Appendix A: Methodology

Appendix A: Methodology

The Wellcome Global Monitor is the first global survey of how people worldwide think and feel about key science and health challenges – including such critical topics as public trust in scientists, attitudes towards vaccines and how inclusive or exclusive people believe the benefits of science to be.

This appendix will focus on key methodological details related to the study, including how the questionnaire was developed, how the survey was fielded and how the data was analysed. The first section of this chapter will focus on the survey methodology of the study, including information about questionnaire development, translation, interviewer training, sampling and data collection, and data preparation. The final section will provide additional information related to the analysis of the survey data, including the use of standardised variables, external metrics, the development of the Trust in Scientists Index and the multivariate analysis exploring the drivers of trust in scientists (see Chapter 3).

Face to face interview taking place in China. *Gallup 2018*

I: Survey Methodology

Topics Explored in this Section

This section provides technical information about how the 2018 Wellcome Global Monitor survey was conducted – including questionnaire development and preparation, sampling approach and data collection methodology, as well as the data weighting process.

The Country Dataset Details for the 2018 Wellcome Global Monitor table at the end of this section provides country-level survey information, including fielding dates, sample size, margin-of-error and survey design effect.

The study was included as a module within the Gallup World Poll. Since 2005, The Gallup World Poll has regularly surveyed people in over 160 countries, representing more than 99% of the world's population aged 15+, using randomly selected, nationally representative samples.

Questionnaire development

The Wellcome Global Monitor was developed using a careful research and design process, which identified the most salient topics related to public attitudes towards science or health that could be meaningfully included on a survey fielded in over 160 countries.

The main steps of the questionnaire development process included an extensive literature review of past research, interviews with leading researchers in this field, a cognitive testing process in ten countries to make sure the questions could be understood across countries and by various demographic groups within any given country, as well as pilot tests in ten countries.

Questions were designed to be easily understood in the local languages, avoiding expressions that are difficult to translate in different languages. Where possible, response options were kept to a simple binary format such as 'yes/no' to lighten the cognitive burden on respondents and limit cultural influences on response styles that are associated with longer scales. Additionally, shorter questions and binary response options may also help reduce the impact of mode differences, or the differences in how people respond to survey questions that appear related to the way the survey was conducted (which, for the Wellcome Global Monitor, was face-to-face or by telephone – please see below).

To learn more about this process, please refer to the Wellcome Global Monitor: Questionnaire Development Report[1].

Questionnaire translation

The questionnaire was translated into the major conversational languages of each country.

First, Gallup created master language questionnaires in English, French, Spanish, Portuguese, Russian, and Arabic. Then, local language translations were performed from the master language version. For example, first, the Russian master language questionnaire was created (translation from English to Russian); and it was then translated from Russian into local languages such as Ukrainian, Kyrgyz and Uzbek.

The key component of quality assurance in translation was an independent check of every questionnaire translation. One of these two translation methods was used in each country:

- **METHOD 1:** Two independent translations are completed. An independent third party, with some knowledge of survey research methods, adjudicates the differences. A professional translator translates the final version back into the source language.
- **METHOD 2:** A translator translates into the target language, and an independent translator back-translates into the source language. An independent third party with knowledge of survey methods reviews and revises the translation as necessary.

Professional translators experienced in translating survey questionnaires were selected who have typically worked for years with Gallup's local data collection network (local translators). All translators received the same set of notes and guidance regarding the meaning of specific items.

Interviewers were instructed to follow the interview script and not to deviate from the translated language.

Interviewer training and quality control

In fielding the 2018 Wellcome Global Monitor, Gallup and its local vendors employed over 3,600 interviewers in the 144 countries. Interviewers for the Wellcome Global Monitor survey participated in standard Gallup training, which includes – among other things – the following topics:

- Research ethics, protecting respondents' confidentiality, staying safe while in the field
- Introductions: starting the interview
- Reading survey questions as on the questionnaire
- Handling questions from respondents
- Closed-end items and open-end items
- Read and rotate patterns
- Skip patterns
- Probing
- Respondent selection
- Household selection and substitution (for face-to-face countries).

During fieldwork, field supervisors and independent validation staff performed a minimum number of validations in each country.

At least 30% of completed face-to-face interviews were validated using accompanied interviews, in person re-contacts or telephone re-contacts. The supervisor/validator evaluated the interviewer's performance in implementing the survey methodology, including starting point selection, random route procedure, correct tracking sheet entry, respondent selection, and proper questionnaire administration (reading each question, not leading the respondent, etc.).

In an accompanying interview, the supervisor was present for at least 50% of the interview (for example, if the interview was 40 minutes in length, the supervisor will have been present for at least 20 minutes). During re-contacts (in person or on the telephone), the respondent was re-contacted to validate the interview.

At least 15% of completed telephone interviews were validated by either listening to interviews live or listening to recorded interviews. Validations verify that the interview was completed, that methodological standards were followed (e.g. respondent selection), and that the questionnaire was administered appropriately (reading each question, not leading the respondent, etc.).

Sampling and data collection methodology

All samples are probability-based and nationally representative of the resident adult population. The coverage area is the entire country, including rural areas, and the sampling frame represents the entire civilian, non-institutionalised, aged 15 and older population (see Face-to-Face Survey Design and Telephone Survey Design sections below). Exceptions include areas where the safety of interviewing staff is threatened, scarcely populated islands in some countries, and areas that interviewers can reach only by foot, animal or small boat (see Table II. below).

Gallup uses telephone surveys in countries where telephone coverage represents at least 80% of the population or is the customary survey methodology. In Central and Eastern Europe, much of Latin America, former Soviet states, nearly all of Asia, the Middle East and Africa, an area frame design is used for face-to-face interviewing. The data collection method used in each country is presented in Table II.

The typical Gallup World Poll survey, of which Wellcome Global Monitor is now part, typically has 1,000 surveys of individuals in each country. In China, India and Russia the sample sizes are 2,000 or greater. In rare instances, the sample size falls between 500 and 1,000 (see Table II).

Face-to-face survey design

First Stage: Stratification and Sampling

In countries where face-to-face surveys are conducted, sampling units are stratified by population size and/or geography, and clustering is achieved through one or more stages of sampling. Where population information is available, sample selection is based on probabilities proportional to population size, otherwise, simple random sampling is used. Samples are drawn independently of any samples drawn for surveys conducted in previous years. The goal is to identify 100 to 125 ultimate clusters (sampling units) consisting of clusters of households.

For face-to-face surveys, Gallup uses three different sampling approaches, depending on the available population information:

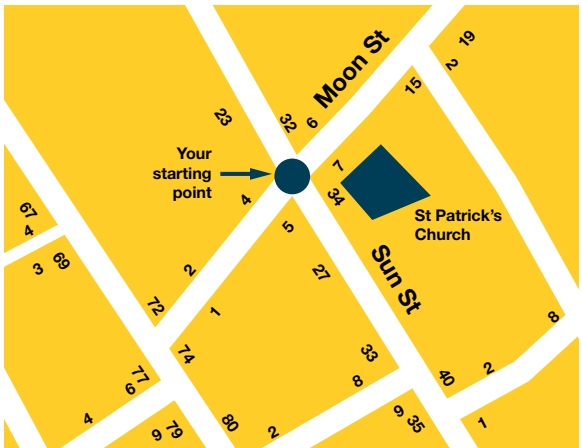
- **METHOD 1:** In countries where Gallup has detailed population information from a recent census or other reliable sources, a stratified single-stage or multiple-stage cluster design is used. Sampling units are selected using probabilities proportional to population size for each sampling stage down to 100 to 125 ultimate clusters, with a fixed number of interviews (eight or ten) completed in each ultimate cluster. If a multiple stage of selection is used, a minimum of 33 Primary Sampling Units (PSUs) are selected.

- **METHOD 2:** In countries with limited population information, Gallup uses a stratified multiple-stage cluster design. PSUs are selected using probabilities proportional to size, and units at subsequent stages are selected using simple random sampling. At least 33 PSUs are selected at the first stage of sampling, with 100 to 125 ultimate clusters selected at the last stage of sampling.
- **METHOD 3:** In countries where only overall population information is available at the strata level (broad geographies/regions), and below that, just the name of units down to the lowest administrative unit are available, Gallup uses a stratified single-stage cluster design. PSUs (for example, wards or villages) are selected using simple random sampling. The sample design results in 100 to 125 PSUs/ultimate clusters.

Second Stage: Household Selection
Random route procedures were used to select sampled households. In each ultimate cluster, the supervisor or field manager pre-selected a starting point/address for the interviewer. Once the interviewer reached the starting point, he or she followed strict rules to determine the households he or she would visit to attempt an interview.

Definition of a Household: All interviews took place at a person's home, which could be anything from a one-room flat to a single house. To be eligible, a household had to have its own cooking facilities, which could be anything from a standing stove in the kitchen to a small fire in the courtyard.

Movement from the Starting Point: Once at the given starting point, the interviewer placed his or her back to the (main) entrance of the structure and moved to the right (rule: always go to the right). Counting three households (excluding the starting point), the interviewer attempted a contact at the third household (main household). A higher interval (five or more) could be employed in dense urban areas or large apartment buildings. Unless an outright refusal occurred, interviewers could make up to three attempts to survey the household.



After visiting this first main household, the interviewer continued to select the third household to the right, and so on. If the interviewer was not successful in completing an interview at a selected household, it was replaced with another household using the same procedure.

The interviewer was instructed to count individual households and not houses, and not to count unoccupied structures. Group quarters (institutions and other group living arrangements such as rooming houses, dormitories and military barracks) were excluded from this survey.

Third Stage: Respondent Selection
The interviewer's next step was to randomly select the respondent within the household. The interviewer listed all household members aged 15 and older who lived in the household. The computer-assisted personal interviewing (CAPI) system then randomly selected the household member to be interviewed. If the country survey was collected using paper and pencil (PAPI), then the selection of the household member to interview was performed using the Kish grid.

If the selected respondent was temporarily unavailable, the interviewer would revisit the household at another time. If the selected respondent refused to take part in an interview or was unavailable for the remainder of the field period, the household was replaced with another household (following the random route procedure).

Telephone survey design

In countries where interviews were conducted by telephone, a dual sampling frame was used (landline and mobile telephone), except for Finland and Libya, which were mobile telephone only.

For each country, landline and mobile samples were generated by one of the following common approaches:

1. Using a pure regression discontinuity design (RDD) approach, where the national numbering plan is used to generate all possible combinations of telephone numbers and a stratified simple random sample is drawn, where the strata for the landline sample is based on geographic regions and for the mobile sample is based on implicit stratification of mobile service providers.
2. List-assisted RDD approach for landlines where directory listing is used to determine the active blocks of telephone numbers from which a stratified simple random sample is drawn.
3. A random sample from a registered listing.

The proportion of landline/mobile phone interviews to be completed in each country was determined based on publicly available information from reliable large-scale nationally representative surveys on landline/mobile access and usage.

For respondents contacted by landline telephone, the respondent was randomly selected within the household (among eligible respondents aged 15 and older). In all Western Europe, Northern America, and developed Asia, a random selection of the respondents was performed by asking for the person aged 15 and older who has the next birthday. For Gulf Cooperation Council (GCC) countries in the Middle East, the respondent was selected by first listing all household members age 15 and above, and the computer-assisted telephone interviewing (CATI) programme randomly selected the household member to be interviewed.

Interviewers made at least five attempts to reach a potential respondent, spread over different days and times of the day.

Data weighting

Data weighting is used to ensure samples are nationally representative for each country and is intended to be used for calculations within a country.

Firstly, Gallup constructs base-sampling weights to account for household size. Weighting by household size (number of residents aged 15 and older) is used to adjust for the probability of selection, as residents in large households will have a disproportionately lower probability of being selected for the sample.

Secondly, to ensure the sample is projectable to the target population, post-stratification weights are constructed to correct for non-response. Population statistics are used to weight the data by gender, age, and, where reliable data is available, education or socio-economic status.

Sampling error/precision of estimates

When interpreting survey results, all sample surveys are subject to various types of potential errors. Errors may occur, for example, due to non-response (where selected respondents are never reached or refuse to participate), interviewer administration error (where a response can be typed incorrectly or misinterpreted by the interviewer), or incomplete or inaccurate answers from the respondent.

The sampling design employed in this study was used to produce unbiased estimates of the stated target population. An unbiased sample will have the same characteristics and behaviours as those of the total population from which it was drawn.

In other words, with a properly drawn sample, we can make statements about the target population within a specific range of certainty. Sampling errors can be estimated, and their measures can be used to help interpret the final data results. The size of such sampling errors depends largely on the number of interviews and the complexity of the sampling design.

The margin of error (MOE), or the level of precision used in estimating the unknown population proportion 'P' can be derived based on the following formula[2]:

$$MOE = 1.96 * \sqrt{(P*(1-P)/n)}$$

where 'n' is the sample size (i.e. the number of completed surveys). Under the most conservative assumption (P = 0.5), the MOE for a sample size of 1,000 will be $1.96* \sqrt{(.25/1000)} = 3.1\%$ under the assumption of simple random sampling.

Table I.A shows the size of the 95% confidence interval half-widths for various sample sizes under the assumption of simple random sampling. They may be interpreted as indicating the approximate range (plus or minus the figure shown) around the sample estimate within which the results of repeated sampling in the same time period could be expected to fall 95% of the time, assuming the same sampling procedures, interviewing process, and questionnaire. For any given sample size, the estimated precision is lowest when P = 0.5 (or 50%). For example, the sample size needed to ensure a sampling error (or half-width of confidence interval) of 0.05 at 95% confidence level is around 400 cases when P = 0.5 (or 50%). A sample size of 300 will produce a sampling error close to 0.057 at 95% level of significance when P = 0.5 (or 50%). With P = 0.4 (or 40%), a sample size of 300 will produce a sampling error of 0.056. Table I.A shows estimated precision levels (or half-widths of confidence intervals) for different values of P and sample sizes under the assumption of simple random sampling.

While Table I.A reflects precision assuming simple random sampling, face-to-face surveys use complex designs involving stratification and clustering. Even for telephone samples, although drawn as simple random samples within each frame, the overall sample design is complex. In addition to design complexities, both modes of data collection require unequal weights to correct for household selection with an unequal probability of selection and non-response adjustments through post-stratification weighting. This introduces a design effect that needs to be taken into account while computing the sampling error (or precision) of the estimates. The design effect is defined as the ratio of the design-based sample variance to the sample variance obtained from a simple random sample of the same size. To calculate the precision of an estimate

Table I.A:
95% confidence interval half-widths for percentages for entire sample or subgroups, in percentage points

Sample sizes near	For percentages near 5/95% + -	10/90% + -	20/80% + -	30/70% + -	40/60% + -	50/50% + -
400	2.1	2.9	3.9	4.5	4.8	4.9
500	1.9	2.6	3.5	4.0	4.3	4.4
600	1.7	2.4	3.2	3.7	3.9	4.0
800	1.5	2.1	2.8	3.2	3.4	3.5
1,000	1.4	1.9	2.5	2.8	3.0	3.1
1,500	1.1	1.5	2.0	2.3	2.5	2.5
2,000	.96	1.3	1.8	2.0	2.1	2.2
2,500	.85	1.2	1.6	1.8	2.0	2.0
3,000	.78	1.1	1.4	1.6	1.8	1.8
4,000	.68	.93	1.2	1.4	1.5	1.5
5,000	.60	.88	1.2	1.3	1.3	1.4

using the complex sampling design with a design effect, one must multiply the precision under the assumption of simple random sampling by the square root of the design effect associated with this estimate.

In other words, the precision of an estimate (p) of an unknown population proportion 'P' may be approximated as:

Precision (p) = {SQRT (Deff)} × SE(p)

where 'Deff' is the design effect associated with the estimate (p)

SE(p)=SQRT{p*(1-p)/(n – 1)}

n = the unweighted sample size

For purposes of simplicity, an estimate of 'Deff_wt' is provided for each country taking into consideration only the variability of weights[3]. In addition to the variability of weights, clustered samples in face-to-face surveys also contribute to the design effect by reducing the effective sample size. The intraclass correlation coefficient for each estimate and the average cluster size impacts the design effect as follows:

Deff_c = (1 + (c-1)*ρ)

Where 'Deff_c' is the design effect due to clustering, 'c' is the average cluster size and 'ρ' is the intraclass correlation coefficient for a particular estimate.
For purposes of illustration, given an average cluster size of 10 and an intraclass correlation coefficient estimate of 0.1, the design effect due to clustering is:

Deff_c = (1 + (10-1)*0.1) = 1.9

Therefore, precision for estimates generated from face-to-face surveys can be approximated by this formula.

MOE = 1.96 *√(P*(1-P)/n) *√(Deff_wt) *√(Deff_c)

Table II.B:
Country dataset details, 2018 Wellcome Global Monitor

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Afghanistan	1–29 Jul, 2018	1,000	1.50	±3.8	Face-to-Face	Dari, Pashto	Gender-matched sampling was used during the final stage of selection.
2018	Albania	23 May–21 Jun, 2018	1,000	1.29	±3.5	Face-to-Face CAPI	Albanian	
2018	Algeria	29 Sep–16 Oct, 2018	1,000	1.57	±3.9	Face-to-Face CAPI	Arabic	Sparsely populated areas in the far south were excluded, representing approximately 10% of the population.
2018	Argentina	19 May–25 Jun, 2018	1,000	1.45	±3.7	Face-to-Face CAPI	Spanish	Those living in dispersed rural population areas were excluded. This represents about 5.7% of the population.
2018	Armenia	1 Aug–29 Sep, 2018	1,000	1.48	±3.8	Face-to-Face CAPI	Armenian	
2018	Australia	13 Aug–21 Oct, 2018	1,003	1.81	±4.2	Landline and Mobile Telephone	English	
2018	Austria	21 Sep–19 Oct, 2018	1,000	1.32	±3.6	Landline and Mobile Telephone	German	
2018	Azerbaijan	24 Aug–13 Sep, 2018	1,000	1.31	±3.5	Face-to-Face	Azeri, Russian	Kelbadjaro-Lacha, Nakhichevan and Nagorno-Karabakh territories not included. These areas represent approximately 14% of the total population.
2018	Bangladesh	30 Apr–13 Aug, 2018	1,000	1.31	±3.5	Face-to-Face CAPI	Bengali	Three hill districts in Chittagong (Rangamati, Khagrachari and Bandarban) and one district in Rangpur (Gaibandha) were excluded for security reasons. The excluded area represents approximately 3% of the population.
2018	Belarus	27 Jun–17 Jul, 2018	1,061	1.42	±3.6	Face-to-Face CAPI	Russian	
2018	Belgium	25 May–22 Jun, 2018	1,004	1.28	±3.5	Landline and Mobile Telephone	French, Dutch	
2018	Benin	13 Jul–6 Aug, 2018	1,000	1.57	±3.9	Face-to-Face CAPI	Bariba, Fon, French, Anago	
2018	Bolivia	17 Aug–10 Oct, 2018	1,000	1.45	±3.7	Face-to-Face CAPI	Spanish	Very remote areas that lack regular public transport were excluded due to the difficulty of access. The exclusions represent approximately 11% of the population.
2018	Bosnia and Herzegovina	12 Jul–9 Sep, 2018	1,000	1.33	±3.6	Face-to-Face CAPI	Bosnian, Croatian, Serbian	
2018	Botswana	10 Nov–1 Dec, 2018	1,002	1.60	±3.9	Face-to-Face CAPI	English, Setswana	
2018	Brazil	19 Jul–22 Aug, 2018	1,000	1.33	±3.6	Face-to-Face CAPI	Portuguese	People living in indigenous lands, closed residential areas and dangerous areas where the safety of interviewers was threatened were excluded during fieldwork. The exclusions represent approximately 1% of the population.
2018	Bulgaria	10 Oct–12 Dec, 2018	1,001	1.44	±3.7	Face-to-Face CAPI	Bulgarian	

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Burkina Faso	25 Jun–6 Jul, 2018	1,000	1.67	±4.0	Face-to-Face CAPI	Dioula, French, Fulfulde, Mòoré	
2018	Burundi	13–23 Sep, 2018	1,000	1.34	±3.6	Face-to-Face CAPI	French, Kirundi	
2018	Cambodia	22 May–16 Jun, 2018	1,000	1.52	±3.8	Face-to-Face CAPI	Khmer	Koh Kong, Steung Treng, and Oddar Meanchey provinces were excluded. These excluded areas represent approximately 3% of the population of Cambodia.
2018	Cameroon	25 Apr–13 May, 2018	1,000	1.65	±4.0	Face-to-Face CAPI	French, English, Fulfulde	The following Arrondissements were excluded due to security concerns from Boko Haram attacks: Goulfey, Blangoua, Fotokol, Zina, Darak, Hile-Alifa, Wasa, Bourrha, Mogode, Koza, Mayo Moskota, Mora, Kolofata, and Tokombéré. Department of Manyu was also excluded due to insecurity. Neighbourhoods with fewer than 50 households were also excluded from the sampling. During fieldwork, the security situation degraded in the Northwest and Southwest regions as armed groups attacked government troops and civilians, and some localities were excluded as a result. In total, the excluded areas represent approximately 16% of the total population.
2018	Canada	5 Jul–13 Aug, 2018	1,012	1.58	±3.9	Landline and Mobile Telephone	English, French	
2018	Chad	5–20 Oct, 2018	1,000	1.69	±4.0	Face-to-Face CAPI	French, Chadian Arabic, Ngambay	Because of security and wilderness, seven regions were excluded from the sampling: Ouaddai, Wadi Fira, Bourkou, Ennedi, Tibesti, Salamat and Sila. Quartiers/villages with fewer than 50 inhabitants were also excluded from sampling. The total population excluded is 20%.
2018	Chile	20 Jul–11 Sep, 2018	1,000	1.39	±3.7	Face-to-Face CAPI	Spanish	A few remote and sparsely populated municipalities were excluded due to difficulties of access. The excluded areas represent less than 1% of the population.
2018	China	6 Jul–19 Aug, 2018	3,649	1.41	±1.9	Face-to-Face CAPI	Chinese	Xinjiang and Tibet were excluded from the sample. The excluded areas represent less than 5% of the population of China.
2018	Colombia	2–20 Jun, 2018	1,000	1.38	±3.6	Face-to-Face CAPI	Spanish	
2018	Comoros	25 Sep–28 Oct, 2018	1,000	1.51	±3.8	Face-to-Face CAPI	French, Comorian	
2018	Congo Brazzaville	21 May–10 Jun, 2018	1,000	1.62	±3.9	Face-to-Face CAPI	French, Kituba, Lingala	
2018	Costa Rica	16 May–30 Aug, 2018	1,000	1.49	±3.8	Face-to-Face CAPI	Spanish	
2018	Croatia	12 Jun–16 Aug, 2018	1,000	1.49	±3.8	Face-to-Face CAPI	Croatian	
2018	Cyprus	23 Apr–16 Jun, 2018	1,011	1.44	±3.7	Landline and Mobile Telephone	Greek, English	
2018	Czech Republic	10 Jul–7 Nov, 2018	1,000	1.21	±3.4	Face-to-Face CAPI	Czech	

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Denmark	3–30 May, 2018	1,000	1.20	±3.4	Landline and Mobile Telephone	Danish	
2018	Dominican Republic	22 Sep–12 Oct, 2018	1,000	1.36	±3.6	Face-to-Face CAPI	Spanish	
2018	Ecuador	2–26 Aug, 2018	1,000	1.45	±3.7	Face-to-Face CAPI	Spanish	
2018	Egypt	4–17 Aug, 2018	1,000	1.36	±3.6	Face-to-Face CAPI	Arabic	Frontier governorates (Matruh, Red Sea, New Valley, North Sinai, and South Sinai) were excluded, as they are remote and represent a small proportion of the population of the country. The excluded areas represent less than 2% of the total population.
2018	El Salvador	18 Jun–10 Aug, 2018	1,000	1.64	±4.0	Face-to-Face CAPI	Spanish	
2018	Estonia	21 Jun–30 Jul, 2018	1,000	1.21	±3.4	Face-to-Face CAPI	Estonian, Russian	
2018	Ethiopia	11 Jun–5 Jul, 2018	1,000	1.44	±3.7	Face-to-Face CAPI	Amharic, English, Oromo, Tigrinya	Six of the nine zones of the Somali region (Degehabur, Werder, Korahe, Fik, Gode, Afder) were excluded due to accessibility, security issues, and nomadism. Additionally, in the Somali regions, Liben Zone, Moyale and Dolo Ado were excluded because of security concerns. All the woredas in Benishangul region, Kamashi Zone were also excluded for security reasons. The exclusions represent 4% of the population of Ethiopia.
2018	Finland	3 May–1 Jun, 2018	1,000	1.53	±3.8	Mobile Telephone	Finnish, Swedish	
2018	France	4–31 Oct, 2018	1,000	1.53	±3.8	Landline and Mobile Telephone	French	
2018	Gabon	27 Jul–24 Aug, 2018	1,000	1.51	±3.8	Face-to-Face CAPI	French, Fang, Teke	
2018	Gambia	21 Jun–14 Jul, 2018	1,000	1.42	±3.7	Face-to-Face CAPI	English, Pulaar, Wolof, Malinke	
2018	Georgia	30 Jul–7 Sep, 2018	1,000	1.36	±3.6	Face-to-Face CAPI	Georgian, Russian	South Ossetia and Abkhazia were not included for the safety of the interviewers. The excluded area represents approximately 7% of the population.
2018	Germany	21 Sep–19 Oct, 2018	1,000	1.41	±3.7	Landline and Mobile Telephone	German	
2018	Ghana	23 Oct–30 Dec, 2018	1,000	1.54	±3.8	Face-to-Face CAPI	English, Ewe, Twi, Dagbani	
2018	Greece	18 Jun–17 Jul, 2018	1,000	1.30	±3.5	Face-to-Face CAPI	Greek	
2018	Guatemala	3 Jul–1 Aug, 2018	1,000	1.30	±3.5	Face-to-Face CAPI	Spanish	
2018	Guinea	23 Aug–11 Sep, 2018	1,000	1.51	±3.8	Face-to-Face CAPI	French, Malinke, Pular, Soussou	
2018	Haiti	22 Aug–5 Nov, 2018	500	1.49	±5.3	Face-to-Face CAPI	Creole	
2018	Honduras	19 Aug–27 Sep, 2018	1,000	1.31	±3.6	Face-to-Face CAPI	Spanish	
2018	Hungary	12 Sep–19 Dec, 2018	1,000	1.36	±3.6	Face-to-Face CAPI	Hungarian	

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Iceland	9 Apr–8 May, 2018	500	1.38	±5.2	Landline and Mobile Telephone	Icelandic	
2018	India	10 Oct–21 Dec, 2018	3,000	1.46	±2.2	Face-to-Face CAPI	Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Marathi, Odia, Punjabi, Tamil, Telugu	Excluded population living in Northeast regions and remote islands. The excluded areas represent less than 10% of the population.
2018	Indonesia	1–27 Jul, 2018	1,000	1.38	±3.6	Face-to-Face CAPI	Bahasa Indonesia	
2018	Iran	9–22 May, 2018	1,005	1.63	±4.0	Landline and Mobile Telephone	Farsi	
2018	Iraq	3 Nov, 2018–6 Jan, 2019	1,000	1.24	±3.5	Face-to-Face CAPI and PAPI	Arabic, Kurdish	
2018	Ireland	23 Apr–22 May, 2018	1,000	1.35	±3.6	Landline and Mobile Telephone	English	
2018	Israel	4 Jul–18 Aug, 2018	1,010	1.26	±3.5	Face-to-Face	Hebrew, Russian, Arabic	The sample does not include the area of East Jerusalem. This area is included in the sample of Palestinian Territories.
2018	Italy	14 Sep–6 Oct, 2018	1,000	1.57	±3.9	Landline and Mobile Telephone	Italian	
2018	Ivory Coast	26 Sep–10 Oct, 2018	1,000	1.69	±4.0	Face-to-Face CAPI	French, Dioula	
2018	Japan	20 Jun–5 Sep, 2018	1,004	1.52	±3.8	Landline and Mobile Telephone	Japanese	Landline RDD excluded 12 municipalities near the nuclear power plant in Fukushima. These areas were designated as not-to-call districts due to the devastation from the 2011 disasters. The exclusion represents less than 1% of the population of Japan.
2018	Jordan	11 Aug–12 Sep, 2018	1,002	1.30	±3.5	Face-to-Face CAPI	Arabic	
2018	Kazakhstan	30 Jun–16 Aug, 2018	1,000	1.43	±3.7	Face-to-Face CAPI	Russian, Kazakh	
2018	Kenya	23 May–9 Jun, 2018	1,000	1.60	±3.9	Face-to-Face CAPI	English, Swahili	Mandera County, Wajir County, Marsabit County, Baringo County, and Garissa County (except some areas in Garissa and Lagdera districts) were excluded due to accessibility and/or security issues. The exclusions represent 8% of the population.
2018	Kosovo	19 Jun–23 Jul, 2018	1,000	1.24	±3.4	Face-to-Face CAPI	Albanian, Serbian	
2018	Kuwait	3 Nov–31 Dec, 2018	1,001	1.34	±3.6	Landline and Mobile Telephone	Arabic, English	Includes only Kuwaitis, Arab expatriates and non-Arabs who were able to complete the interview in Arabic or English.
2018	Kyrgyzstan	14 Jun–28 Jul, 2018	1,000	1.51	±3.8	Face-to-Face CAPI	Kyrgyz, Russian, Uzbek	
2018	Laos	2 Oct–24 Nov, 2018	1,001	1.25	±3.6	Face-to-Face CAPI	Lao	Excluded Xaisomboun Province and some communes that are unreachable and/or have security considerations. The excluded areas represent approximately 10% of the population.

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Latvia	17 Aug–18 Nov, 2018	1,021	1.25	±3.4	Face-to-Face CAPI	Latvian, Russian	
2018	Lebanon	29 Jun–25 Jul, 2018	1,000	1.30	±3.5	Face-to-Face CAPI	Arabic	Hermel, Baalbek, and Bint Jbeil under the strict control of Hezbollah were excluded. The excluded areas represent approximately 10% of the population.
2018	Liberia	23 May–10 Jun, 2018	1,000	1.38	±3.6	Face-to-Face CAPI	English, Pidgin English	
2018	Libya	23 April–11 May, 2018	1,003	1.63	±3.9	Mobile Telephone	Arabic	
2018	Lithuania	13 Sep–31 Oct, 2018	1,000	1.38	±3.6	Face-to-Face CAPI	Lithuanian	
2018	Luxembourg	4 Oct–1 Nov, 2018	1,000	1.47	±3.8	Landline and Mobile Telephone	French, German	
2018	Macedonia	25 Jun–14 Aug, 2018	1,008	1.46	±3.7	Face-to-Face CAPI	Macedonian, Albanian	
2018	Madagascar	15 May–9 Jun, 2018	1,000	1.50	±3.8	Face-to-Face CAPI	French, Malagasy	Regions that were unsafe or unreachable were excluded from the sample. The excluded areas represent approximately 20% of the total population.
2018	Malawi	10–18 Aug, 2018	1,000	1.39	±3.7	Face-to-Face CAPI	Chichewa, English, Tumbuka	
2018	Malaysia	12 Jul–15 Sep, 2018	1,000	1.40	±3.7	Face-to-Face CAPI	Bahasa Malay, Chinese, English	
2018	Mali	9–18 May, 2018	1,000	1.56	±3.9	Face-to-Face CAPI	French, Bambara	The regions of Gao, Kidal, Mopti and Timbuktu were excluded because of insecurities. Quartiers and villages with fewer than 50 inhabitants were also excluded from the sample. The excluded areas represent 23% of the total population.
2018	Malta	15 Apr–16 May, 2018	1,004	1.55	±3.8	Landline and Mobile Telephone	Maltese, English	
2018	Mauritania	18 Sep–1 Oct, 2018	1,000	1.66	±4.0	Face-to-Face CAPI	French, Poulaar, Wolof, Hassaniya, Soninké	
2018	Mauritius	10 Apr–14 May, 2018	1,000	1.56	±3.9	Landline and Mobile Telephone	Creole, English, French	
2018	Mexico	5 Oct–29 Nov, 2018	1,000	1.39	±3.7	Face-to-Face CAPI	Spanish	
2018	Moldova	10 Jul–24 Aug, 2018	1,000	1.19	±3.4	Face-to-Face CAPI	Romanian/ Moldavian, Russian	Transnistria (Pridnestrovian Moldavian Republic) excluded for the safety of interviewers. The excluded area represents approximately 13% of the population.
2018	Mongolia	2–24 Jun, 2018	1,000	1.20	±3.4	Face-to-Face CAPI	Mongolian	
2018	Montenegro	3 Jun–6 Sep, 2018	1,000	1.40	±3.7	Face-to-Face CAPI	Montenegrin, Serbian	
2018	Morocco	14 Jul–3 Aug, 2018	1,001	1.34	±3.6	Face-to-Face CAPI	Moroccan Arabic	Excludes the Southern Provinces. The excluded area represents approximately 3% of the population.
2018	Mozambique	3 Aug–17 Oct, 2018	1,000	1.56	±3.9	Face-to-Face CAPI	Portuguese, Xichangana, Emakhuwa	

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Myanmar	11 Jun–16 Jul, 2018	1,000	1.32	±3.6	Face-to-Face CAPI	Burmese	Chin, Kayah, and Kachin states were excluded. The excluded areas represent less than 5% of the population.
2018	Namibia	31 Jul–24 Sep, 2018	1,005	1.60	±3.9	Face-to-Face CAPI	English, Oshivambo, Afrikaans, Kwangali	
2018	Nepal	22 Jun–23 Jul, 2018	1,000	1.32	±3.6	Face-to-Face CAPI	Nepali	
2018	Netherlands	6 Apr–4 May, 2018	1,001	1.40	±3.7	Landline and Mobile Telephone	Dutch	
2018	New Zealand	30 Jul–9 Sep, 2018	1,002	1.73	±4.1	Landline and Mobile Telephone	English	
2018	Nicaragua	21 May–22 Jun, 2018	1,000	1.34	±3.6	Face-to-Face CAPI	Spanish	
2018	Niger	29 Jun–23 Jul, 2018	1,000	1.54	±3.8	Face-to-Face CAPI	French, Hausa, Zarma	
2018	Nigeria	14 May–2 Jun, 2018	1,000	1.66	±4.0	Face-to-Face CAPI	English, Hausa, Igbo, Pidgin English, Yoruba	The states of Adamawa, Borno and Yobe were under a state of emergency due to Boko Haram activity and were excluded for safety and security reasons. These states represent 7% of the population.
2018	Northern Cyprus	8 May–2 Jul, 2018	1,000	1.61	±3.9	Landline and Mobile Telephone	Turkish	
2018	Norway	23 May–26 Jun, 2018	1,000	1.50	±3.8	Landline and Mobile Telephone	Norwegian	
2018	Pakistan	15 Jul–17 Aug, 2018	1,000	1.42	±3.7	Face-to-Face CAPI	Urdu	Did not include Gilgit-Baltistan. The excluded area represents less than 1% of the population. Gender-matched sampling was used during the final stage of selection.
2018	Palestinian Territories	28 Jul–14 Aug, 2018	1,000	1.43	±3.7	Face-to-Face CAPI	Arabic	Areas with security concerns close to the Israeli borders, areas that are accessible only to special Israeli permit holders, and areas with population concentrations of fewer than 1,000 people were excluded. The excluded areas represent less than 2% of the population. The sample includes East Jerusalem.
2018	Panama	21 Jun–1 Nov, 2018	1,000	1.38	±3.6	Face-to-Face CAPI	Spanish	
2018	Paraguay	5 Oct–30 Nov, 2018	1,000	1.34	±3.6	Face-to-Face CAPI	Spanish, Jopara	
2018	Peru	18 May–15 Jun, 2018	1,000	1.44	±3.7	Face-to-Face CAPI	Spanish	
2018	Philippines	14 Sep–1 Oct, 2018	1,000	1.41	±3.7	Face-to-Face CAPI	Filipino, Iluko, Hiligaynon, Cebuano, Maranao, Waray, Sorsoganon	Some areas were excluded from the sampling frame due to security concerns (such as barangays considered as war zones in Marawi) and areas that are remote or inaccessible. The excluded population from these areas represent less than 1% of the population.
2018	Poland	2 Oct–28 Dec, 2018	1,000	1.25	±3.5	Face-to-Face CAPI	Polish	
2018	Portugal	24 Apr–4 Jun, 2018	1,001	1.64	±4.0	Landline and Mobile Telephone	Portuguese	

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Romania	19 Jun–23 Aug, 2018	1,002	1.40	±3.7	Face-to-Face CAPI	Romanian, Hungarian	
2018	Russia	24 Jun–4 Oct, 2018	2,000	1.40	±2.6	Face-to-Face CAPI	Russian	People living in very remote or difficult to access areas were excluded. The excluded areas represent approximately 5% of the population.
2018	Rwanda	20–30 Jul, 2018	1,000	1.32	±3.6	Face-to-Face CAPI	English, Kinyarwanda	
2018	Saudi Arabia	2–15 May, 2018	1,016	1.54	±3.8	Landline and Mobile Telephone	Arabic, English	Includes Saudis, Arab expatriates, and non-Arabs who were able to complete the interview in Arabic or English.
2018	Senegal	26 Jun–6 Jul, 2018	1,000	1.36	±3.6	Face-to-Face CAPI	French, Wolof	
2018	Serbia	6 Jun–16 Aug, 2018	1,000	1.41	±3.7	Face-to-Face CAPI	Serbian	
2018	Sierra Leone	6–31 Jul, 2018	1,000	1.57	±3.9	Face-to-Face CAPI	English, Krio, Mende	
2018	Singapore	29 Jun–17 Sep, 2018	1,000	1.44	±3.7	Face-to-Face CAPI	English, Chinese	Some condominiums were excluded due to restricted access. This exclusion represents no more than 12% of the population.
2018	Slovakia	23 Aug–18 Oct, 2018	1,000	1.32	±3.6	Face-to-Face CAPI	Hungarian, Slovak	
2018	Slovenia	12 Apr–20 May, 2018	1,000	1.45	±3.7	Landline and Mobile Telephone	Slovene	
2018	South Africa	27 Jul–15 Aug, 2018	1,000	1.38	±3.6	Face-to-Face CAPI	Afrikaans, English, Sotho, Xhosa, Zulu	
2018	South Korea	15 Jun–14 Sep, 2018	1,014	1.31	±3.5	Landline and Mobile Telephone	Korean	
2018	Spain	14 Sep–6 Oct, 2018	1,000	1.48	±3.8	Landline and Mobile Telephone	Spanish	
2018	Sri Lanka	21 Jul–13 Sep, 2018	1,109	1.44	±3.5	Face-to-Face CAPI	Sinhala, Tamil	
2018	Swaziland	10 Aug–16 Sep, 2018	1,000	1.38	±3.6	Face-to-Face CAPI	Siswati, English	
2018	Sweden	3–31 May, 2018	1,001	1.50	±3.8	Landline and Mobile Telephone	Swedish	
2018	Switzerland	4 Oct–7 Nov, 2018	1,000	1.54	±3.8	Landline and Mobile Telephone	German, French, Italian	
2018	Taiwan	4 Jun–6 Jul, 2018	1,000	1.45	±3.7	Landline and Mobile Telephone	Chinese	
2018	Tajikistan	29 Oct–29 Dec, 2018	1,000	1.45	±3.7	Face-to-Face CAPI	Tajik	
2018	Tanzania	22 Jun–17 Jul, 2018	1,000	1.53	±3.8	Face-to-Face CAPI	English, Swahili	
2018	Thailand	24 May–19 Aug, 2018	1,000	1.48	±3.8	Face-to-Face CAPI	Thai	Three provinces in the South region (Pattani, Narathiwat, and Yala) were excluded for security reasons; in addition, a few districts in other provinces were excluded. The excluded areas in total represent less than 4% of the population.
2018	Togo	15–26 Jul, 2018	1,000	1.75	±4.1	Face-to-Face CAPI	French, Ewe, Kabiye	
2018	Tunisia	28 Apr–15 May, 2018	1,001	1.26	±3.5	Face-to-Face CAPI	Arabic	

Year	Country	Data collection dates	Number of interviews	Design effect[a]	Margin of error[b]	Mode of interviewing	Interviewing languages	Exclusions (samples are nationally representative unless noted otherwise)
2018	Turkey	23 Jul–7 Sep, 2018	1,000	1.43	±3.7	Face-to-Face CAPI	Turkish	
2018	Turkmenistan	18 Jun–20 Jul, 2018	1,000	1.30	±3.5	Face-to-Face CAPI	Turkmen, Russian	
2018	Uganda	13–22 Sep, 2018	1,000	1.44	±3.7	Face-to-Face CAPI	Ateso, English, Luganda, Runyankole	Three districts in the Northern region were excluded for security reasons: Kotido, Moroto and Nakapiripirit. The excluded areas represent approximately 4% of the population.
2018	Ukraine	11 Jul–18 Aug, 2018	1,000	1.50	±3.8	Face-to-Face CAPI	Russian, Ukrainian	Due to the situation in the east of Ukraine, occupied and conflict areas in Donetsk and Luhansk oblasts were excluded. The excluded areas represent approximately 9% of the population.
2018	United Arab Emirates	24 April–24 May, 2018	1,005	1.26	±3.5	Landline and Mobile Telephone	Arabic, English	Includes only Emiratis, Arab expatriates and non-Arabs who were able to complete the interview in Arabic or English.
2018	United Kingdom	23 April–22 May, 2018	1,000	1.37	±3.63	Landline and Mobile Telephone	English	
2018	United States	12 Jul–23 Aug, 2018	1,006	1.63	±4.0	Landline and Mobile Telephone	English, Spanish	
2018	Uruguay	3 Oct–13 Dec, 2018	1,000	1.34	±3.6	Face-to-Face CAPI	Spanish	
2018	Uzbekistan	23 Jun–16 Aug, 2018	1,000	1.55	±3.9	Face-to-Face CAPI	Uzbek, Russian	
2018	Venezuela	27 Sep–28 Nov, 2018	1,000	1.37	±3.6	Face-to-Face CAPI	Spanish	The Federal Dependencies were excluded due to remoteness and difficulty of access. Exclusions represent less than 1% of the population.
2018	Vietnam	22 Jun–19 Aug, 2018	1,012	1.31	±3.5	Face-to-Face CAPI	Vietnamese	Eleven provinces were excluded: An Giang, Đắk Lắk, Điện Biên, Gia Lai, Ha Giang, Ha Tinh, Kien Giang, Kon Tum, Nghe An, Quang Binh, Thanh Hóa. The excluded areas represent approximately 19% of the population.
2018	Yemen	11 Sep–10 Nov, 2018	1,000	1.54	±3.8	Face-to-Face CAPI	Arabic	Gender-matched sampling was used during the final stage of selection. Al-Mahrah, Saada Governorate and the island of Socotra were excluded due to their small sizes and remoteness. These excluded areas represent less than 4% of the population. In addition, due to the ongoing security situation, half the PSUs were replaced with a similar PSU in the same province.
2018	Zambia	20 Aug–13 Sep, 2018	1,000	1.52	±3.8	Face-to-Face CAPI	Bemba, English, Lozi, Nyanja, Tonga	
2018	Zimbabwe	19 May–22 Sep, 2018	1,000	1.43	±3.7	Face-to-Face CAPI	English, Shona, Ndebele	

a. The design effect calculation reflects the weights and does not incorporate the intraclass correlation coefficients. Design effect calculation: n(sum of squared weights)/[(sum of weights)(sum of weights)].

b. The margin of error is calculated around a proportion at the 95% confidence level. The maximum margin of error was calculated assuming a reported percentage of 50% and accounts for the design effect. The margin of error calculation: $\sqrt{(0.25/N)}1.96\sqrt{(DE)}$.

II: Data Analysis Methodology

Topics explored in this section

this section provides further information related to the data analysis presented in this report, including how different country groupings were defined; how key personal information such as education, employment and income were standardised across countries; how the Wellcome Trust in Scientists Index was developed and how the multivariate analysis into this variable was conducted.

The analysis in this report sought to answer – or at least begin to answer – the key research questions that motivated this study. In some instances, this simply entailed reporting on the topline results for each country in the study; however, very often, more complex data techniques were required to better understand why and how attitudes to science and health differed across the world, or parts of the world, or within a certain population. This section will explore the analytical tools and techniques that were employed in this analysis.

Country groupings used in the analysis

The main classifications of countries used in this report include:

Geographic Region: The analysis is classified as belonging to one of eighteen different geographic regions, largely corresponding to the continental 'sub-region' or 'intermediary' regions used by the United Nations Statistics Division (UNSD)[4]. Note the region of the 'Middle East' is not used by the UNSD; instead, the intermediary region of 'Western Asia' contains the countries defined in this report as the 'Middle East' (see Error! Reference source not found. below), along with Armenia, Azerbaijan, Cyprus and Georgia. To make the regions more accessible to the common reader, these regional definitions were altered slightly. Kosovo and Northern Cyprus were not included in any regional definition as the United Nations does not group them and there was no clear regional grouping to place them in. However, Taiwan, which is also not grouped by the UN, was placed with East Asia.

Country-income level

Countries were divided into four groupings: low, lower-middle, upper-middle and high-income economies, as defined by the World Bank[5]. According to the organisation, the groups are defined as follows[6]:

- **Low-Income:** Gross national income (GNI) per capita of $995 or less (in 2017)
- **Lower-Middle-Income:** GNI per capita of $996 and $3,895
- **Upper-Middle-Income:** GNI per capita of $3,896 and $12,055
- **High-Income:** GNI per capita above $12,055.

Overall, 27 countries included in the study were classified as low-income economies, 34 as lower-middle-income economies, 38 as upper-middle-income economies and 45 as high-income economies. One country included in the study, Northern Cyprus, is not included in the World Bank definition and was not included in these categories.

Standardisation of income, education and employment groups

Key personal information such as income, education and employment can be defined and/or measured differently in countries, which can create challenges when attempting to compare cross-country results[7]. For this reason, this report of the Wellcome Global Monitor examines these characteristics using the standardised definitions of income and education as developed by the Gallup World Poll; additionally, employment status is defined in a manner that is consistent with those of the Bureau of Labor Statistics in the United States[8].

Box II.A:
Regional groupings used in this report

In analysing the results from 144 countries in the Wellcome Global Monitor, this report makes use of regional groupings used by the United Nations Statistics Division. Countries are grouped into 18 categories as follows:

- **North Africa**: Algeria, Egypt, Libya, Morocco, Tunisia
- **Eastern Africa**: Burundi, Comoros, Ethiopia, Kenya, Madagascar, Malawi, Mauritius, Mozambique, Rwanda, Tanzania, Uganda, Zambia, Zimbabwe
- **Central Africa**: Cameroon, Chad, Congo Brazzaville, Gabon
- **Southern Africa**: Botswana, Namibia, South Africa, Swaziland
- **Western Africa**: Benin, Burkina Faso, Ghana, Guinea, Ivory Coast, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, The Gambia, Togo
- **Central America and Mexico**: Costa Rica, Dominican Republic, El Salvador, Guatemala, Haiti, Honduras, Mexico, Nicaragua, Panama
- **South America**: Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Paraguay, Peru, Uruguay, Venezuela
- **Northern America**: Canada, United States
- **Central Asia**: Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan
- **East Asia**: China, Japan, Mongolia, South Korea, Taiwan
- **Southeast Asia**: Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, Vietnam
- **South Asia**: Afghanistan, Bangladesh, India, Iran, Nepal, Pakistan, Sri Lanka
- **Middle East**: Iraq, Israel, Jordan, Kuwait, Lebanon, Palestine, Saudi Arabia, Turkey, United Arab Emirates, Yemen
- **Eastern Europe**: Belarus, Bulgaria, Czech Republic, Hungary, Moldova, Poland, Romania, Russia, Slovakia, Ukraine
- **Northern Europe**: Denmark, Estonia, Finland, Iceland, Ireland, Latvia, Lithuania, Norway, Sweden, United Kingdom
- **Southern Europe**: Albania, Bosnia Herzegovina, Croatia, Cyprus, Greece, Italy, Malta, Macedonia, Montenegro, Portugal, Serbia, Slovenia, Spain
- **Western Europe**: Austria, Belgium, France, Germany, Luxembourg, Netherlands, Switzerland
- **Australia and New Zealand**: Australia, New Zealand

Education

Countries have unique ways of classifying education levels, and these classifications need to be preserved during data collection for weighting purposes. However, to make comparisons across countries by educational attainment, consistent categories also needed to be created. All education descriptions can be placed within three categories: primary, secondary and tertiary. All responses regarding education are coded into their relevant category for global comparison.

- **Primary:** Functional equivalent to completing primary education or lower secondary or less, the level that is closest to completing up to eight years of education. The exact definition will vary by country.
- **Secondary:** Functional equivalent to completing some secondary up to some tertiary education. This typically refers to individuals who have completed between nine and fifteen years of education but have not yet completed the equivalent of a bachelor's degree. The exact definition will vary by country.
- **Tertiary:** Functional equivalent to completing four years of post-secondary tertiary education, or the equivalent of a bachelor's degree. This typically refers to individuals who have completed approximately sixteen or more years of education. The exact definition will vary by country.

Income

To provide household income measurements, Gallup asked respondents two questions. The first question asked respondents about their monthly household income in local currency before taxes. Respondents were instructed to include all income from all wages and salaries in the household, remittances from family members living elsewhere, and all other sources. If the respondents hesitated to answer or had difficulty answering the first question, they were presented with a set of income ranges in their local currency and were asked which group they fell into.

- What is your total MONTHLY household income in (country), before taxes? Please include income from wages and salaries, remittances from family members living elsewhere, farming, and all other sources.
- (If don't know or refused, ask:) Would you say your total MONTHLY household income is ____?

Estimates for respondents answering the second income question were imputed using hot-deck imputation, but restricting imputing values to the reported range. Estimates for respondents who did not answer either income question were imputed using the same method, with no restriction of range. In this imputation process, each missing value is replaced with an observed value from another unit that has characteristics similar to the missing unit.

The hot-deck imputation procedure matched respondents with answers and without answers (called 'donors' and 'beggars' respectively) by a set of external independent variables that are expected to be related to both household income and non-response to the household income survey question. For imputing household income, the list of these variables included survey items related to respondents' feelings about household income, ratings of standards of living, reporting of not having enough money for food, household size, and other variables that may vary by country, such as urbanicity. Below is an illustration:

Louise did not report her exact household income but reported $10k–20k in the follow-up closed-ended item. Her household income was imputed by finding a respondent with the same or very similar characteristics on the survey variables who did report their income and whose reported income was between $10K and $20K. That respondent's income value was used to fill in Louise's household income.

After the imputation of income ranges and missing values, income data were annualised, and per capita annual income was calculated by dividing household income by the total number of persons living in the household. Per capita annual income was used to create income quintiles within each country dataset.

Employment

Gallup classified respondents into one of six categories of employment based on a respondent's combination of answers to a series of questions about employment.

- **Employed full time for an employer:** A respondent is considered employed full time for an employer if he or she is employed by an employer and if he or she works for this employer for at least 30 hours per week.
- **Employed full time for self:** Respondents are considered employed full time for themselves if they are self-employed and if they work for at least 30 hours per week.
- **Employed part time do not want to work full time:** Respondents who work either for an employer or themselves and do not work more than 30 hours per week at either job are categorised as employed part time. Additionally, when asked, these respondents indicate that they do not want to work more than 30 hours per week.
- **Employed part time, want to work full time:** Respondents who work either for an employer or themselves and do not work more than 30 hours per week at either job are categorised as employed part time. Additionally, when asked, these respondents indicate that they do want to work more than 30 hours per week.
- **Unemployed:** A respondent is unemployed if he/she reports not being employed in the last seven days, either for an employer or for himself or herself. The respondent must also report actively looking for a job in the past four weeks AND being able to begin work in the last four weeks.
- **Out of the workforce:** Respondents who are out of the workforce, were not employed within the last seven days, either for an employer or for themselves, are not looking for work, AND/OR are not available to start work. Respondents may be full-time students, retired, disabled or homemakers; however, some respondents will not fall into any of these scenarios.

About gallup world poll metrics used in the analysis

The Gallup World Poll has developed over two dozen indices that summarise how people feel about social, political and economic matters. The Gallup World Poll Research Methodology and Codebook provides detailed information about each of these indices, including the specific survey questions used, the larger concepts measured and additional technical information.

However, two indices – the National Institutions Index and the Communications Access Index – were featured in the analysis[9] and, for the sake of convenience, additional information will be provided here.

National institutions index

The National Institutions Index reflects citizens' confidence in key institutions prominent in a country's leadership: the military, the judicial system, the national government and the honesty of elections.

Index questions

- Do you have confidence in each of the following, or not? How about the military?
- Do you have confidence in each of the following, or not? How about the judicial system and courts?
- Do you have confidence in each of the following, or not? How about the national government?
- Do you have confidence in each of the following, or not? How about the honesty of elections?

Index construction

Index scores are calculated at the individual record level. For each individual record, the following procedure applies: the four items are recoded so that positive answers are scored as a '1' and all other answers (including 'don't know' and 'refused') are assigned a score of '0'. If a record has no answer for an item[10], then that item is not eligible for inclusion in the calculations. An individual record has an index calculated if it has valid scores for at least two of the four items. A record's final index score is the mean of valid items multiplied by 100. The final country-level index score is the mean of all individual records for which an index score was calculated. Country-level weights are applied to this calculation.

Communications access index

The Communications Access Index measures respondents' access to telephone and internet for personal use.

Index questions

- Do you have a landline telephone in your home that you use to make and receive personal calls?
- Do you have a mobile phone that you use to make and receive personal calls?
- Do you have access to the internet in any way, whether on a mobile phone, a computer, or some other device?

Index construction

Index scores are calculated at the individual record level. For each individual record, the following procedure applies: the first two questions (landline telephone and mobile phone) are used to determine whether a respondent has a phone and is used to create the phone component of the index. If respondents answer 'yes' to either question, they are assigned a score of '1' for the phone component and a '0' if they do not have a phone. For the remaining question, positive answers are scored as a '1' and all other answers (including 'don't know' and 'refused') are assigned a score of '0'. An individual record has an index calculated if it has valid scores for both components. A record's final index score is the mean of items multiplied by 100. The final country-level index score is the mean of all individual records for which an index score was calculated. Country-level weights are applied to this calculation.

About external data sources used in the analysis

To better understand how people's attitudes to science and health are shaped by their larger environment, the analysis integrated different data about the countries included in the study. The table below specifies the type of data and source. For all statistics, Gallup used the most recent estimate available, typically 2017 or 2018, but for some statistics that are updated on an infrequent basis (either in general or for specific countries), these estimates could date as far back as the year 2000. The number of countries without data varied by statistic, but science-related indicators, such as GDP as a percentage of research and development, had the highest amount of missing data.

Table II.A:
Summary of external statistics used in analysis

Statistic	Unit	Description	Source/notes	Countries without data
GDP per Capita, PPP	Current international dollars	GDP per capita based on purchasing power parity (PPP)	2017 data as reported on the World Bank website.	2
Gini Coefficient	Index score, 0 to 100	Measure of inequality	World Bank estimate. The most recent estimate was used, dating back to 2000.	12
Annual GDP Growth Rate	Average	The simple average of GDP growth rate from 2008–2017	The average was calculated using data available on the World Bank website.	2
Life Expectancy at Birth	Years of age	Hypothetical life expectancy at birth	World Bank, 2016 or most recent value. Taiwan estimates provided by the country's statistical agency.	1
Cause of Death by Non-Communicable Diseases	Percentage of total deaths	Includes deaths by diseases such as cancer or heart disease	2016 data as reported by the World Bank.	4
Mortality Rate	Per 1,000 births	Number of infants who die before reaching one year of age	2016 data as reported by the World Bank.	2
Research and Development as % of GDP	Percentage of GDP	Gross domestic expenditures on research and development (R&D)	World Bank estimate. Uses most recent estimate available.	33
Number of Researchers in R&D	Per millions of people	Number of researchers working in R&D per million	World Bank estimate. Uses most recent estimate available.	28

Descriptive statistics

In this report of the first wave of Wellcome Global Monitor data, the priority of the analysis was to understand the current state of public attitudes to science and health, especially in those countries and regions of the world where this type of research was infrequent or non-existent. As such, the analysis in this report often relies on the use of descriptive statistics, which present the topline results for each survey question, typically at a country or multinational level.

All results presented by country are weighted to enhance the representativeness of the data (see Survey Methodology for more on this). Because Gallup normalises country-level weights to equal the overall sample size – which for most countries is approximately equal to 1,000 survey results that were aggregated across more than one country (for instance by region or country-income level) were adjusted by the 15+ population size of the countries included in the analysis. This gives larger countries more weight than smaller countries.

For instance, consider the region of East Asia, which consists of five countries for this report but is dominated population-wise by China. The table below shows the country-by-country results for the percentage of people who said they know 'some' or 'a lot' about science. We then show the simple regional average – the average result across the countries, without accounting for population. Finally, the result for the population-adjusted result for the region is shown, which is the statistic used in this report.

Example: Percentage of people who say they know 'a lot' or 'some' about science in countries of East Asia.

Table II.B: Difference between the simple average for an aggregated group of countries compared to the population-adjusted result

Country	A lot/some
South Korea	49%
Taiwan	43%
Mongolia	41%
Japan	32%
China	23%
Simple regional average	38%
Population-adjusted average	26%
Number of researchers in R&D	28

Development of the Trust in Scientists Index

Conceptual Background

An important research objective of the Wellcome Global Monitor was to produce a comprehensive measure of trust in scientists. The first stage of the study focused on understanding this concept better – and how it should be measured. A literature review of the existing research suggested more specific ways in which the public should have confidence in the scientific community, including the accuracy of the work that scientists do, the transparency of their work and funding sources, and their motivation to serve the public good[11]. Furthermore, the research explores whether trust in the scientific community applies equally to scientists regardless of the type of institution they work in, be it for-profit, not-for-profit, or an academic organisation.

Considering these findings, it was clear that the survey would need to capture trust attitudes in several different contexts, requiring several different survey items. The items, or indicators, could then be analysed to measure the underlying concept of interest and overall trust in scientists. As with all questions on the Wellcome Global Monitor, these indicators also needed to be designed in a way that they would allow for cross-country comparability and researchers could be confident that they would be answered in a meaningful way in the 140+ countries in the study.

The testing phase of the project was used to ensure that the final battery of questions met these criteria. In total, seven questions were asked on the 2018 Wellcome Global Monitor that were intended to measure overall confidence in scientists.

Box II.B: Wellcome Global Monitor survey items measuring general trust in science

Response options for all items:
A lot, some, not much or not at all.

- In general, would you say that you trust science?
- How much do you trust scientists in this country?
- In general, how much do you trust scientists to find out accurate information about the world?
- How much do you trust scientists working in colleges/universities in this country to do their work with the intention of benefiting the public?
- How much do you trust scientists working in colleges/universities in this country to be open and honest about who is paying for their work?
- How much do you trust scientists working for companies in this country to do their work with the intention of benefiting the public?
- How much do you trust scientists working for companies in this country to be open and honest about who is paying for their work?

Each question was measured on a four-point scale, indicating that respondents had 'a lot' of trust, 'some' trust, 'not much' trust or 'not at all'.

About the methods used in the index development

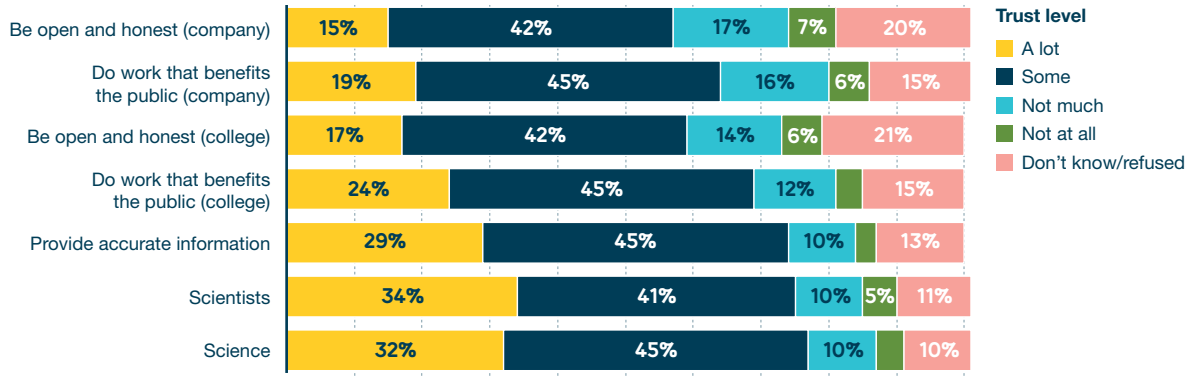
The theoretical framework suggested by the literature review was that the seven indicators, or some subset of them, would represent one single concept: overall trust in scientists. This framework would be tested using confirmatory factor analysis (CFA).

In CFA, indicators included in the model are assumed to measure the underlying variable of interest (the factor) but also contain measurement error. The error terms of the indicators, in the standard CFA approach, are assumed to be independent of each other. CFA requires that the number of factors is specified; here, the number of factors was specified as one.

Handling of 'don't know'/'refused' responses

Prior to the modelling, a review of the survey data highlighted one possible area of concern: across the seven items that were candidates for inclusion in the index, a third of the overall global population did not offer an opinion (e.g. indicated a selection between the possible answer options of 'a lot', 'some', 'not much', and 'not at all'). On a question-by-question basis, the rate of invalid responses (i.e. not offering an opinion) was especially high for the two questions about whether scientists working for companies or colleges are 'open and honest about who is paying for their work' (see Chart II.A below). However, for each of the seven questions, the rate of invalid responses was at least 10%.

Chart II.A:
Items reviewed for trust in scientists index, global results



In some countries or regions, large percentages of people provided an invalid response to at least one of the seven questions, such as Cambodia, Togo, Kuwait and the United Arab Emirates (see table below). In Cambodia, nearly half of people said 'don't know' or refused to answer at least five of the seven questions.

Table II.C:
Countries with the highest percentage of 'don't know' or 'refused' rates to at least ONE of the seven items on trust in science

Country	Percent
Cambodia	76%
Togo	69%
Kuwait	69%
UAE	66%
Iran	66%

Data in percentages of countries who gave a specified number of invalid responses, ranked by the percentage who gave an invalid response to at least one of the seven questions

These rates of invalid responses posed problems for the representativeness of the index, at least in some countries. As such, researchers considered imputing their responses or using statistical methods to provide the best estimate as to how a person would have answered a question. Researchers ultimately decided this action. Imputing – or using an algorithm to estimate as to how a person who did not answer one of the questions might have responded – did not seem appropriate given that this was the first wave of data collection.

As a result, only respondents who answered all of the survey questions being looked at in a substantive manner (answering 'a lot,' 'some,' 'not much,' or 'not at all') were included in the CFA stage of the index development process. In the actual calculation of the index, however, individuals who answered at least three of the five index questions were included, somewhat easing the problem of representativeness.

Model development

The global pooled sample was used in fitting the initial model. The initial model tested all seven items, treating the data as ordinal. Errors were assumed to be uncorrelated.

Standard goodness-of-fit measures suggested this model was a poor fit, including Comparative Fit Index (CFI) (0.855), Root Mean Square Error of Approximation (RMSEA) (0.165) and p-value for Chi-square greater than 0.05. To improve model fit, the "trust in science" item was removed from the solution, as the question has a somewhat different focus than the other items in the series. Moreover, an analysis of the reliability of the items at the country level suggested that the item about corporate scientists being open and honest should also be excluded, as reliability tended to improve at the country level when this item was excluded. (See Box 3.2 in Chapter 3 for the final 5 items in the Wellcome Global Monitor Trust in Scientists Index).

The final model included the remaining five items, with reasonably strong goodness-of-fit measures, with Chi-square below 0.05 and CFI equal to 0.975. RMSE was slightly high at 0.99.

Reliability test

For the five-item model at the global level, Cronbach's alpha, a test of how reliable a group of indicators are, was equal to 0.801, a result generally accepted as high and above the commonly used threshold of 0.7[12]. Examining the reliability of these items on a country-by-country basis reveals that in 124 of the countries in the study, Cronbach's alpha is above 0.7. In another 14 countries, Cronbach's alpha falls between 0.65 and under 0.7, a range generally considered acceptable[13]. The lowest Cronbach's alpha score at the country level is 0.561, in Austria.

Index Calculation

The Trust in Scientists Index takes the simple average of the five items[14] returning a score between one and four. To receive an Index score, respondents must have answered at least three of the five questions in a substantive manner (i.e. they did not answer 'don't know' or 'refused'); otherwise, respondents are considered as registering 'no opinion'. To ease interpretation of the Index value, the results are sometimes presented in a categorical descriptive fashion, with respondents possessing either low, medium or high trust. In terms of its numerical span, the 'low' category seems the largest, but this is in part because this category encompasses essentially two response options from the original scale: 'not much', and 'not at all'. Due to the positive skew of the data, a more granular category system was deemed inadvisable.

An alternate method of calculating the Index was tested using the unstandardised regression weights from the CFA, i.e. providing the 'latent' score rather than a direct average. Applying these weights would give a slightly higher representation of the question about whether a person trusts college scientists to do work that benefits the public (1.24) and if college scientists do work that benefits the public (1.17), compared to the other three items, which have weights equal or nearly equal to 1.

The two scoring approaches produce largely similar results. The correlation between the latent index scores calculated for every country in the study and the average index score is 0.999. The country 'rankings' are also largely identical between the two approaches, with Spearman's Rho also equal to 0.999. Considering this, the approach of taking the mean score was adopted, if only for its simplicity.

Regression analysis to identify predictors of trust in scientists index

Another aim of the analysis was to build on the existing research and estimate the relationship between individual and country-level predictors on trust in science. The dependent variable in this model is the Trust in Science Index (hereafter 'Trust') previously described. We want to estimate the effect of several individual and country-level variables on Trust, including[15]:

- **Demographic Factors:** Gender, Age, Urbanicity
- **Education and Science Knowledge:** Educational Attainment, Whether a Person has Been Taught Science at any Level, Highest Level of Education in Science, Knowledge Test
- **Wellbeing:** Life Evaluations, Income Quintile, Feelings about Household Income
- **Access to Communications:** Communications Access Index, Communications Use Index
- **Religiosity:** Have Religion, Religion Important
- **Country-Level Variables:** Inequality (Gini), Gini Groups, Country GDP, World Bank Income Classification of Country, Country GDP Growth, Country Life Expectancy, Country Mortality Rate, R&D as a Percentage of GDP.

These variables were selected by the Wellcome Global Monitor team based on their theoretical relevance to explain *Trust*. Preliminary data screening identified a large amount of missing data for 'R&D as a Percentage of GDP', comprising 22% of the total unweighted sample and 33 out of 144 countries. Missing values were imputed via Expectation Maximisation at the country level, using all other country-level predictors and the regional classification variable as auxiliary variables. Additionally, the 'Mortality Rate per Live 1,000 Births' variable showed some non-normal properties (skewness = 1.18, SE of skewness = 0.001), which were corrected through a logarithmic transformation. A preliminary regression analysis showed no signs of collinearity.

Using the final transformed variables, we estimate a series of nested hierarchical linear models using the Stata xtmixed command, with two levels (individual and country) and sampling weights at the individual level, and a projection factor at the country level to represent the correct number of adults in the population. Weights are scaled by size, with individual-level weights be scaled summing to the sample size within each country. Country-level weights are left unchanged.

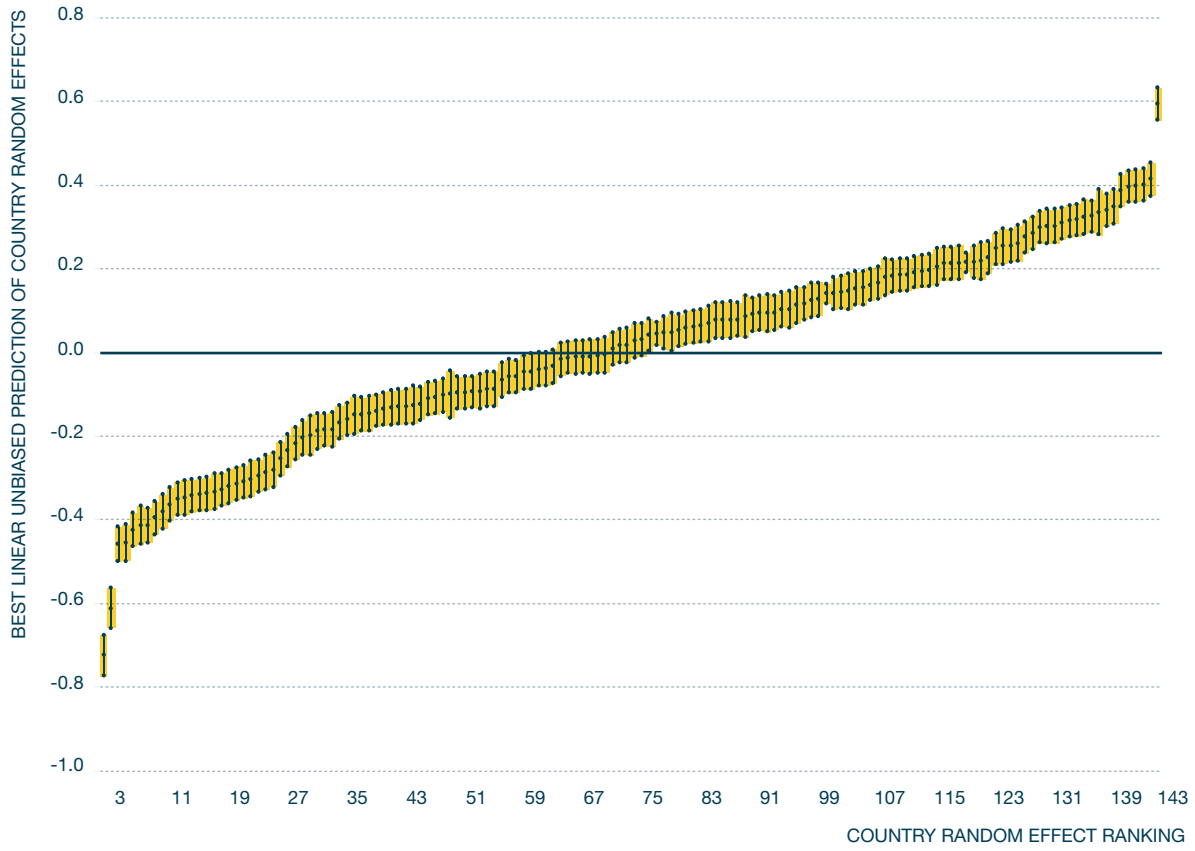
Results of regression analysis

We are interested in identifying the main correlates of *Trust* in science at the individual and country level, and the relative importance of predictors at both levels. For this nested modelling approach, a null intercepts-only respondent-level model with no predictors is estimated as the starting point,

$$T_{ij} = \beta_0 + U_{0j} + e_{ij} \qquad (1)$$

where the *Trust* score *T* of respondent i in country j are equal to the total population mean β_0, plus a country-specific effect *U* for each country *j*, plus an individual-level error e_{ij}. The country effects U_{0j} and individual-level errors e_{ij} are assumed to follow a normal distribution with mean = 0 and variance σ^2_{u0} (country) and σ^2_e (individual). The overall mean β_0 (across countries) is estimated as 2.99. The between-country variance in T is estimated as $\sigma^2_{u0} = 0.03$, and the within-country between-individual variance is estimated as $\sigma^2_e = 0.34$, for a total variance of 0.03 + 0.34 = 0.37. The intraclass correlation coefficient (ICC) is 0.03/0.37 = 0.08 , indicating that 8% of the variance in Trust can be attributed to differences between countries. Country-level residuals U_{0j} for T and their standard errors are presented in Chart II.B. Country-level random effects range from -0.72 to 0.59, though 141 out of 144 countries fall within the narrower range of -0.5 to 0.4. Togo and Mauritania stand out as showing particularly low Trust scores, with country intercepts of -0.72 and -0.61 respectively. Uzbekistan, on the other extreme, shows a particularly high Trust score, with a country intercept of 0.59.

Chart II.B:
Best linear unbiased prediction of country random effects



In order to determine what factors are associated with Trust at the individual level, we add to equation (1) a vector X of predictors $(X_1,\ldots,X_p)$ varying for each individual i in each country j, with coefficients β_1 $(\beta_1,\ldots,\beta_p)$:

$$T_{ij} = \beta_0 + \beta_1 X_{ij} + U_{0j} + e_{ij} \qquad (1)$$

Comparison between this model and the intercepts-only null model shows that T is predicted better after the inclusion of individual-level predictors, according to both the Akaike information criterion (AIC) and Bayesian information criterion (BIC). An inspection of variable-specific coefficients shows that only a small subset of individual-level variables has an effect that is significantly different from zero. In order to avoid convergence problems (e.g. Tabachnik & Fidell, 2007) , we simplify into a parsimonious model only including significant predictors: urbanicity, highest level of science education, income quintile, feelings about household income, communications access, and communications use. Since communications access and use are very strongly correlated (r = 0.96) we only keep communications access, the stronger predictor. Likewise, income quintiles and feelings about household income are collinear, so we keep feelings about household income as the stronger predictor. In order to determine what factors are associated with Trust at the country level, we add to equation (1) all the country-level fixed effects, represented by the new term $\beta_2 Y_j$:

$$T_{ij} = \beta_0 + \beta_2 Y_j + U_{0j} + e_{ij} \qquad (2)$$

where $\beta_2 Y_j$ represents a vector of fixed coefficients β_2 on country-level covariates Y_j. An inspection of variable-specific coefficients shows that only a small subset of country-level variables has an effect that is significantly different from zero, including GDP growth and inequality levels (Gini coefficient). We finally add all remaining individual and country-level predictors, after removing insignificant individual-level predictors for parsimony:

$$T_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 Y_j + U_{0j} + e_{ij} \qquad (3)$$

The final model is presented in the table below. Both AIC and BIC show this model to be superior to either of the previous models. The model's ICC = 0.067, lower than for the null (ICC = 0.091), and indicates that the final model can explain some (though not all) of the variation across countries, suggesting that some of the between-country differences are due to

compositional differences on the individual-level variables. A closer examination of predictors shows science education to be one of the strongest correlates of Trust: receiving science education up to college level is associated with an increase in Trust of 0.27 points (on a 0–4 scale). Income also plays an important role. All other things being equal, 'finding it very difficult to get by on present income' was associated with a 0.20 drop in Trust. Location plays a role above and beyond income and education. Other variables in the model being equal, individuals in rural locations tend to have a higher Trust. Relative to living in a rural area, living in a small town or village is associated with a decrease in Trust of 0.07, whereas living in a big city or suburb of a big city is associated with a decrease in Trust of 0.08. Finally, country-level variables also have significant effects. Each percentage point increase in GDP growth is associated with an increase of 0.02 points in Trust. Inequality, on the other hand, is negatively associated with Trust. A unit increase in the Gini coefficient is associated with a 0.01-point decrease in Trust.

Table II.D: Confirmatory model – selected regression analysis output

					95% CI	
Trust in Scientists Index	Coef.	Robust Std. Err.	z	P>z	Low	High
Urbanicity – small town or village	-0.07	0.02	-2.97	0.00	-0.11	-0.02
Urbanicity – big city or suburb of big city	-0.08	0.02	-3.27	0.00	-0.12	-0.03
Science education – at primary only	0.09	0.02	4.54	0.00	0.05	0.13
Science education – up to secondary	0.18	0.03	6.24	0.00	0.13	0.24
Science education – up to college	0.25	0.06	4.48	0.00	0.14	0.36
Feelings about Hh income – getting by on present income	-0.09	0.03	-2.59	0.01	-0.16	-0.02
Feelings about Hh income – finding it difficult on present income	-0.16	0.04	-4.29	0.00	-0.24	-0.09
Feelings about hh income – finding it very difficult on present income	-0.20	0.06	-3.38	0.00	-0.31	-0.08
Communications access index	0.00	0.00	0.97	0.33	0.00	0.00
Gdp growth	0.02	0.01	2.34	0.02	0.00	0.03
Gini coefficient	-0.01	0.00	-4.12	0.00	-0.01	0.00
Constant	3.22	0.12	26.61	0.00	2.98	3.45
Intraclass correlation	**0.069**	**0.012**			**0.048**	**0.097**

AIC = 7.64e+09, BIC = 7.64e+09.

Coefficients for 'DK'/'Refused' categories omitted.

The model above shows relatively low fit, with an ordinary least squares (OLS) model including country-fixed effects showing an R2 = 0.11, indicating that the model can explain a maximum of 11% of the variation in Trust scores. A stronger model should expand the range of individual and country-level predictors of Trust. A focus on individual-level predictors is particularly warranted, considering that the remaining amount of unexplained country-level variance in our final model is low. With this in mind, we proceeded with an exploratory round of analysis incorporating all the indexes in the Gallup World Poll. Gallup indexes span multiple political, social and economic topics that make up the interconnected components of the Gallup World Poll.

Our exploratory analysis identified the National Institutions Index as a particularly strong predictor of Trust. The National Institutions Index reflects citizens' confidence in key institutions prominent in a country's leadership: the military, the judicial system, the national government and the honesty of

elections. We incorporated the individual predictors that go into the National Institutions Index, with the exception of 'honesty of elections', where data was only available for 97 countries. The results show that confidence in national institutions, including the military, judicial system and courts, and the national government all have a strong association with Trust; e.g. having confidence in the military is associated with an increase of 0.16 in Trust.

Table II.E: Exploratory model – selected regression analysis output

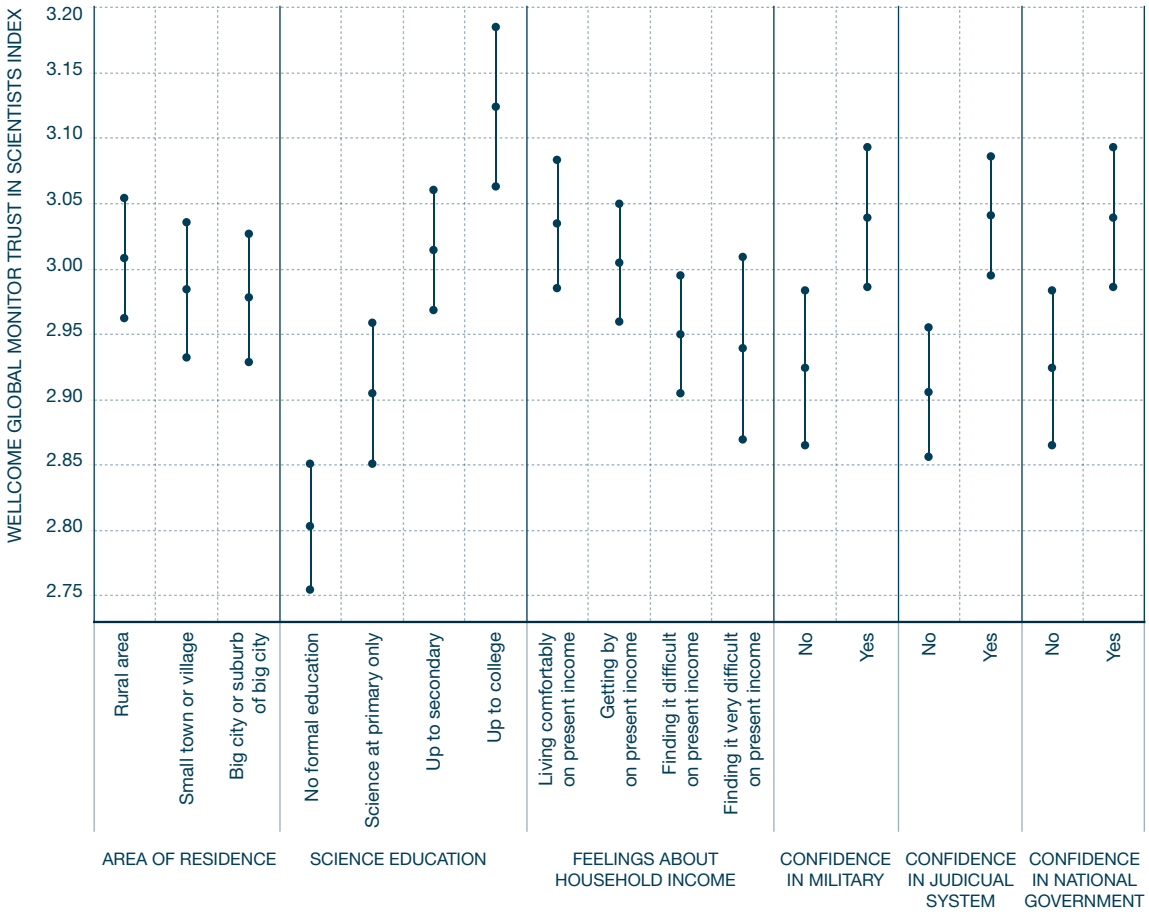
					95% Conf. Interval	
Trust in Scientists Index	Coef.	Std. Err.	z	P>z	Low	High
Urbanicity – small town or village	-0.02	0.01	-2.56	0.01	-0.04	-0.01
Urbanicity – big city or suburb of big city	-0.03	0.02	-1.46	0.14	-0.07	0.01
Science education – at primary only	0.10	0.02	6.35	0.00	0.07	0.13
Science education – up to secondary	0.21	0.02	11.80	0.00	0.18	0.25
Science education – up to college	0.32	0.03	10.69	0.00	0.26	0.38
Feelings about Hh income – getting by on present income	-0.03	0.01	-2.30	0.02	-0.05	0.00
Feelings about Hh income – finding it difficult on present income	-0.08	0.01	-6.56	0.00	-0.11	-0.06
Feelings about Hh income – finding it very difficult on present income	-0.10	0.03	-3.48	0.00	-0.15	-0.04
Communications access index	0.00	0.00	2.56	0.01	0.00	0.00
Gdp growth	0.00	0.01	-0.17	0.87	-0.02	0.01
Gini coefficient	-0.01	0.00	-4.70	0.00	-0.01	0.00
Has confidence in the military	0.16	0.01	16.86	0.00	0.14	0.18
Has confidence in judicial system and courts	0.14	0.01	9.47	0.00	0.11	0.16
Has confidence in national government	0.11	0.03	3.37	0.00	0.05	0.18
Constant	2.78	0.10	27.89	0.00	2.59	2.98
Intraclass correlation	**0.05**	**0.01**			**0.03**	**0.08**

AIC = 5.81e+09, BIC = 5.81e+09.

Coefficients for 'DK'/'Refused' categories omitted.

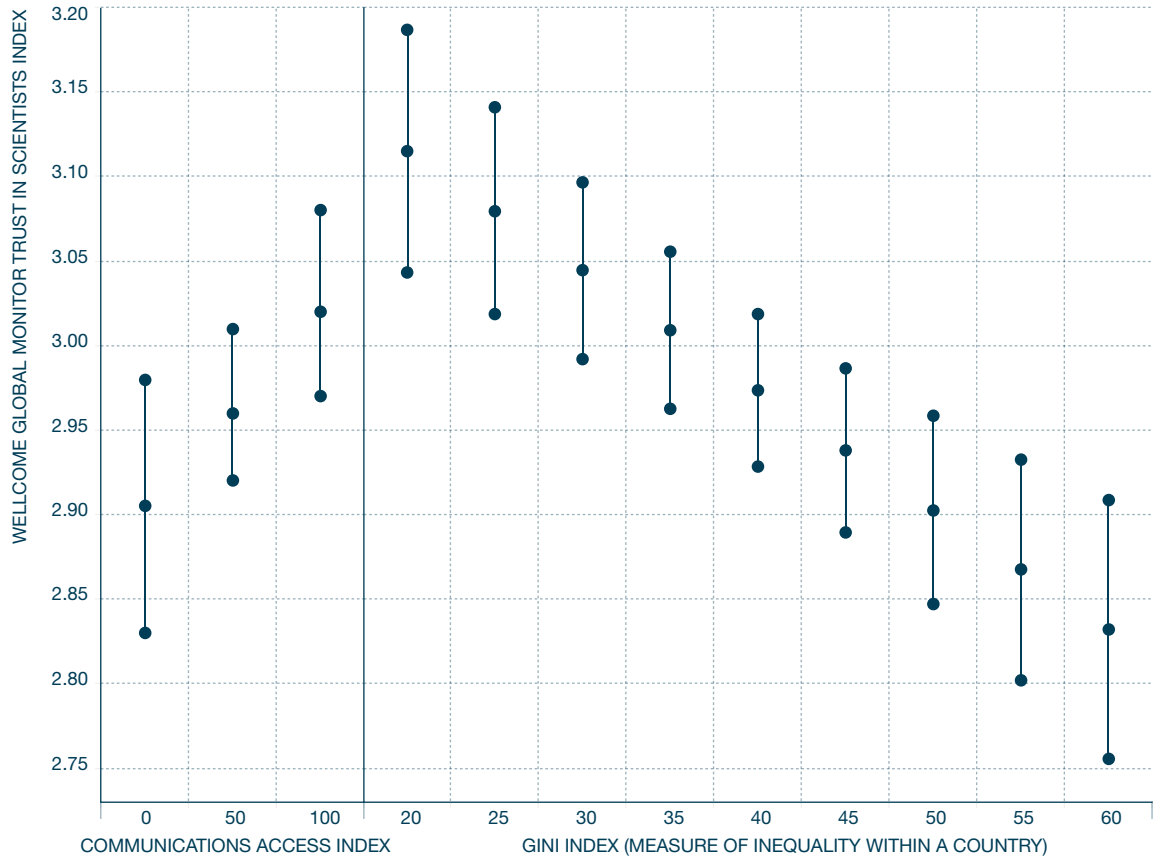
For ease of interpretation, margin plots are presented for each of the variables in the final model, excluding those that no longer show an effect significantly greater than 0 after the inclusion of confidence in national institutions, i.e. GDP growth. Margin plots present the model's predicted Trust in Scientists Index score for a given variable level, keeping all other factors at their average level. For example, the model predicts that an individual with 'No formal education' in science would have a Trust in Scientists score of 2.8, while an individual with 'Up to college' education in science would show a Trust in Scientists score of 3.1. Margin plots make it easier to visualise the relative significance of each variable. For example, it is clear that 'Urbanicity', even if it may be a significant predictor of Trust in Scientists, only makes a small difference, compared to a factor such as science education, or a country-level variable such as the Gini coefficient.

Chart II.C:
Margin plots for categorical variables



All margins calculated as linear predictions (fixed portion). Error bars represent the 95% confidence interval for each prediction.

Chart II.D:
Margin plots for continuous variables



All margins calculated as linear predictions (fixed portion). Error bars represent the 95% confidence interval for each prediction.

This final model, while improved, still shows a relatively modest explanatory power (R2 = 0.15). Subsequent waves of the Wellcome Global Monitor will aim to expand the range of individual and country-level predictors of Trust. Given the ample room to improve the model fit, such efforts should probably begin with in-depth formative research to ascertain the likely determinants of Trust, particularly at the individual level.

Regression analysis appendix

Demographic Factors

- Gender (1 = Male, 2 = Female).
- Age (in Years).
- Urbanicity (1 = Rural Area, 2 = Small Town or Village, 3 = Big City or Suburb of Big City, 99 = Missing/DK/Refused).
- If person has a specific religion (1=Yes, 0=All Other Responses)
- If person says religion is important in their daily life, or 'religiosity' (1=Yes, 0=All Other Responses)

Education and Science Knowledge

- Educational Attainment (1 = Completed Primary Education or Less, 2 = Secondary–3-Year Tertiary, 3 = Completed Four Years of Education Beyond High School and/or Received a 4-Year College Degree, 997 = Missing/DK/Refused).
- Whether a person has been taught science at any level (0 = No Science Education at Any Level, 1 = Has Learned Science at Some Level, 997 = Missing/DK/Refused).
- Highest Level of Education in Science (0 = No Formal Education, 1 = Science at Primary Only, 2 = Up to Secondary, 3 = Up to College, 997 = Missing/DK/Refused).
- Knowledge Test (0 = Did Not Give Right Answer on Both Items, 1 = Did Give Right Answer on Both Items). Test Items Included: 'Studying Diseases is a Part of Science' and 'Writing Poetry is a Part of Science'.

Wellbeing

- Life Evaluations: Cantril Self-Anchoring Striving Scale (Cantril, 1965). The question uses a scale from 0 to 10 and asks respondents:
  - Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. Suppose we say that the top of the ladder represents the best possible life for you, and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time, assuming that the higher the step the better you feel about your life, and the lower the step the worse you feel about it? Which step comes closest to the way you feel?
- Income Quintile: Per Capita Annual Income in International Dollars, divided into quintiles within countries (1 = Poorest 20%, 2 = 21%–40%, 3 = 41%–60%, 4 = 61%–80%, 5 = Richest 20%).
- Feelings About Household Income (1 = Living Comfortably on Present Income, 2 = Getting by on Present Income, 3 = Finding it Difficult on Present Income, 4 = Finding it Very Difficult on Present Income, 5 = Missing/DK/Refused).

Access to Communications

- Communications Access Index: The Communications Access Index measures respondents' access to telephone and internet for personal use, based on the following three items:
  - Do you have a landline telephone in your home that you use to make and receive personal calls?
  - Do you have a mobile phone that you use to make and receive personal calls?
  - Do you have access to the internet in any way, whether on a mobile phone, a computer, or some other device?

  The first two questions (landline telephone and mobile phone) are used to determine whether a respondent has a phone and is used to create the phone component of the index. If respondents answer 'yes' to either question, they are assigned a score of '1' for the phone component and a '0' if they do not have a phone. For the remaining question, positive answers are scored as a '1' and all other answers (including 'don't know' and 'refused') are assigned a score of '0'. An individual record has an index calculated if it has valid scores for both components. A record's final index score is the mean of items multiplied by 100.
- Communications Use Index: The Communications Use Index measures respondents' access to telephone and internet for personal use, and use of the internet in the past seven days, based on the following four items:
  - Do you have a landline telephone in your home that you use to make and receive personal calls?
  - Do you have a mobile phone that you use to make and receive personal calls?
  - Do you have access to the Internet in any way, whether on a mobile phone, a computer, or some other device?
  - Have you used the Internet in the past seven days, whether on a mobile phone, a computer, or some other device?

Confidence in Institutions

- Do you have confidence in each of the following, or not? How about the military? (1=yes; 0=all other responses)
- Do you have confidence in each of the following, or not? How about the judicial system and courts? (1=yes; 0=all other responses)
- Do you have confidence in each of the following, or not? How about the national government? (1=yes; 0=all other responses)

Country-Level Variables

- Indicators in this analysis include:
  - GINI Coefficient (value between 0-100)
  - World Bank Country Income Level (1=Low income, 2=Lower-middle income, 3=Upper middle income, 4=High income)
  - Country GDP Growth Rate (simple average, 2008-2017)
  - Country GDP per capita PPP (current international dollars)
  - Country Mortality Rate (number of deaths per live 1,000 births)
  - Country Average Life Expectancy (average age of death)
  - Expenditures on research and development as a percentage of GDP (% of GDP)

Further information about these data series, including their source and recency, is provided in Table II.A above. Please note that data transformations were performed on some of these series in the analysis.

Endnotes

1 Retrieved online at: https://wellcome.ac.uk/sites/default/files/wellcome-global-monitor-questionnaire-development-report_0.pdf
2 This formula is calculated at the 95% confidence level, i.e. $\alpha = .05$, resulting in $z\alpha/2 = 1.96$.
3 The design effect was defined formally by Kish (1965, Section 8.2, p. 258) as 'the ratio of the actual variance of a sample to the variance of a simple random sample of the same number of elements.' Based on Kish's approximate formula {design effect = (sample size)*(sum of squared weights)/ (square of the sum of weights)}
4 Retrieved online at: https://unstats.un.org/unsd/methodology/m49/
5 Retrieved online at: https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries
6 All definitions are according to the World Bank's most recent update of these categories, in July 2018. This update uses 2017 Gross National Income information.
7 As discussed in the Gallup World Poll Methodology and Codebook (pages 12–14).
8 See page 14 of the Gallup World Poll Methodology and Codebook.
9 Both measures were used as explanatory variables in the drivers of trust analysis. Additionally, country results regarding views on the exclusivity of the benefits of science, confidence in healthcare and the categorical trust in science index were compared to the National Institutions Index.
10 This occurs when the respondent has not been asked the item, typically because the question was omitted from that country's Gallup World Poll questionnaire.
11 For further information, please see the literature review of the 'Wellcome Global Monitor Development Report'.
12 Retrieved online at: https://www.ijme.net/archive/2/cronbachs-alpha.pdf
13 Retrieved online at: https://data.library.virginia.edu/using-and-interpreting-cronbachs-alpha/
14 This approach assumes equal weights for the items, an assertion that will be tested and, if necessary, revised in future waves of the Wellcome Global Monitor.
15 See appendix for comprehensive variable descriptions.