	Alshoneh Alshmalieh	Tabaqat Fahl	120	90.00%	5.00%
Wastiyyah	Wastiyyah	Wastiyyah	120	90.00%	5.00%
Governorate: Jerash			600	95.00%	5.00%
Jerash	Qasibah	Alm'arad	300	90.00%	5.00%
Bab Amman	Bab Amman	Bab Amman	170	90.00%	5.00%
Jerash	Qasibah	Jarash Alkubrah	130	90.00%	5.00%
Governorate: Mafrq			1240	98.00%	5.00%
Al Badiuh alshamaliuh algharbiuh	Al Badiuh alshamaliuh algharbiuh	Alza'tary & Almansheah	120	90.00%	5.00%
Aum Qutain	Aum Qutain & Makfieah	Aum Qutain & Makfieah	120	90.00%	5.00%
Al Badiuh alshamaliuh	albaduih alshamaliuh	Bani Hashem	120	90.00%	5.00%
Al Badiuh alshamaliuh	Dair Alkahf Aljadedah	Dair Alkahf Aljadedah	120	90.00%	5.00%
Khaladiyah Jadedah	Khaladiyah	Khaladiyah	210	90.00%	5.00%
Mafrq	AlManshiat	Manshiat Bane Hasan	120	90.00%	5.00%
Mafrq	Rhab	Rhab Aljadedah	170	90.00%	5.00%
Al Badiuh alshamaliuh	Sabha	Sabha & Defiane	120	90.00%	5.00%
Al Badiuh alshamaliuh algharbiuh	Sama Serhan	Serhan	140	90.00%	5.00%
Region: South			3000	98.00%	5.00%
Governorate: Aqaba			850	95.00%	5.00%
Aqaba city	Qasibah	Aqaba city	540	95.00%	5.00%
Al quayruh	Al quayruh	Hud Aldisah	100	90.00%	5.00%
Wadi Araba	Wadi Araba	Qatar & Rahmah	120	90.00%	5.00%
Wadi Araba	Wadi Araba	Wadi Araba	90	90.00%	10.00%

Governorate: Karak			950	98.00%	5.00%
Faguo	Faguo	Abdulah Bin Ruaha	210	90.00%	5.00%
Hazman	Hazman	Hazman	130	90.00%	5.00%
Mu'ab Aljadedah	Mu'ab Aljadedah	Mu'ab Aljadedah	370	90.00%	5.00%
Qatraneh	Qatraneh	Qatraneh	120	90.00%	5.00%
Qatraneh	Qatraneh	Sultani	120	90.00%	5.00%
Governorate: Maan			600	95.00%	5.00%
Maan	Al Jafer	Al Jafer	120	90.00%	5.00%
Maan	Alsharah	Alsharah	130	90.00%	5.00%
Maan	Iel Jadedah	Iel Jadedah	150	90.00%	5.00%
Shobak	Shobak	Shobak Aljadedah	200	90.00%	5.00%
Governorate: Tafeilah			600	95.00%	5.00%
Qasibah	Al Tafeilah	Al Tafeilah Alkubrah	300	90.00%	5.00%
AlBasira	AlBasira	Qadesiah	300	90.00%	5.00%
National Sample			12000	98.00%	2.00%

The following table shows the number of clusters and households per each municipality.

Table 3: New Sample Allocation among Municipalities

Municipality	Number of Sample Clusters	Number of Sampled Households
Amireah	3	45
Amman Capital	100	1000
Jizah	13	130
Na'oor	21	210
Um Al-Rasas	12	120
Ain Al-Basha	43	430
Alshoneh Al-Wasta	21	210
Fuhais	12	120
Mahes	12	120
Swaimah	5	75
Al-Hashemiyah	28	280
Bierain	12	120
El-Hallabat	12	120
Zarqa	99	990
Jabal bne Hamedah	13	130
Madaba Al-Kubrah	47	470
Al-Shoaleh	12	120
Alyarmook Aljadedah	12	120
Dair Abi Sa'id Jadedah	12	120

Irbid alkubrah	95	950
Mazar Jadedah	12	120
Rabeat Al Koorah	12	120
Sahel Horan	20	200
Tabaqat Fahl	12	120
Wastiyyah	12	120
Alza'tary & Almansheah	14	140
Aum Qutain & Makfieah	10	100
Bani Hashem	6	90
Dair Alkahf Aljadedah	11	110
Khaldiyah	21	210
Manshiat Bane Hasan	13	130
Rhab Aljadedah	20	200
Sabha & Defianeh	12	120
Serhan	14	140
Alm'arad	30	300
Bab Amman	17	170
Jarash Alkubrah	13	130
Ajlun Alkubrah	30	300
Janed	17	170
Shafa	13	130
Abdulah Bin Ruaha	21	210
Hazman	14	140
Mu'ab Aljadedah	24	360
Qatraneh	12	120
Sultani	8	120
Al Tafeilah Alkubrah	36	360
Qadesiah	24	240
Al Jafer	10	150
Alsharah	8	120
Lel Jadedah	13	195
Shobak Aljadedah	9	135
Aqaba city	61	610
Hud Aldisah	9	90
Qatar & Rahmah	4	60
Wadi Araba	9	90
Total	1,155	1,2000

3.5. Weights

The sample was designed to be self-weighting at the stratum (municipal) level. However, due to inevitable non-response and replacements during implementation of the survey, post data collection adjustments through weights were needed to return the sample to its original design. Furthermore, weights were also needed for analysis at the governorate level. The basic weight for each household (HH) equals the reverse

of the probability of the selecting the HH in the sample (it was calculated by multiplying the probability of selecting at all stages of sample selection).

The following are the weight formulas that were used.

Calculation of the probability of selecting a sampling unit:

Probability of selecting the I cluster from h stratum (p_{hi})

$$p_{hi} = \frac{n_h \times M_{hi}}{M_h}$$

Where:

n_h = the number of primary sampling units to be selected from h stratum.

M_h = number of households in the h stratum from the frame

M_{hi} = number of households in the cluster I in the stratum h from the frame.

Probability of drawing the household from the cluster p_{hij}

$$p_{hij} = \frac{m_{hi}}{M_{hi}}$$

Where:

p_{hij} = The probability of selecting the j household from the i cluster from h stratum.

m_{hi} = Number of households selected from the cluster i from stratum h.

M_{hi} = number of households in the ith cluster in the h stratum.

The basic weight of household j in cluster i in the stratum h is equal to the inverse of probability of drawing the household in the sample and its code is W_{hij} and this is equal to:

$$W_{hij} = \frac{M_h}{n_h * m_{hi}}$$

It can be seen that if the number of households selected from each cluster at the stratum level is constant (as in the example of 10 households from each cluster), the sample will be self-weighting at the municipality level. It is important to adjust the weights to take into considerations non responses in each cluster. Due to the fact that the weights will be calculated at cluster level it is worth adjusting the weights at this level, and it is recommended that non responses be kept to the minimum to reduce bias and non-sampling errors. It is also recommended that the replacement samples be kept to a minimum, and replacements only done when required.

When the number of completed questionnaires is less than the selected number of households in the specific cluster, it is recommended that the basic weight be multiplied by the adjusted factor by applying this equation:

$$AdjW_i = \frac{m_{hi}}{m_{hi}^*}$$

Where:

$AdjW_i$: Adjusted factor for cluster i in the stratum h

m_{hi} : Number of households selected from the cluster i in the stratum h

m_{hi}'' : Number of completed questionnaires from the cluster i in the stratum h

So, the final weight for each household in the i th cluster in the stratum h and where the code is W_{hij} and the equation is:

$$W_{hij} = \frac{W_{hij} \times m_{hi}''}{m_{hi}''}$$

3.6. Relative Weight

A relative weight is calculated to change the sample to a self-weighting one. The relative weight is calculated for each observation and the summation of relative weights will be equal to the total number of observations. This method provides high flexibility during statistical analyses. For example, it can produce results at the regional level by making a summation of the governorate results. The relative weight for each household from a specific cluster is equal to the adjusted weight of the cluster divided by the result of mean weight multiplied by the number of completed questionnaires.

3.7. Sketches from the Department of Statistics (DOS):

The DOS provided maps or “sketches” of the selected clusters. The maps were satellite images of the selected clusters with a unique number for each building in the cluster. These maps were provided to the field teams to guide them in the identification of the preselected buildings.

Replacement Processes

Before starting data collection, each field team was given 10 preselected buildings from the core sample and 5 from the replacement sample. Due to refusals and non-residential buildings, some replacements of the samples were needed.

The replacement process was done per the following two levels:

- Replacement inside the block. There are two scenarios where we need replacement inside the block, and these are:
 - Building replacement: Requested in cases there was no answer from any of the houses inside the building after three attempts or in case the building itself was not a residential unit.
 - Household replacement: Requested in cases the household rejected the visit in the first place, or if the Kish Grid selected respondent was not available after three trials.
- Replacement of the block itself: In case we were not able to secure half of the required interviews for the block, we requested additional blocks within the same municipality.

4. Respondent selection

Within each randomly selected household, respondents were then randomly selected through a Kish grid to eliminate selection bias. The Kish grid was programmed on the tablet, which made it straightforward for all enumerators to follow. In the household, the enumerator listed all household members 18 years and above by name, age, and gender. The electronic Kish grid then randomly selected the respondent. Enumerators were instructed to list adult household members from the youngest to the oldest. The purpose of this instruction was to ensure that the listing did not miss any household member. However, the Kish grid was not dependent on the order, as the randomization was electronic and completely random. Under no circumstances were enumerators allowed to substitute an alternate member of a household for the selected respondent. If the respondent refused to participate or was not available after three call-backs, a replacement was sought.

5. Response Rates

The overall response rate for this survey was 76%. The response rate was calculated according to the standards of the [Marketing Research and Intelligence Association](#), a Canadian not-for-profit association “representing all aspects of the market intelligence and survey research industry”. The Empirical Method of Response Rate Calculation was the specific method used, and it is illustrated with actual figures from the current survey in the table below.

Empirical Calculation for Data Collection			
#	Cases	Figures from the survey	Notes
1	Total number of interviews attempted	33,569	
2	Invalid	2,178	Non-housing units
Unresolved (U)			
3	No answer	1,622	Could not determine if housing unit or not; could not gain access to building
In-scope - non-responding (IS)			
4	Language problem	4	Selected respondent did not speak Arabic
5	Illness, incapable	31	Selected respondent was incapable of participating in the interview
6	Selected respondent not available	2,688	
7	Household refusal	3,889	Person who received the interviewer did not allow him/her access to the household
8	Respondent refusal	113	Selected respondent refused to participate in the in the interview
9	Qualified respondent break-off	54	
In-scope - Responding units (R)			
10	Language disqualify	4	
11	No one 18+	13,087	
12	Other disqualify	1,288	Gender quota filled
13	Completed interviews	12,266	12,266 interviews were completed, but 303 interviews were cancelled during quality control. Therefore, the final number of interviews was 11963.
	Response Rate = R/(U+IS+R)	76%	

6. Questionnaire Review

MESP provided the questionnaire on February 5, 2018 which Mindset then reviewed and translated.

Mindset reviewed the questionnaire in three ways:

1. **Internal review and feedback:** Upon receiving the questionnaire from MSI/MESP, Mindset initiated an internal review of the questions and their flow.
2. **A mock / cognitive interview with four potential respondents:** Once the internal review and feedback were done, Mindset set out to gauge the understanding of the questionnaire with four cognitive interviews with two Jordanians and two Syrian respondents from rural and urban areas. These interviews were conducted between February 22, 2018 and February 25, 2018 for an average of 4 hours each. As a result of these interviews, Mindset saw the need to link the questionnaire questions to the nationality of the respondents, provide a comprehensive introduction and add details to the questions for clarity.
3. **A pilot session:** 188 ki0 interviews were conducted to pilot the questionnaire and test the electronic script. Enumerator feedback was taken into consideration and incorporated into the questionnaire edits.

Mindset submitted amendments to MESP for approval. The main changes concerned verifying question skipping rationale such as the below examples:

1. If the respondent hadn't elected, he/she were still asked how the elections affected their life.
2. If the respondent had no children, they were still asked how many times they visit their school
3. At the start of the interview, the enumerator could only continue or end the questionnaire based on the respondent's aptitude to proceed. However, another option had to be added in case the house was empty.
4. Question 10B didn't provide a text box for the enumerator to enter the "other" answer.
5. If the respondent had just moved into a new home, an "other" option should've been added when asked about the value of the last water bill.
6. When asked about how long the water was unavailable last year, the survey didn't accept the answer "I Don't Know" and wouldn't proceed to the next question.
7. When asked if he/she were the parent/legal guardian of a child who attends school and the answer was no, the survey did not proceed to the next question and asked about the child. Syrian respondents were still asked questions about decentralization and elections and questions about school even if they didn't have any children.

7.1. Translation Process:

Once the wording of the questionnaire was finalized, the translation from English to Arabic was initiated. The translation followed the steps below:

- The questionnaire was translated by a professional translator.
- The translation was reviewed by senior project staff and amendments were conducted accordingly as seen fit.
- The translation was reviewed a second time by a different senior project member by comparing the translation with the English version of the questionnaire.

In addition to producing an accurate translation, this process also ensured that key project staff are fully engaged in the questionnaire and are ready to train interviewers and answer their questions during training and research.

7.2. Scripting Process

The questionnaire was scripted on an ODK-based system that included rigorous controls to prevent and flag illogical answers. The tool was pre-tested and modified prior to scripting on the system and after scripting to ensure that all quality assurance rules were applied correctly. An export of dummy data was done prior to commencement of data collection for assurance that the data is compatible with the needed format.

This allowed for the submission of quantitative data in SPSS and STATA formats.

8. Interviewers and Training

The interviewing team consisted of 139 enumerators and 32 supervisors. (We trained a higher number than was required for the field in order to have replacements ready.) All team members underwent a structured and thorough four-day training given by our senior field coordinator, Ibtisam Al Qayyam as shown in table x below. MSI/MESP staff were also present throughout the four trainings to monitor and assist in any clarifications.

Table 4: Training Schedule

Session	Date	Attendance
Central Governorates	May 27, 2018 – May 30, 2018	47 Enumerators 11 Supervisors
Northern Governorates	June 4, 2018 – June 7, 2018	36 Enumerators 12 Supervisors
Southern Governorates	June 20, 2018 – June 24, 2018	24 Enumerators 6 Supervisors
Additional Training for newly recruited enumerators	July 1, 2018 – July 4, 2018	32 Enumerators 3 Supervisors

The training provided comprehensive background information on the study and the questionnaire for the first two days where days 3 and 4 were specifically designed to train the enumerators on the tablet and answer any questions.

Day 1 – Topics Covered:

- Training agenda and rules
- Team introductions
- Study partners introduction – Mindset and MSI/MESP
- Study objectives
- The sample
- The research methodologies
- Review of the questionnaire content

Day 2 – Topics Covered:

- Review of the questionnaire content– continued
- Ethical standards in research
- Quality assurance and quality check

Day 3 – Topics Covered:

- Tablet training

Day 4 – Topics Covered:

- Mock interviews between supervisors and enumerators
- Role playing and scenario testing

9. Field Team Composition

At the start of the project, Mindset trained 139 enumerators and 32 supervisors from all over Jordan to conduct interviews. But as data collection started, the number of enumerators and supervisors decreased to 100 enumerators and 25 supervisors as 39 and 7 withdrew or were withdrawn, respectively. Tables 5 and 6 below provides information about the regional deployment and gender of the survey enumerators and supervisors.

Table 5: Composition of Enumerators and Supervisors by Region

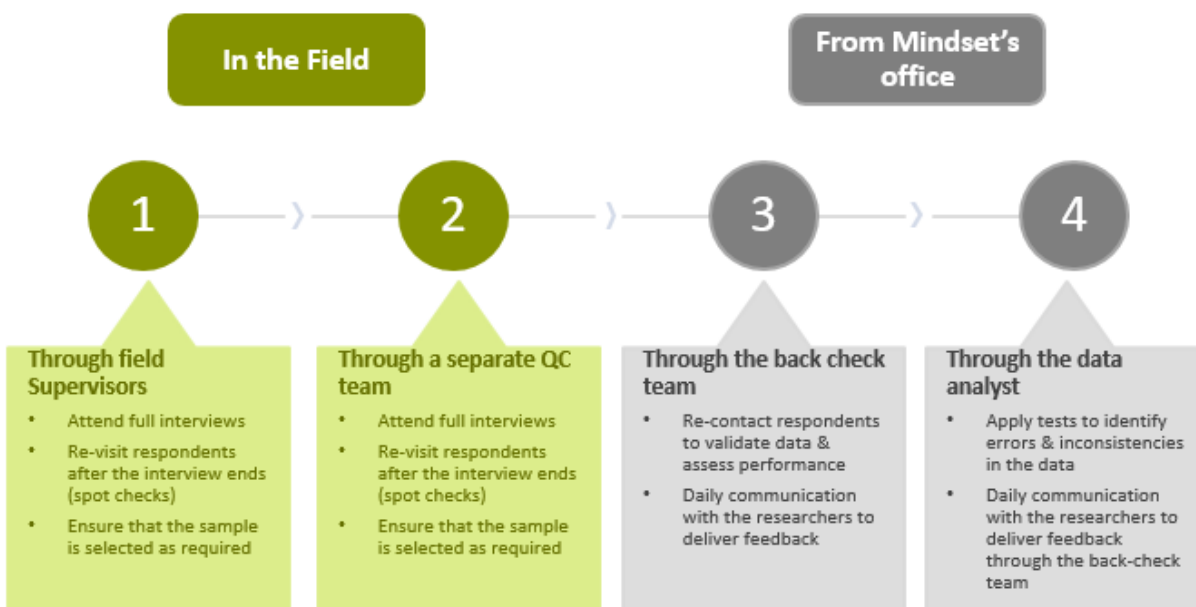
Number of/ Per Region	North	Center	South	Total
Enumerators	28	48	24	100
Supervisors	7	12	6	25

Table 6: Composition of Enumerators and Supervisors by Gender

Number of/ Per Region	Male	Female	Total
Enumerators	13	87	100
Supervisors	25	0	25

10. Quality Control Measures

Quality control processes for the survey were guided by principles of validity, reliability, timeliness, and integrity. These principles have been made operational through three steps: data collection at the enumerator level, an independent Quality Control (QC) team, and at the MESP level to ensure data quality and transparency. Some of these steps, such as the QC team, have been contracted specifically for the project while others are institutionalized processes at Mindset. Mindset hired 6 quality control supervisors to ensure best practice in the execution of this research project. Most of the interviews were subject to some form of quality control. Quality checks were implemented through the following four steps;



10.1. Through field supervisors:

- A return visit to the residence where an interview took place by the enumerator (face to face): 988 interviews, 8.6%
- During fieldwork, supervisors accompanied enumerators to ensure the proper methodology was observed (direct observation): 518 interviews, 4.3%

10.2. Through the QC team:

- A return visit to the residence where an interview took place by the enumerator (face to face): 267 interviews, 2.2%
- During fieldwork, supervisors accompanied enumerators to ensure the proper methodology was observed (direct observation): 304 interviews, 2.5%

10.3. Through the back-check team:

- Back-checks included calling back randomly selected businesses to verify key question and to monitor the performance of all enumerators. Additionally, faulty responses identified by the data processing expert were re-contacted for verification (The QC back-check form can be found in Appendix A): 2009 successful phone calls, 16.8%
- As a result of the back-checks conducted, a total of 303 interviews, or 2.5%, were omitted from the dataset due to contradicting or missing information.
- Listening to interview audios: 8418 interviews (70.3%)
- Verification by GPS data: 4364 interviews (36.5%)

10.4. Through the data analyst:

- The data processing experts performed several levels of data cleaning for cohesion, logic, and completeness of data.

The QC team additionally monitored the field supervisors and enumerators per the following points:

- The commitment of the supervisors to the right blocks
- The commitment of the supervisors and enumerators to the right buildings
- Interview duration
- The overall assessment points
- Assist supervisors with directions and location accuracy
- Attend interviews and re visit respondents after the interview ends
- Review any unvisited respondents or refusals

Table 7: Quality Control Summary

Procedure	Description	Percentage
Face to face	A return visit to the residence where an interview took place by the enumerator	10.5%
Direct Observation	During fieldwork, supervisors accompanied enumerators to ensure the proper methodology was observed	6.8%
Data cleaning	The data processing experts performed several levels of data cleaning for cohesion, logic, and completeness of data.	100%
Listening	The data process team listened to the recorded interviews for quality assurance and enumerator feedback	70.3%
Back-checks	Back-checks included calling back randomly selected businesses to verify key question and to monitor the performance of all enumerators. Additionally, faulty responses identified by the data processing expert were re-contacted for verification. The QC back-check form can be found in Appendix A.	16.8%
Deleted cases	As a result of back-checks, some interviews were deleted due to missing or contradicting information	2.5%
GPS Verification	Office verification of the GPS collected data	36.5%
General field supervision	Field supervisors were required to ensure that the specified respondents are being interviewed as per requirements.	100%

II. Coding and Data Entry

II.1. Coding

Coding of open-ended questions started on the second day of fieldwork. The data processing team was responsible for entering the codes daily and highlighting invalid answers for the call-back team. Senior project members reviewed and approved the codes. Moreover, during data cleaning, the data processing officer reviewed all the entered codes to ensure they are valid for each question

II.2. Data Cleaning and Processing

Data cleaning was done on an on-going basis from the second day of data collection.

- I. Common errors are collected by the data processing officer and relayed daily to the research team.

2. Data errors are divided into three types:
 - Logic errors. Those are referred to the call back team for collection and verification.
 - Data entry errors. Those are referred to the data cleaning team for correct entry.
 - Open ended errors. Those are spelling mistakes which are also referred to the data cleaning team for correct entry.
 - Other checks that were done:
 - Single response: contains 1 response
 - Text response: contains words only
 - Numeric values: contains numbers only
 - Exclusive answers: contains 1 response only
 - Skips: ensure skip patterns are followed
 - The option “Other” in open-ended questions: response is entered if “other” is selected and response is different from original options / codes
3. After all errors were addressed and modified in the system, a final cleaning of the full dataset was done. The cluster, block and governorate numbers were linked to the questionnaire number which operations checked manually daily and verified that the completed questionnaires were rightly linked as per their location.

12. Project Schedule

The project started on Feb 5th when Mindset received the English questionnaire from MESP. Mindset fulfilled the requested sample as per original plan on the Oct 31st, including the ongoing data cleaning.

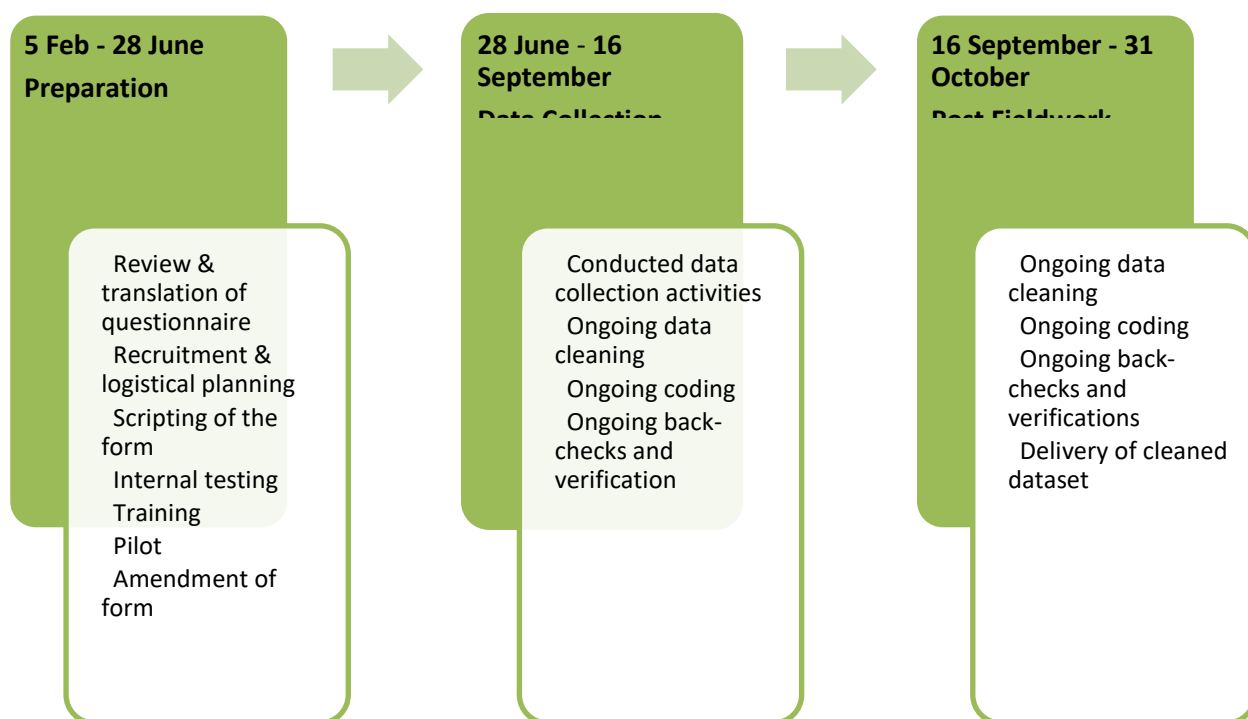


Table 8. Key Dates

Task	Date
Questionnaire review and translation	Feb 7, 2018 – Feb 9, 2018
Recruitment of enumerators and supervisors	May 7, 2018
Enumerator training	May 27, 2018 – June 24, 2018
Tablet training	May 27, 2018 – June 24, 2018
Logistical planning	June 7, 2018 – June 11, 2018
Permit approval	June 19, 2018
Pilot	June 20, 2018 – June 26, 2018
Data collection	June 28, 2018 – September 16, 2018
Data entry, cleaning, and processing	June 30, 2018 – September 30, 2018

13. Study Challenges

Mindset encountered several challenges throughout the study, listed below:

- **Non-response rate:** We faced a non-response rate of 24%.
- **Sketch Problems:** There was a delay in receiving the sketches from the Department of Statistics (DOS) since some of the sketches required by the sampling expert were unavailable at DOS. And upon checking them, once received, we had to conduct re-visits \. Additionally, the study is done on municipal level, but the information available at the Department of Statistics is at a district level.
-

- **Wrongly Labelled Buildings:** There were multiple repeated buildings between the core and replacement sample. There were also many wrongly numbered buildings which affected the sample list for the study.
- **Resistance from Local Communities:** In certain areas such as Al Shoneh and Al Sawimeh, local communities refused for the data collection team to enter the area and conduct interviews. Security authorities stopped research teams on multiple occasions, which prevented them from working areas (particularly in Fuheis considering the events of Balqa and the terrorist attack).
- **Change of Enumerators:** Several enumerators in Al Aqaba withdrew due to the difficult nature of the work. Moreover, some enumerators and supervisors in Al Tafilah governorate were asked to leave the project due to their lack of commitment to work rules and ethics.

Appendix: Post-Research Data Processing and Quality Control Procedure Checklist

Data Quality/Cleaning Checklist

Sight Checks

1. Do all SPSS variable labels and value labels in the dataset match the final questionnaire? Do they have the correct skip patterns?
2. Does the numbering of the response options in the dataset match the numbering of the options in the final questionnaire (e.g., make sure no items were reverse-coded, etc.)?
3. Does the codebook adhere to the questionnaire?

Data Cleaning

1. Does the structure (multiple/single response) of all questions in the data match the structure in the codebook?
2. Are there any missing values that should not be missing?
3. Do any of the questions have filters that were not properly followed or administered?
4. Is there any extraneous data to remove?
5. Have missing values been recoded (e.g. applying a new code to a question: e.g. refused to answer)?
6. Have open coded questions been back-coded so that “other” responses are fit into properly categorized answers whenever data filters are not affected by these changes?

Perform Logic Checks (marginal/crosstabs)

1. Were filter questions or skip patterns properly executed (cross-tabulate variables to see if respondents were isolated properly using filters/skip)?
 - If minor errors found was there forward cleaning of data? (which may include removing extraneous data of later questions that have filters that were not properly followed or administered during the research)
2. Are questions that allow for multiple responses (such as first answer/second answer; multiple dichotomies) coded properly or in a way that makes sense?
3. Are there any outliers?

Check Para/Meta Data

1. Are paradata and metadata variables specified in the technical specifications included in the data file?
2. Do sampling variables in the data file match the pre-survey sampling design?

Interviewer Checks

1. Are interviewer and supervisor workloads consistent with the contract/technical specifications for the project (e.g., number of interviews per interviewer, number of supervisors used)?
2. Is the daily distribution of interviews consistent with the contract and logically feasible for an interviewer (e.g., number of interviews per day)?
3. Do the dates and locations of the interviews match the stated dates and locations in the work plan?

4. Is the average time of interview reasonable given the questionnaire length? Can any excessively short or long interviews be explained satisfactorily?
5. Are there any overlapping interviews by the same interviewer on the same day?
6. Are there any instances of interviewer "teleportation" (e.g., interviewer moves across the country in a single day, in a way that is impossible)?
7. Are there any interviewers who had the same responses for questions across all his/her interviews?
8. Are there interviewers with high item non-responses and missing values in the data?
9. Are there any interviews/cases that have the same answers across a series of questions?

Duplicates

1. Does the dataset have any duplicate cases (e.g., duplicate IDs)?