

DOING SCIENCE IN TIMES OF COVID-19: A SURVEY

Methodology Memo

Karlijn Roex (k.l.a.roex@uva.nl), Giovanni Colavizza (g.colavizza@uva.nl)

University of Amsterdam

Abstract | This NWO Corona Fast-Track Data study is a two-wave (non-panel) survey of academic researchers working at Dutch research institutions. Through this survey we aimed to study scientists' information seeking and sharing behaviour with regard to the latest scientific information on COVID-19. In the midst of the COVID-19 pandemic, the expertise of researchers from medicine and biology, but also social sciences, humanities and economics is urgently needed to counter misinformation. In order to gain timely and otherwise lost insights on how to improve science communication during this crisis, we collected the data in two survey waves. The first wave survey data were collected between 9 June and 31 August; the second between 26 October and 30 November 2020. In this document we provide some technical methodological background information on the data.

Sampling process

For each survey wave, a separate sample was taken, rather than relying on a fixed panel. Respondents were recruited through a sampling process that was partly convenience, partly random, and partly stratified. Partly convenience, as we first collected data within the universities where we were affiliated (University of Amsterdam and Leiden University). The remaining institutions were chosen randomly. Finally, the sample was partly stratified, as we used universities as an embedding stratum of where to invite respondents. In this stratified selection we deliberately made sure to include some technological universities, to ensure getting a wider range of academic disciplines. In an invitation e-mail or news item potential respondents were informed about the study and invited to complete a web-based survey. We optimized the sampling process through the research project, which is why the sampling process of Wave 1 is slightly different from that of Wave 2. We will describe both sampling processes below.

Wave 1 (09 June – 31 August 2020)

For Wave 1, the respondents were recruited through a multi-stage convenience sampling process – meaning that we first used one method to recruit respondents and then another one to recruit more. The stages were:

Stage 1, June: Advertisement via faculty, interfaculty and university-wide newsletters from the two institutions where the project researchers (Giovanni Colavizza and Karlijn Roex) were based. Those institutions were:

- o University of Amsterdam (UvA)
- o Leiden University

Stage 2, July: Direct e-mailing of university personnel of a random sample of Dutch universities whose e-mail addresses were collected through Web scraping. Table 1 shows the sample of universities (or university centres).

Table 1. Universities in sample.

Institution	No. of e-mails sent	Bounced
Erasmus MC (university medical centre)	2,545	247
Leiden University Medical Centre (LUMC)	2,044	466
Utrecht University	4,734	102
Tilburg University	3,172	525
Radboud University	2,589	314
Radboud University Medical Centre (RUMC)	2,165	0
Eindhoven University of Technology	4,558	0
Amsterdam Medical Centre (AMC)	692	42

The number of researchers we have reached through the newsletters in early June is unknown. In potential, this could have been 2,824 research staff at UvA and 2,372 at Leiden University (Rathenau

Instituut, 2020).¹ In addition, in early July we sent a total of 22,499 e-mails. We can assume that 10% were inoperative e-mail addresses (1,696 bounced back, which corresponds to 7.5%), so that the actual number of people reached through direct e-mailing was approximately 20,249.

Moreover, we can assume that many contacted people, while being employees or affiliates, were not actual academic researchers (i.e., non-scientific staff). Regarding this low specificity, we improved the sampling for Wave 2. We can assume that only 40% to 60 % of the contacted persons were actual scientific staff. The Rathenau database of university staff indicates that, on average, 53.8% of the staff was scientific personnel in 2019.² In the case that 40% of the reached contacts were scientists, we have contacted 8,099 researchers through direct e-mailing. Adding those reached with newsletters, this makes a total for Wave 1 of 13,295 researchers. Of this, we received 777 responses – giving a net response rate of 5.8%. For the given research population, we do not know whether this response rate is usual or not. For instance, the Labour Force Survey – one of the most important household surveys in The Netherlands – has a response rate of 60%.³ However, that survey was (at least partly) conducted with home visits by an interviewer. The mode (visit, telephone, online) determines the response rate⁴, and therefore this response rate is not comparable with ours. A 2002 online survey administered to Dutch households, offering them incentives to participate, had a response rate of 20.4%.⁵ A 2014 Taylor & Francis survey directed to academic researchers in the European Union, achieved a net response rate ('net' here means after omitting bounced back e-mails) of 9.2%.⁶

1

<https://www.rathenau.nl/nl/wetenschap-cijfers/wetenschappers/personeel-aan-de-universiteiten-en-umcs/wetenschappelijk-en>.

2

<https://www.rathenau.nl/nl/wetenschap-cijfers/wetenschappers/personeel-aan-de-universiteiten-en-umcs/wetenschappelijk-en>.

³ Cobben, F. (2009). *Nonresponse in sample surveys: methods for analysis and adjustment*. The Hague: Statistics Netherlands.

⁴ Deutskens, E., Ruyter, De. K., Wetzels, M. & Oosterveld, P. (2004). Response rate and response quality of internet-based surveys: an experimental study. *Marketing Letters*, 1, 21-36.

⁵ Deutskens, E., Ruyter, De. K., Wetzels, M. & Oosterveld, P. (2004). Response rate and response quality of internet-based surveys: an experimental study. *Marketing Letters*, 1, 21-36.

⁶ Frass, W., Cross J. & Gardner, V. (2014). *Taylor & Francis Open Access Survey. Annex C. European Union*. London: Routledge/ Taylor & Francis. p. 4.

Wave 2 (26 October – 30 November 2020)

In contrast to the Wave 1 sampling process, this time we recruited all respondents in the same way: approaching them through direct e-mail. Table 2 shows the sample of institutions from which scientists were receiving our invitation.

Table 2. Universities in sample.

Institution	No. of e-mails sent	Bounced
University of Amsterdam (UvA)	2,813	16
Leiden University	3,619	106
Delft University of Technology	2,433	150
Erasmus University	1,188	3
Radboud University	102	2
Utrecht University	109	0
Maastricht University	537	7

For Wave 2, we checked for all manually-gathered e-mail addresses whether the corresponding staff members actually were academic researchers. In total, we sent 10,801 e-mails. This time, only 2.6% of the e-mails were bounced. However, we received many auto-replies that indicated that people have left office (e.g., have left their job). We can therefore assume that slightly more than only the bounced-back e-mails were inoperative e-mail addresses (around 3-5% of all sent e-mails). This makes the actual number of researchers reached through direct e-mailing, at best, 10,476. Of this, we received 1,167 responses – giving a response rate of 11.4%. The higher response rate for Wave 2 may be explained by the stricter criteria we used to reach participants (carefully checking online whether the university staff are actually researchers). Moreover, the more detailed information we provided about respondents' privacy in the Wave 2 invitation, may have gained trust.

Questionnaire

We used the software Qualtrics for designing the online questionnaire and for collecting the data. This software enabled us to make a user-friendly questionnaire for both PC-users as well as mobile device users. The questionnaire can be seen in the document ‘Questionnaire’, and in the document ‘Codebook’ we list the survey items.

Important note on Wave 1 survey data on *des_influence_1* and *des_influence_2*

Due to an error in the item *des_influence_1*, one should be cautious in interpreting any results of analyzing the question about the idealized influence of scientific experts on COVID-19. In order to help researchers to make their own decisions in how to interpret results from their analyses with this item, we provide a report of the error here.

The item *des_influence_1* was meant to deal only with the idealized influence of experts in *popular news outlets*, the item also dealt with social media – which was meant to be the exclusive terrain of item *des_influence_2*, not *des_influence_1*. Unfortunately, this is not what appeared in the eventual survey. Therefore, many respondents could have interpreted the item *des_influence_1* as if this item deals with social media. We corrected this error in the questionnaire before sending it out for Wave 2, so this issue only applies to Wave 1 data. Figure 1a shows the correct version, while Figure 1b shows the incorrect version.

Figure 1 Error in item *des_influence_2*

(a) Correct version

UNIVERSITY OF AMSTERDAM
Faculty of Social and Behavioural Sciences

And how large should their influence ideally be at this moment, in your opinion?

0 10 20 30 40 50 60 70 80 90 100

popular national or international news outlets

social media (e.g. Twitter)

national policymaking

← →

Detailed description: This is a screenshot of a survey question. At the top, it shows the University of Amsterdam logo and the Faculty of Social and Behavioural Sciences. The question asks for the ideal influence of three entities: popular news outlets, social media, and national policymaking. A horizontal scale from 0 to 100 is provided. Three sliders are shown, each with an orange dot at the 0 position. Navigation arrows are at the bottom.

(b) Incorrect version

UNIVERSITY OF AMSTERDAM
Faculty of Social and Behavioural Sciences

And how large should their influence ideally be at this moment, in your opinion?

0 10 20 30 40 50 60 70 80 90 100

popular national or international news outlets and social media

social media (e.g. Twitter)

national policymaking

← →

Detailed description: This is a screenshot of the same survey question as in (a), but with an error. The first slider label, "popular national or international news outlets and social media", incorrectly combines the two categories from the correct version. The other elements, including the question, scale, and navigation arrows, are identical to the correct version.

Wave 2 improvements

Our survey had an ‘End comment’ section where respondents could leave their comments about the survey (see the datasets. Reading those is recommended). After collecting the data, we carefully read their comments and some gave rise to some improvements of the questionnaire. In this, we had to balance the importance of standardized measurements (i.e., if we have an entirely different questionnaire, we cannot make any meaningful comparisons between Wave 1 and 2) and making the survey more convenient for respondents to complete. Many respondents were worried about the so-called ‘estimation questions’, in which they were asked to estimate the extent to which they used certain types or channels of information. A number of respondents had wished that there was a follow-up question through which they could communicate the reliability of their judgement to us. We included such questions in the Wave 2 questionnaire.

Moreover, before asking respondents whether they had recently corrected any misinformation they had seen, in Wave 2 we asked them whether they had seen any misinformation at all.

Representativity analysis and weighting recommendations

We compared several socio demographic characteristics of our sample to statistics on the full research population (i.e., academic researchers at Dutch institutions). These statistics are available at the Rathenau Institute’s website (2019; 2020).⁷ Tables 3-6 below display the sample and population distributions for several key demographic variables. Our sample appeared to be highly representative regarding age and career stage (*cr_stage*), and therefore no weighting was needed for these factors. Still, survey weights are recommendable that ensure that the resulting weighted data are as much as representative of the population of academic researchers in The Netherlands as possible. Our sample is especially showing an overrepresentation of some academic fields and institutions.

We recommend using a raking estimator to calibrate the weights so that they would match the Rathenau proportions on gender, research institution (available at request with the authors) and field. We recommend only to weigh for institutions that are widely included in the sample. Therefore, it is advised to be explicit in reports about the fact that conclusions from analysis on

⁷ ([Cijfers over universitair personeel, WOPI, Rathenau Instituut, 2019](#) and [Het personeel bij de universitair medische centra, Rathenau Instituut, 2020](#)).

these data are not unproblematically generalizable to, for instance, researchers at Maastricht University or Wageningen University.

Table 3 Sample and population distributions of institution and gender⁸

⁸ For each variable, respondents with missing values on that variable were omitted. For the population gender distribution, we found that the proportion of women was much higher in academic hospitals ([separate Rathenau data](#)) - namely 61.59% - than the 40.59% in the non-medical faculties ([general Rathenau statistics](#) where the medical faculty was omitted). We calculated an overall gender distribution using information on the total FTE of scientific staff in the academic hospitals and the other faculties. We used the following function: proportion of women in scientific staff of academic hospitals * share of FTE of scientific staff in academic hospitals compared to all scientific staff + proportion of women in scientific staff in other faculties * share of FTE of scientific staff in all other faculties. The proportion of women is much higher in the universities that have an academic hospital (e.g., Radboud University) compared to those that have not (e.g., Tilburg University). Rathenau data on the academic hospital of the University of Amsterdam (AMC) was missing.

	Wave 1 (N = 471)			Wave 2 (N = 707)			Rathenau (fte = 27005)		
	Women	Men	Other	Women	Men	Other	Women	Men	Other
Leiden University	0.56	0.44	0	0.49	0.51	0	0.53	0.47	NA
Utrecht University	0.45	0.54	0.01	0.36	0.64	0	0.52	0.48	NA
Groningen University	0	1	0	0	1	0	0.52	0.48	NA
Erasmus University	0.64	0.36	0	0.63	0.37	0	0.57	0.43	NA
Maastricht University	NA	NA	NA	0.59	0.41	0	0.53	0.47	NA
University of Amsterdam	0.65	0.35	0	0.57	0.42	0.01	0.44	0.56	NA
Vrije Universiteit Amsterdam	NA	NA	NA	0.50	0.50	0	0.54	0.46	NA
Radboud University	0.51	0.49	0	0.53	0.47	0	0.53	0.46	NA
Tilburg University	0.63	0.36	0.01	1	0	0	0.44	0.56	NA

Delft University	0	1	0	0.37	0.62	0.01	0.26	0.74	NA
Eindhoven University	0.27	0.73	0	0.38	0.62	0	0.28	0.72	NA
University of Twente	NA	NA	NA	1.0	0	0	0.30	0.70	NA
Wageningen University	1	0	0	NA	NA	NA	0.46	0.54	NA
Other	0.53	0.42	0.05	0.33	0.67	0	0.47	0.53	NA

Table 4 Sample and population distributions of academic field

	Wave 1 (N = 590)	Wave 2 (N = 942)	Rathenau (fte = 27005)
Science/ Mathematics	0.35	0.36	0.33
Law & Legal Studies	0.03	0.08	0.05
Economics/ Business & Management	0.05	0.08	0.05
Humanities	0.12	0.19	0.07
Social Sciences	0.17	0.23	0.13
Clinical & Medical Sciences/ Health Studies	0.27	0.06	0.38

Table 5 Sample and population distributions of career stage⁹

	Wave 1 (N = 593)	Wave 2 (N = 946)	Rathenau (fte = 27005)
Full professor	0.09	0.12	0.11
Associate	0.07	0.09	0.09
Assistant	0.20	0.23	0.20
Postdoc	0.13	0.12	0.14
PhD	0.39	0.34	0.32
Retired still engaged in research	0.02	0.03	NA
Retired not engaged in research	0.01	0.01	NA
Other	0.10	0.06	0.14

Table 6 Sample and population distributions of age

	Wave 1 (N = 491)	Wave 2 (N = 713)	Rathenau (fte = 27005)
> 25	0.03	0.02	0.03
25-29	0.29	0.25	0.27
30-34	0.19	0.18	0.20
35-39	0.12	0.15	0.13
40-44	0.09	0.08	0.09
45-49	0.06	0.07	0.07
50-54	0.06	0.06	0.07

⁹ Rathenau Institute did not have a separate category for ‘postdoctoral researchers’, but only for ‘other scientific personnel’. That category contained 7518 FTE, and was divided 50-50% amongst our categories ‘postdoctoral researchers’ and ‘other’.

55-59	0.06	0.06	0.06
60-64	0.06	0.07	0.06
65-69	0.03	0.05	0.01
70 =<	0.01	0.02	NA

Acknowledgements

This dataset was funded under the NWO Corona Fast-Track Data scheme. Project id: 440.20.028; Project title: “Collecting systematic survey data on scientists’ information-seeking and information-spreading behaviour in a time of crisis”. The recruitment of respondents for the second survey would not have been successful without the careful work of Claudia Vlad, Daisy Goddard and Simone Astarita. We also thank Iro Fragkaki, Mrinalini Luthra, Marta Mangiarulo, Vera Toepoel and Marije Bomers for their helpful comments on earlier versions of the questionnaire. We also thank our communication department and colleagues for helping in distributing the survey, and Iris Proff and Arno Mul for designing the website for our project.

How to cite

Roex, K. & Colavizza, G. (2020). Doing Science in Times of Covid-19: A Survey. [Data set].

Zenodo. <http://doi.org/10.5281/zenodo.4313278> .