



seriss

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

Deliverable Number: D 7.7

Deliverable Title: Blueprint for comparative web panel

Work Package: WP7 A survey future online

Deliverable type: Report

Dissemination status: Public

Submitted by: ESS ERIC (HQ/City)

Authors: Curtis Jessop (HQ/City), Gianmaria Bottoni (HQ/City), Elena Sommer (Share ERIC), Elissa Sibley (HQ/City), Rory Fitzgerald (HQ/City)

Date Submitted: 3 September 2019

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 654221.





www.seriss.eu  @SERISS_EU

SERISS (Synergies for Europe's Research Infrastructures in the Social Sciences) aims to exploit synergies, foster collaboration and develop shared standards between Europe's social science infrastructures in order to better equip these infrastructures to play a major role in addressing Europe's grand societal challenges and ensure that European policymaking is built on a solid base of the highest-quality socio-economic evidence.

The four year project (2015-19) is a collaboration between the three leading European Research Infrastructures in the social sciences – the European Social Survey (ESS ERIC), the Survey of Health Ageing and Retirement in Europe (SHARE ERIC) and the Consortium of European Social Science Data Archives (CESSDA AS) – and organisations representing the Generations and Gender Programme (GGP), European Values Study (EVS) and the WageIndicator Survey.

Work focuses on three key areas: Addressing key challenges for cross-national data collection, breaking down barriers between social science infrastructures and embracing the future of the social sciences.

Please cite this deliverable as: Jessop C., Bottoni G., Sommer E., Sibley E. and Fitzgerald R. (2019). Blueprint for comparative web panel. Deliverable 7.7 of the SERISS project funded under the *European Union's Horizon 2020 research and innovation programme* GA No: 654221. Available at: www.seriss.eu/resources/deliverables

Contents

0. Executive summary	5
1. Introduction	6
Purpose of this blueprint	6
Building on the CRONOS Panel	6
Aims of a cross-national web panel	6
Content of the blueprint	6
2. The case for a cross-national web panel	6
Cross-European research	6
Benefits of web panel data collection	7
3. General principles and aims	9
Aiming at representative sample	9
Panel set-up	9
Cross-national set-up and comparability	9
Free and easy access to data and documentation	10
Cost, quality and quantity balance	10
4. Research focus	10
Cross-sectional and Longitudinal data collection	10
International data collection	11
Core & rotating modules	11
Selections of the topics	12
Methodological work	12
5. Designing and running a cross-national panel study	13
Sampling	13
The parent survey	13
Countries to be included	14
Relaxing requirements	14
Recruitment	14
Standardised protocols	14
Maximising recruitment rates and efficiency	15
Participant engagement post-recruitment	15
Questionnaire development	16
Scheduled and responsive questionnaire development	16
Web survey design	16
Translations	17
Sample management	17
Longitudinal sample	17
Live sample updates	17
Data security	18
Sample management systems	18
Fieldwork design	19
Mode	19
Fieldwork length	19

Frequency of contact	20
Panel longevity.....	20
Targeting fieldwork designs	21
Participant communications and engagement	21
Study branding and messaging	21
Translations.....	21
Mailings.....	22
Information sources.....	22
Incentives	23
Weighting	23
Ethical considerations.....	24
6. Organisational structure.....	24
7. Data outputs	26
General principles.....	26
Data included in outputs	27
Timing of releases.....	28
8. Impact and dissemination.....	28
Connections to parent study.....	28
Data archiving and documentation.....	28
Reporting.....	28
Core/repeat data collection.....	29
Ad-hoc data collection	29
Methodological work.....	29
Online impact and events & conferences	29
9. Costs	29
Types of costs incurred.....	30
Central costs.....	30
Local costs.....	30
10. Funding and governance	31
REFERENCES	32

0. Executive summary

This deliverable describes the design for a cross-national probability-based web panel recruited off the back of an established face-to-face survey. The document is based on our experiences in developing and running the world's first cross-national probability-based web panel, that is "The CROss-National Online Survey Panel" (CRONOS). CRONOS was piloted in three countries – Great Britain, Slovenia and Estonia – following up respondents to Round 8 of the European Social Survey. CRONOS fieldwork started in February 2017 and continued for 12 months. The CRONOS project had the following objectives:

- Evaluate the feasibility of establishing the first cross-national probability-based web panel using the achieved sample from existing cross-sectional surveys
- Foreground a methodology for building new and efficient web-based survey infrastructures for Europe based on state of the art procedures and technology
- Develop a blueprint for a comparative probability-based web survey

Following the successful implementation of CRONOS, this document outlines a design for a cross-national web panel, using CRONOS evidence and accepted best practice, taking into account the challenges encountered during the project, and suggesting how to setup and implement a web panel considering all the related issues.

1. Introduction

Purpose of this blueprint

This Blueprint is for researchers looking to conduct cross-national, probability-based sample surveys. It provides an outline design for a cross-national web panel, indicating things to consider and our recommended approach.

This blueprint is written with reference to the parent survey, the European Social Survey (ESS), and the follow-up web panel – CRONOS. However all the principles outlined in this document can be applied equally to other surveys.

More specifically, the blueprint is for a cross-national web panel in Europe recruited on the back of a face-to-face survey on general topics related to attitudes and behaviours relevant for the context (Europe) that aims to foster comparative research.

Building on the CRONOS Panel

The design primarily builds on our experiences developing and running the CRONOS¹ web panel (Villar et al. 2018; Villar and Sommer 2017; Bottoni and Sommer 2019), which in turn is based on evidence from national probability-based web panels (Sommer 2017).

Following the successful implementation of the CRONOS Panel, this blueprint builds on that design, using evidence and our professional judgement, taking into account the challenges encountered, and considering how we might do things differently if we were to run a similar study in the future.

Aims of a cross-national web panel

The immediate goal of developing a cross-national web panel following this approach is not to replace face-to-face surveys, but complement them. Although a replacement may happen in the future, face-to-face interviewing is still the ‘gold standard’ in survey data collection, in particular due to its typically higher response rates and coverage of the offline population. In addition, more research needs to be done to evaluate the quality of data collected via online surveys particularly in respect to representativeness.

Content of the blueprint

A key challenge for the design of a web panel is to balance the aims of conducting surveys more quickly and at lower cost while ensuring that data quality is not undermined.

After outlining the potential benefits and suggested purpose of a cross-national web panel, this blueprint will focus on the practical design decisions, which relate to this balance such as whether or not to include the offline population, incentive, and communication strategies.

2. The case for a cross-national web panel

Cross-European research

Over the past 10 years, several national probability-based panels have been established in Europe, Australia and the USA. However, these studies are designed to deal with specific national aims and needs and the different designs mean it is problematic to combine them to be employed for cross-

¹ More information about the CRONOS Panel can be found at https://www.europeansocialsurvey.org/methodology/methodological_research/modes_of_data_collection/cronos.html, and at <https://seriss.eu/>.
www.seriss.eu

national comparative research. With the exception of the world's first cross-national probability-based web panel (CRONOS), which worked as proof of concept for the feasibility of an European web panel, there are still no cross-national web panels capable of providing an international and/or comparative view on socio-economic phenomena and a comparative perspective on social phenomena.

As Durkheim stated: "Comparative sociology is not a particular branch of sociology: it is sociology itself" (1964, p.139). European countries show enormous cultural, socio-economical and constitutional/institutional variation. This diversity gives unique opportunities to study the interaction between individuals' attitudes and behaviours and their national structural environment. In addition, in Europe, where socio political turbulence is challenging and questioning the process of integration, and the key issues (e.g. climate change, large-scale immigration) can be addressed only in a combined and mutual manner, there is a strong demand for European-wide measurement of citizens' attitudes, beliefs and orientations.

This diversity cannot be investigated without a systematic, continuous, and harmonised study to ensure comparability of the different measures collected. By implementing an input-harmonised approach (i.e. central management with consistent recruitment, setup and maintenance) a cross-national web panel can facilitate functional equivalence, a necessary condition to conduct high-quality cross-national/cross-cultural research (Survey Research Center 2016).

Establishing a cross-national European web panel therefore would provide the scientific community with additional opportunities compared to existing national panels that already exist, providing a tool to collect frequently cross-national data on the most relevant societal issues.

Benefits of web panel data collection

A number of input-harmonised cross-national studies already exist (for example the European Social Survey). However, these are mostly conducted using face-to-face interviewers - a costly and time-consuming (although high-quality) approach. This Blueprint outlines a web-panel approach to data collection which has several characteristics that makes it an attractive alternative to classical modes of data collection.

Firstly, the web panel approach means that questionnaires can be set up and delivered in a relatively short time as there is not interviewer capacity limit, or contact requirement for fieldwork progress (Evans and Mathur 2005; Sue and Ritter 2012). This can offer a number of benefits:

- Short-term change: A shorter fieldwork period allows the more frequent collection of data, allowing researchers to detect changes in the short term (adding this perspective to the long-term change that generally is the focus of face-to-face surveys like the European Social Survey).
- Responsiveness: A shorter set-up and fieldwork time means the survey will be able to respond more quickly to events (social, political, economic) in a specific country or across Europe employing a comparative perspective. This allows researchers and policymakers to quickly collect and analyse data to promptly address issues in societies.

Whilst web survey do offer greater speed, it should be noted that cross-national web data collection is still less swift than a single national survey. Questionnaire design, translation and data processing take longer in a cross-national environment and the availability of staff across multiple countries for implementation also needs to be considered.

Secondly, web panel surveys potentially offer cost savings if they are built onto existing data collection (Schmidt 1997). As there are no costs for interviewers, survey fieldwork costs can be relatively low, and having a named and engaged sample improves response rates while using a piggy-back approach means the recruitment fieldwork costs are entirely borne by the parent study. These savings are also highly scalable – the larger the sample size, the greater the savings. This is particularly beneficial where the study is being run across multiple countries. This lower cost reduces the barriers to survey data collection, allowing for a greater volume of survey data to be conducted to cover new topics or explore others in more detail, when previously there may have been insufficient funding to do so.

In addition, using a web panel design allows for longitudinal, as well as cross-sectional data collection. Surveying the same panel members over time allows for analysis of individual-level change, helping to understand whether or not groups of people and individuals are changing over time.

Finally, by using web data collection, there may be some benefits to data quality. As a self-completion mode, data can be more accurate due to the absence of interviewer effects (respondents may feel more comfortable providing honest and more objective answers). In addition, web surveys may make it easier to implement more interactive forms of questions, taking advantage of multimedia functionality.

Conducting a follow-up web panel recruited off the back of an existing face-to-face survey increases the value of the parent survey in addition to being a benefit in its own right. Indeed, operating a follow-up web panel would provide at least three main connected advantages to the parent survey. A web follow-up survey will provide additional questionnaire space. It would allow the possibility for the face-to-face parent survey to field more items more frequently, protecting and enhancing its time-series and creating additional possibilities. Another benefit is related to the value that the web follow-up survey adds to the parent survey in terms of faster responsiveness to important events within a country or across Europe. This is particularly important if we consider that modern societies develop in a faster way compared to the older ones and, especially in the era of the social media and news on demand, attitudes and behaviours change quicker and new ones arise. Finally, the web panel would complement the face-to-face parent survey with the addition of a longitudinal component. While implementing a panel approach in face-to-face surveys is extremely complex and requires increasing resources as the panel proceed over the years, a web panel would make the implementation of the longitudinal element more feasible and easier in terms of resources.

However, there are also challenges related to web panel surveys (Couper 2000). While the lack of interviewer may reduce social desirability biases, it may make the administration of certain types of data collection (e.g. cognitive tests, bio-measures, data linkage consent requests, collection of text data) more problematic. While the longitudinal nature may offer analytical benefits, conditioning effects may affect panel members' answers (in particular knowledge/awareness questions). Perhaps most importantly, some researchers are cautious about the representativeness of web panel surveys (Callegaro et al. 2014). Representing the general population poses several problems when trying to do so via web surveys since there is no country where the whole population has internet access, and internet penetration varies between European countries. In addition, the lack of an interviewer may reduce response rates relative to interviewer-led approaches, while longitudinal samples are subject to attrition and systematic aging, potentially introducing additional bias into the sample.

3. General principles and aims

This section presents the five general principles of a cross-national online panel recruited off the back of an established cross-national probability-based face-to-face survey.

Aiming at representative sample

Ideally, a representative sample of general population should be achieved for each participating country in an online panel. The piggy-backing approach makes use of the probability-based methods of the parent survey to select the sample. However, one point of major considerations is the coverage of the offline population that might affect the representativeness of the online panel sample. A careful decision should be taken on how to deal with this issue to try to achieve a representative sample in countries with a relatively high offline population (e.g. setting an upper age limit or covering the offline population with other modes). In addition, weights for the panel data should be produced and made publicly available and panel representativeness should be regularly assessed against national benchmarks. Another challenge is achieving a large enough sample for advanced statistical analyses. To increase the initial panel sample size, samples of new panellists should be added after each new wave of the parent survey. In addition, panel maintenance strategies targeting at lowering panel attrition are crucial for sample representativeness.

Panel set-up

An online panel, being a flexible data collection vehicle, allows the possibility to react to important political and societal events on a more timely basis than is currently possible in most cross-national face-to-face surveys. In the long-term perspective, an online panel would offer a possibility for more frequent and timely data collection as compared to face-to-face mode. The aim of a CRONOS style online panel is, however, not to replace the parent face-to-face study but to complement it with an additional flexible data collection option that provides the potential for cross-national longitudinal research and new types of analyses at relatively low cost. To achieve the advantages associated with longitudinal research, the online panel should be designed and implemented as a long-term endeavour. The long-term use can also justify the relatively high investment for setting-up the online panel infrastructure at the start.

Cross-national set-up and comparability

The main innovation of the approach presented in this document is the cross-national input-harmonised framework for such a panel. A critical minimum mass of countries (4-6 countries) would be required to facilitate meaningful substantive comparisons. The countries participating in the online panel need to achieve a sufficient response rate (and a large enough net sample size) in the parent survey to enable a large enough online panel sample. The effective sample size also needs to be considered. Further, a long-term commitment to participate in the parent survey and the online panel are essential for planning. Participating in the follow-up online panel might be particularly attractive for countries where national probability-based academic online panels do not exist yet. However, even in the case where existing national online panels are in place, being included in comparative data analyses increases the value of the panel for all participating countries.

Cross-national comparability can be maximised through using an input-harmonisation approach reflected in the standardised recruitment in all participating countries as well as a centralised online panel management including centralised tools for sample administration and data collection. In addition, building on the input harmonised parent study provides comparable demographic information that itself was collected using comparable approaches. Further, comparability can be

achieved through equivalence whereby exactly the same questions should be asked in all participating countries (with exception of country-specific modules) and the quality of translation and transferability of concepts in the wider international context should be evaluated by questionnaire and translation experts before finalising the surveys. As far as possible and sensible, all participating countries should also adhere to standardised panel maintenance procedures (e.g. in terms of incentives and the type and content of materials for panellists) and use the same procedures for online and offline communication with panellists. Where standardisation is not possible, the chosen optimisation for local procedures should not add complexity to the central infrastructure and deviations should be documented.

Free and easy access to data and documentation

The data (including survey data, paradata, and administrative data), documentation, and weights should be made available free of charge for non-commercial users. Data access should be easy for potential users, including data users outside academia (e.g. policy makers or journalists). The data should be archived following the FAIR principles (meaning that data are findable, accessible, interoperable, and re-usable). The online panel data need to be linked to the interview data of the parent survey. However care should be taken to ensure that the disclosure risk does not increase as data is added with later waves. Further, to make the panel data attractive to the wider research community and to increase its impact, there should be open calls for researchers to submit questions to be considered for inclusion in the panel free of charge.

Cost, quality and quantity balance

A CRONOS style online panel should aim at achieving cost, quality and quantity balance. Thus, the potential coverage issues and smaller sample sizes as compared to the parent survey are counter-balanced by opportunities of speedy and more frequent data collection at relatively low cost as well as opportunities for longitudinal research. To optimise quality standards, collaboration with different experts and regular quality assessment of collected data are recommended. The decision on frequency and the length of the online surveys (quantity of collected data) depends on a balanced combination of available budget and infrastructure capacities, the needs of the scientific community and policy makers, as well as the burden for panellists.

4. Research focus

The world's first cross-national probability-based web panel (CRONOS) demonstrated the feasibility of recruiting a panel off the back of an existing face-to-face survey in terms of costs, representativeness, response rate and data quality although the final sample sizes were rather small (Villar et al. 2018). Having demonstrated the feasibility of a cross-national web panel, the research focus of future CRONOS style panels should be on enhancing and developing further European comparative research by facilitating additional data collection on substantive topics. In other words, the focus should not be on methodological research as much as in CRONOS-1 but rather focused on the substantive contribution such a panel could make.

Cross-sectional and Longitudinal data collection

The general topic of a future CRONOS style cross-national web panel would be the study of social, political, and cultural beliefs, attitudes and behaviours, their differences across countries and changes in the distributions across time as well as the explanation of those attitudes and behaviours (and their changes) and related social phenomena.

This aim will be achieved through two forms of data collection:

- Cross-sectional: the additional data collection will provide more space to:
 - Field items/topics left out from the parent survey
 - Explore a topic fielded in the parent survey in more detail
 - Repeat modules that were fielded in previous parent survey rounds (protecting and enhancing parent survey time-series)
 - Field new items/modules
 - Foster collaboration with other European or non-European surveys (e.g. fielding same items)
- Longitudinal: this will complement the face-to-face parent survey with a longitudinal component:
 - Fielding a longitudinal rotating core module to measure attitudes as frequently as expected to change (however, this will add complexity since the measures cannot be assumed to be independent anymore – correlated residual errors)
 - Collecting event history data (same complexity as above)

The longitudinal measures would represent an additional value since it is not just having more data but, if those measure are attached to cross-sectional data collected face-to-face, this will facilitate data analysis not previously possible and support more detailed substantive analyses (e.g. following-up same respondents to study how an event impacted on their attitudes across time).

International data collection

As previously reported, the focus of the cross-national web panel will be the study of attitudes and behaviours across European country in a comparative perspective. Therefore, the main part of the panel will be devoted to collecting measures that have relevance at European level.

This said, one of the advantages of setting up a panel study is also the possibility of fielding country-specific waves. Indeed, one wave per year could be dedicated to field country-specific surveys. This would be of particular value in those countries where there are no national panels, providing them with an opportunity to field national-relevant surveys using a cost and time-effective, but high quality, approach.

Core & rotating modules

The content for each wave should consist of different combinations of the following:

- Core modules: these will include questions and topics concerning the main and general themes of the panel (i.e. study of social, political, and cultural beliefs, attitudes and behaviours as well as differences across countries and changes across time). These should be pre-determined and scheduled, repeating as appropriate for the topic.
- Demographic section: even though the basic demographics are already measured in the parent survey and therefore available for the web panel, some demographics should be collected every 6 months in order to study changes and ensure they are not out of date.
- Ad-hoc data collection modules: this content will be less planned than core content, and includes space for external applications with researchers invited to submit their proposals to include topics and indicators in a wave, as well as for the study to carry content in response to emerging international events.

Selections of the topics

The selection of the topics for the cross-national web panel should be performed taking into account its primary objective of developing and furthering the comparative research across countries of interest. The topics selected must be:

- Theoretically driven: survey instruments and indicators need to be designed to take account of theoretical arguments since bespoke data collection using surveys is particularly well suited to answering such questions
- Based on existing successful measurement and empirical evidence (previous studies – not necessarily in a comparative setting, pre-testing), although new instruments should be considered as well in order to avoid a conservative bias
- Relevant for the general theme of the cross-national panel (comparative research in an international setting) with particular reference to relevance to all countries in the panel.

Areas of substantive interest

The selection of topics should be based firstly on the consideration of including that specific topic to foster substantive research in a European comparative setting. In doing that, the topics should therefore be focused on elements that will help in understanding social, political, cultural and structural changes across the countries in the panel.

A further element to take into account refers to the responsiveness toward events in Europe/World. Indeed, a web panel is much more flexible compared to a face-to-face survey. For this reason, the panel should consider fielding questions regarding events happening in a society enhancing and complementing measures in the parent survey.

Methodological work

Although methodological research should not be the primary focus of the panel, web surveys do provide more flexibility compared to face-to-face surveys for administering more complex experimental designs. For this reason, a small section of the panel questionnaire space (e.g. a secondary focus) could work as a platform for methodological experiments. For example, the panel could work as a platform to pre-test items that will be eventually fielded in the parent survey.

In addition, the panel could also work to improve web survey implementation, experimenting with different contact modes, question wording, translation approaches and, more in general, implementing different randomised experiments concerning respondent behaviour. This should be beneficial for the study itself as well as wider literature as findings from these experiments can then be implemented as part of the study's core design and inform other similar studies.

Finally, the panel could work as an experimental platform to use new technology and allow the collection of new and alternative forms of data. These could include:

- Administrative or social media data linkage
- Self-admin bio measures
- Geo-demographics
- Diary data

However, we should consider that the main aim of the panel is to provide the scientific community with data collected for general use, not for specific clients or a subset of particularly statistically

literate researchers. For this reason, everything that adds complexity (e.g. autocorrelation of the measures) should be carefully evaluated.

5. Designing and running a cross-national panel study

Sampling

The parent survey

As the research focus of this cross-national panel study will be to allow for comparison across countries, it is important that as much as possible the sampling approach is equivalent across the countries included in the panel. For a scientifically oriented study the design should also be based on probability sampling methods.

This blueprint is for a panel recruited using a ‘piggy-back’ approach; that is to say panel members are recruited from an existing cross-national survey infrastructure. The sampling frame for a cross-national panel study that uses a piggy-back approach to recruitment is defined by the sampling frame used for the ‘parent’ study. It is therefore important that the parent study selected also meets the criteria for the panel study. In this blueprint, the goal is to have a panel that is representative of the general ‘adult’² population of a country, recruited using random probability sampling, so the parent study selected should also use this approach.

Beyond this, it is also important that equivalent sampling approaches are used in all countries included, using the best methods available for the country - population registers where they are available and address-based sampling where they are not. With the development of appropriate protocols and training, this would allow for the consistent application of specific sampling criteria across countries (e.g. the recruitment of an entire household or a random individual, or the inclusion/exclusion of institutions). SHARE, ESS and EVS have detailed protocols and approaches that can be drawn on to try to achieve the equivalence in sampling design.

An additional consideration in selecting the parent study is the sample size it would provide to the panel. Based on the findings from the CRONOS study (Villar et al. 2018), approximately one in three participants in the parent study, or one in seven of all eligible people sampled at the start of the sampling for the parent study, will end up taking part in a given wave of the panel study (though this varies by country). The achieved sample size for a given Panel survey should be sufficiently large for the effective statistical analysis of country-level data (or higher if sub-group analysis is required). Assuming a target of 800 complete panel interviews per country per wave, this would require that the parent study has an eligible issued sample size of c.4,000 to 5,500 per country, though this may vary for countries with particularly high or low response rates³. It may be that to achieve this the Panel may need to be recruited from multiple waves of the parent study.

² We would recommend defining this as people aged 18+ to eschew the ethical and practical challenges of asking for parental consent for those aged 15 to 17. Some web panels put upper age limits on their panels (e.g. GESIS: 70, and GIP/ELIPSS: 75), as low internet penetration amongst older people raise concerns about how well they can be represented. However, we do not recommend this for this study as the focus is the population as a whole, and it would be preferable to exclude this group from analysis should there be particular concerns, rather than from the data collection.

³ These will likely vary, but figures are based on the overall response rates of 15% to 22% seen in CRONOS.

In general the parent study design would not be influenced by the web panel follow-up. However, consideration might be given to oversampling groups known to have higher non response to the parent study or who are known to be more likely to attrit once the panel operated.

Countries to be included

The countries to be included in the sample for the panel study should firstly be determined by the research objectives of the study. In this instance the panel study aims to act as a general social survey for international social science research rather than answering very specific research questions, and as such it should aim to include as wide a range of countries as possible, to maximise its potential utility.

However, in order for a country to be eligible to be a member of the panel study, it must also already be part of the parent study in order to ensure that its recruitment protocols are consistent with the other countries taking part in the panel study. They should also commit to the longer-term support for the online panel, as it is over time that it offers the greatest value.

In addition to this, as an online panel, that country should have a minimum level of internet penetration as those without internet access will not be able to participate, and should this be a large or very distinct proportion of the population then this may significantly bias the sample. The precise cut-off point is difficult to estimate, and should depend on the extent to which those with and without access to the internet differ in terms of their responses to variables of interest, which is often unknown.

Relaxing requirements

The use of a piggy-back recruitment approach as outlined above offers many advantages, most notably substantial cost savings in accessing a high-quality, consistently-recruited, probability-based sample⁴. However, it also creates limitations – finding a parent study that meets the precise requirements may be difficult. As such, those requirements may need to be prioritised and relaxed. The extent of that relaxing should also be balanced against the benefits of using a piggy-back approach compared to alternative approaches such as recruitment from a fresh sample.

Recruitment

Standardised protocols

In order to maintain the principle of a random probability sample design, all eligible people who take part in the parent survey should be invited to join the online panel study. This should include those who report not personally having access to the internet as they may gain access in the future, or be able to access the internet through alternative means (e.g. a public library). It is up for each survey to decide whether they want to invite those who do not participate in the parent study. However, it should be taken into account that this group would be missing the valuable profiling information from the parent study (potentially including the random selection of an individual), and would be unlikely to elicit high take-up and therefore substantially improve the sample quality (see also Villar and Sommer 2017). At the moment, there is no research available about the inclusion of panellists who did not participate in the parent survey.

As with the sampling and recruitment for the parent study, the recruitment approach to the panel should be standardised across countries as much as possible, developed centrally in collaboration with

⁴ See, for example, experimentation with the GESIS Panel refreshment (Schaurer et al, 2019).

local teams. This should cover the briefing of interviewers, the materials they present to participants, and the wording of the recruitment question. All three should aim to maximise levels of informed consent by addressing any concerns and answering any questions participants might have. The approach should also be compliant with GDPR (which may vary depending on the project's legal basis for processing data), and confirm the contact details of those that agree to join the panel.

Maximising recruitment rates and efficiency

As a random-probability sample, the quality of the sample recruited to the panel should be optimised by maximising the recruitment rates⁵. However, increasing the proportion (and number) of people that are recruited to the panel also increases the costs of running the panel, as more resources are spent on (for example) sending letters or providing unconditional incentives. The level of resources placed into recruitment to the panel should therefore also consider the marginal costs of an additional panel member, and be weighed against the impact on the overall response rate, rather than the recruitment rate⁶. If the marginal costs of panel size are high (for example, where unconditional incentives are used), effort should be focused on maximising response rates among panel members, rather than at recruitment, and costly protocols such as additional incentives for interviewers or participants should be avoided, in order to maintain the cost-efficiency of the panel design.

In addition, some panels (e.g. PEW American Trends Panel in the USA) place greater recruitment efforts on target groups known to have lower response rates. At the recruitment stage efforts to incentivise respondents from groups known to be less likely to join the panel and/or are more likely to attrit should be considered. Example would be interviewer incentives for recruitment or higher offered incentives for participation. In addition, the messaging at the recruitment stage could be tailored.

Participant engagement post-recruitment

Face-to-face survey fieldwork can take several months to complete depending on the scale of the survey and the capacity of the organisation conducting the fieldwork. As a result, panel members recruited early in the parent survey's fieldwork period may have a long waiting period before the first panel survey is administered (once all panel members have been recruited, to ensure representativeness). To maintain engagement for this group, we recommend using a 'welcome mailing' (preferably sent by post and email) to provide panel members with additional information about their participation and confirm their joining, and a 'welcome survey' run in parallel to the parent study to keep panel recruits engaged and help accustom them to the process of accessing a questionnaire as part of the panel. These goals should be reflected in the questionnaire content (focusing on 'engaging' content, and checking/collecting information that may be of use for running the panel). As the survey will not be administered to all panel members at the same time, it may not be appropriate for substantive research. We recommend that the mailing and survey are timed to minimise the gap between contacts, with multiple batches sent and the welcome mailing sent c.1/3 of time between recruitment and Wave 1, and the welcome survey sent 2/3 of the time between recruitment and Wave 1.

⁵ Much of the overall non-response to a panel survey will be from non-response to the parent survey as well. We assume that response-maximisation strategies are already in place for the parent study, however, moving some resources aimed at response-maximisation for the panel and employing them during recruitment to the parent survey may be mutually beneficial depending on their relative marginal impact.

⁶ Evidence from the NatCen Panel (Jessop, 2018) showed that a 65% higher recruitment rate was reduced to a 20%-30% (or three to four percentage point) difference in overall response rates.

Once the panel is operating efforts to incentivise respondents from groups known to be more likely to attrit should be considered. Example would be higher incentives for participation. In addition, the messaging and contact procedures could be tailored.

Questionnaire development

Scheduled and responsive questionnaire development

One of the goals of developing a cross-national web panel is to take advantage of its ability to deliver high-quality data outputs more quickly and at lower cost than traditional probability-based methods which rely on face-to-face data collection. As well as applying this principle in the fieldwork design, it should also be reflected in the questionnaire development stage.

The majority of questionnaire content administered via the web panel would be pre-planned, and should be carefully developed to take account of the additional challenges of cross-national measurement. However, the level of resources used for development and testing should be proportionate to effort put into the mainstage fieldwork. For example, telephone or online cognitive testing, 'hall testing' with a convenience sample, or a systematic desk review may be more cost- and time- appropriate than face-to-face cognitive interviewing. However, more 'thorough' approaches may be appropriate where, for example, the questions are planned to be repeated across multiple panel waves, or where the questions are completely new (as opposed to existing questions being adapted for the online mode).

Where a survey is fielded in reaction to unpredicted events, and the purpose is to collect and report high quality data more rapidly, the questionnaire development may need to be abridged further in order for the data to be produced in a timescale that is still relevant. However, there should always be cross-national expert review and advance translation efforts⁷ at a minimum before questions are fielded. Ensuring there is an infrastructure for this task is an essential part of building such a panel.

Web survey design

A key focus of the questionnaire development should be making the questionnaire mode-appropriate. While some principles remain consistent with other self-completion and interviewer-administered modes of survey data collection, web surveys are different in that different participants, using different devices to complete the survey, may experience the questionnaire differently. Evidence in CRONOS suggests that on average 40% of panel members in Great Britain, 35% in Slovenia and 25% in Estonia completed a survey via a smartphone and this is likely to rise in future. The questionnaires should be designed to be 'smartphone-friendly', ensuring that, when viewed in a smartphone browser, questions are still accessible.

The mode of interview also affects the appropriateness of length of interview. Shorter questionnaires are recommended for web surveys (relative to face-to-face interviews), with guidelines recommending limits of c.15 to 20 minutes to ensure that participants are willing to participate and continue to engage with the survey, and ensure data quality of later questions and later waves is not negatively affected (Callegaro at al. 2015). However, panel samples tend to be more engaged and there is some evidence that longer interviews are not so problematic for future participation (Lynn, 2014). In that context, many national web panels in Europe typically have interview lengths of c.20 to

⁷ Regarding advance translation efforts, for example if there are 25 countries in the study, a country can translate the pilot version in advance of all countries doing their translation. This can help to identify items that might be difficult to translate in advance.

30 minutes (Blom et al, 2016). Irrespective, web surveys' fixed costs are substantially lower than face-to-face surveys, and as such an optimal design would take advantage of this to ensure any survey is no longer than necessary to maximise data quality.

Translations

A key challenge for the questionnaire design for a cross-national web panel will be the translation of questions. As with any international research study, consideration must be given that literal translation may not equate to equivalent meaning, as words may have culturally-specific connotations, and different cultures may (for example) use similar scales differently. These variations should also be considered specifically in the web-survey context, which may be different to face-to-face or telephone surveys.

The development of equivalent measures across countries can require significant questionnaire development and testing. As noted above, this should be balanced against the context of aiming to produce research outputs that are timely and cost-effective.

Sample management

Longitudinal sample

Despite the research focus likely to be mostly cross-sectional, the sample management processes required for managing a web panel are most similar to a longitudinal study as surveys are administered to the same people over time. The key feature of this is that sample information will need to be collected and maintained over time in order to ensure the effective operation of the study. For example, participants' contact details (e.g. email address) may change, and this information should be tracked to ensure that future communications (e.g. invitations to take part in a survey) reach the correct person. Alternatively, someone may wish to withdraw from the study or may become ineligible, in which case their records will need to be updated to ensure they do not receive further contact.

There can be additional requirements – for example, panel studies offer the opportunity for 'dependent interviewing', where previously held information about the panel member is used to personalise a questionnaire script. For example, to reduce participant burden/improve data quality, they may be initially asked if their labour market status has changed, rather than re-asking the question. The presence of large volumes of historic data allow for many opportunities for targeted approaches to fieldwork and data collection, but for these to work, the relevant data need to be effectively managed and fed through from wave to wave. We would recommend that a panel is designed to take advantage of this.

Live sample updates

The relatively fast-paced nature of online panel research means that the sample management systems and processes must allow for data to be as close to 'live' as possible. For example, when reminder emails are sent out, it is important that the sample data is based on up-to-date information on who has or has not completed the survey already to avoid sending reminders to people who have already completed the survey. In addition, the study processes and systems needs to be responsive to demands that may be made under the GDPR.

Data security

The management of international sets of data, which include potentially sensitive and (indirectly) identifiable information about people, requires careful consideration of the data security of its management – both from a legal and ethical perspective.

As much as is possible, appropriate and legally required should be done to maintain the security of participants' information, applying principles of good data management, for example:

- The systematic and considered management and processing of data, ensuring that what data are held and where they are stored is always thought through and documented, separating out data which need to be maintained as part of the 'sample' to run future waves (e.g. contact details, or substantive variables needed for routing), and cross-sectional data which are fixed and not required (e.g. substantive or paradata variables that are not needed at future waves).
- Data reduction should be practised to minimise the data held about an individual in one location to what is necessary for the required processes, such that the risk of disclosure is minimised and, should disclosure occur, the risk of harm is minimised (i.e. so variables are less likely to be disclosive in combination and, if a panel member is identified, the amount of information that can be found out about them is minimised).
- The control and documentation of access to data, to ensure that only those who need, and are trained, to access identifying information are able to do so. For example, there is no need for the national team responsible for managing the sample for one country to have access to the sample information of panel members from another country.
- As a panel study, the project may be open-ended, in which case the storage of sample information may be justified indefinitely. However, it may also be the case that panel members will only be included for a certain amount of time. Sample management processes should be clear regarding what data and variables are to be stored, and for how long. Even if panel members are kept in the study indefinitely, specific sample files (such as those used to create letters) do not need to be maintained. For such files, any required paradata should be extracted and the files deleted securely and promptly as soon as possible.

The international nature of the study may also add complications to the legal requirements for sample management. Depending on the countries involved, there may be specific requirements related to the transfer and storage of panel members' information in different countries. For a European panel it is optimal that the personal data be stored and processed centrally in an archive located in an EU or EEA country.

Sample management systems

The longitudinal management of a cross-national web panel sample, accounting for updates from multiple sources as quickly as possible and producing multiple output files for different purposes, will require technical solutions that are not typically covered by 'off the shelf' sample management systems. Given the importance of sample management to the successful and secure running of a web panel, we would recommend that the development of a bespoke sample management system would be a worthwhile investment of resources.⁸ This should be able to securely hold and update the information required to run surveys, including contact details and fed-forward information for surveys. It should be responsive, with data as close to 'live' as possible to ensure contacts are made with the correct people at the correct address; integrated, with data flowing easily between the

⁸ A cross-national sample management system is being developed under a new project, Social Sciences and Humanities Open Cloud (SSHOC; GA No 823782).

sample management system and others (e.g. the survey or email communications software); and secure, with only necessary personal data held, and it only made available to those who need it to run the systems.

Fieldwork design

Mode

Implementing a mixed-mode design through the inclusion of an 'offline' mode such as paper self-completion or telephone/face-to-face interviewer-administered surveys, or the provision of web-enabled devices can be relatively costly and/or time-consuming to implement. Doing so would therefore risk undermining the goal of the web panel to collect data in a quicker and more cost-effective manner. In addition, it risks introducing mode effects into the data which may undo any benefits of improved sample representativeness, and this blueprint therefore outlines a single-mode (web) design.

However, the appropriateness of this approach will depend on the specific context of the study – for example, the budget and time available may make a mixed-mode approach more feasible, or the research questions and sub-populations of interest may mean that coverage of those without internet access is more important. Currently in Europe a mixed mode design would be required in order to maximise representativeness in most if not all countries: while some countries in Europe have most of their population reporting having used the internet (e.g. Iceland and Netherlands at 96%), many still have significant minorities that report having not (e.g. Germany: 17%, Spain: 21%, Hungary: 28%)⁹. If those who do not use the internet are excluded from the study, and they are different to the rest of the population in terms of the measures of interest, this may bias the survey results.

Fieldwork length

Web surveys do not require the same length of time to field as interviewer-administered approaches, as there is no field-force capacity issues to consider. However, a longer fieldwork length may still be desirable as it should allow more participants, who may be busy or unavailable, the opportunity to take part and therefore potentially increasing the sample's representativeness. It should also reduce the risk of one country's results being affected by specific national events (e.g. holidays). However, this should be balanced against the goals of fast turnaround fieldwork, and that by extending fieldwork, those answering later may be doing so in a different context to those taking part sooner.

In general, we would recommend a 4-week fieldwork period, but as above this will depend on the research objectives of that particular wave – if data are required more urgently, then it may make sense to reduce the fieldwork length further .

The empirical evidence based on Cronos showed that in the last wave 74.6% of panellists responded to the survey within a week (30% of them responded the day the invitation was sent), 12.9% of panellists the second week, 6.2% the third week, 3.1% the fourth week, 2.5% the fifth week, and 0.7% the sixth week. In the same way, in wave 2 (the first with this data available due to a data loss in wave 1, see user guide Villar et al. 2018) there is a similar response structure. Indeed, the 77% of the panellists responded within the first week.

⁹ Source: ESS Wave 8
www.seriss.eu

Frequency of contact

Increasing the number of surveys administered to panel members through higher frequency contact can increase the value-for-money of a web-panel as fixed costs are spread over more waves. However, higher levels of contact may be excessively burdensome on panel members, resulting in higher levels of attrition from (or refusal to join) the panel, or lower quality answers. Given the surveys are incentivised, there are also ethical (or legal) implications in especially high volumes of surveys. Conversely, low levels of contact may result in disengagement.

There are also logistical considerations for survey frequency. While it is possible to administer two surveys to the same panel member simultaneously, it is operationally riskier to do so and may be more confusing or burdensome for panel members (considering they may be receiving reminder communications for both at the same time). We would therefore recommend avoiding overlapping fieldwork periods where possible.

With a c.4-week fieldwork period, this would put an upper limit of c. 12 waves per year. However, it should be considered that the upper limit of 12 waves per year would be particularly resource-intensive in a cross-national context and with academic teams running the study. Instead, falling below 4 waves per year, with c. 3 months between surveys, may risk disengagement, and reduce the cost efficiencies of setting up a 'panel' approach as opposed to an ad-hoc follow-up study. The CRONOS Panel ran 6 waves per year with no evidence of negative impact on data quality, and other national studies run with similar frequency (e.g. the NatCen Panel runs 6 to 8 per year).

Within these bounds of 4 to 12 waves per year, the frequency of data collection should be dictated by the study requirements and available resources: What is the demand for 'core' or ad-hoc measures? How frequently do they need to be collected to address research questions? How much space should be allowed for 'responsive' data collection? Is there a requirement to conduct 'national' waves?

Panel longevity

The planned panel life should be defined up-front and it should be made clear to panel members whether they will be asked to take part in surveys for a fixed period of time (and if so, for how long), or indefinitely.

Over time, panel members may become increasingly unlike the population they were supposed to represent. Recruited at a single point in time, panel members are a representative cohort and will become relatively old compared to the target population over time - after one year, a sample that was initially representative of people aged 18+ will effectively be a representative sample of people aged 19+. In addition, attrition from the Panel over time (as panel members request to leave) will potentially negatively affect sample sizes and representativeness. Panel members may also experience 'conditioning' effects, where the process of participating in the panel and completing surveys may affect their answers, for example by making them consider issues more and therefore changing their opinions, or by learning survey question styles. Finally, the level of burden placed on panel members should be considered. If over-burdened, this may affect their answers and participation, but it should also be considered from an ethical perspective.

These issues would suggest that it is preferable to keep panel life relatively short. However, that conclusion should be placed into context. While 'aging' effects are important, they may not have a substantial impact at the population level, but be more of a concern for particular sub-group analysis. Attrition in web-based panels is relatively low as there is less direct contact with the panel members

(and therefore fewer opportunities to request to leave), with response rates declining very slowly and therefore being unlikely to have a substantial impact on the sample profile. Conditioning effects may not be substantial, although repeating ‘awareness’ measures, for example, should be avoided. While participant burden should be considered, the survey frequency we propose is relatively light (compared to, for example, commercial web panels) and burden is arguably lower than studies that take multiple hours at a single wave, and participation would always be voluntary.

In addition, there are benefits to a greater panel longevity – indeed the cohort would need to be maintained for the length of any longitudinal projects. The longer the panel sample is maintained for, the less requirement there is for fresh recruitment from the parent study or (if the parent study is repeated regularly) the greater the opportunity to expand the size of the panel.

In practice, panel longevity will be constrained by the scale and frequency of the parent study and the research requirements. A frequent or large parent study, with only short-term longitudinal research questions can have a short longevity, but a small or infrequent parent study, and/or one with requirements for longer-term longitudinal analysis would need to maintain the sample for longer. We would recommend that any panel study is set up to be maintained indefinitely rather than with a pre-determined end point. Even if initial requirements may suggest that this would not be necessary, it allows flexibility should requirements change and avoids the negative consequences that would flow from having to ask for additional consent. ‘Standard’ cross-sectional fieldwork may be conducted with the most recently recruited panel members to minimise any negative effects, and older cohorts held in reserve for use in special circumstance (for example longitudinal studies, or where ‘boost’ samples are required).

Targeting fieldwork designs

As with survey questions, fieldwork protocols may not work consistently well across different countries. Although there should be a central framework within which countries should operate (including hard limits on fieldwork dates), fieldwork protocols should be adapted to specific country contexts to optimise response rates and sample representativeness – for example the levels or types of incentives used or the messaging used to encourage participation

Participant communications and engagement

Study branding and messaging

The Panel study should have a consistent branding, based on (but distinct from) the parent study, to help develop and maintain panel member engagement. The precise messaging and branding may vary between countries (for example the extent to which the study’s cross-national nature is emphasised), but the overall messaging should emphasise the importance of the study and the panel member’s continued participation, and addressing any potential concerns (for example around data protection). While incentives should be mentioned, messaging should avoid implying a ‘transactional’ relationship which might negatively impact data quality, with key benefits emphasised of having the opportunity to have their views and experiences represented.

Translations

All participant communications should be translated, with care taken to ensure equivalent meaning and show sensitivity to the optimal wording for a given country. With multiple mailings across multiple waves, we recommend that the content of mailings is standardised across waves as much as possible to minimise the costs and time associated with translating documents. For example, while information

about the content of a specific wave may change, the instructions for how to access the survey should stay the same.

Mailings

Panel members should be sent multiple communications to keep them engaged and informed, and encourage survey participation. Reflecting the web methodology, and to reduce costs, the primary mode of communication should be email, however, postal mail should still be used to supplement these, reaching those who do not check their emails regularly and adding credibility to the study (in particular in earlier waves). Text messaging should also be used to reach additional groups and in particular improve participation rates (Bosnjak et al. 2008).

The first mailing panel members should receive should be a ‘welcome mailing’, bridging the gap between their recruitment in the parent study and the first panel survey wave, providing them with additional information about the study, and confirming their joining.

Subsequently, panel members should receive communications related to their participation in specific waves, including information on how to take part and the content of the survey. Panel members should receive an invitation mailing at the start of fieldwork, and multiple reminders spread across the fieldwork period. In general, each additional mailing will have some marginal impact on participation, but over-contacting panel members may come across as harassment and increase costs disproportionately. Reminder contact should therefore not exceed more than once in any one week, with the overall aim of contacting panel members at a range of times of day and days of the week, through multiple modes and with a variety of messages.

Finally, ‘between-wave’ mailings should be used to maintain panel member engagement – providing feedback on findings, or on how the study has made an impact, and prompting to provide up-to-date contact details via ‘change-of-details’ cards, etc. If survey waves are frequent, then this content may be included in the survey wave communications.

Information sources

Panel members may wish to find out additional information to that provided in mailings or at recruitment. As a web panel, the primary source of this should be online, with panel members given access to a website/portal to access information. Different language versions should be available for each language supported by the study. It should include: 1. key information to meet legal and ethical requirements such as a privacy statement, the legal basis for processing data, and how the data will be used; 2. motivational content emphasising the importance of taking part, demonstrating impact, and addressing potential concerns panel members might have; 3. Links to the latest live surveys, and support for how to take part.

Panel members should also be able to review and update their information and preferences (e.g. email address or whether or not they are sent letters) via the website. This may simply be a web form which gets sent on to an appropriate member of the study team. However, it would be preferable that panel members can log in to their own ‘portal’ where they can update their information and contact preferences directly. When logged in, panel members should also be able to directly access ‘their’ questionnaire, rather than following a generic URL.

Not all panel members will be able to find the information they need online, either due to access issues or the information not being available. Panel members should therefore be able to access additional support by contacting the study team, with a different contact for each participating country and

adequate resourcing in place to manage these interactions. Contact details should be included on the website and in all letters and emails sent to panel members.

Incentives

Incentives should be used in recognition of the value of the time and effort that panel members put into completing surveys, and to encourage participation. Unconditional incentives are generally considered to be more effective than conditional (or ‘on-completion’) incentives (e.g. Singer et al., 1999). However, they are also relatively expensive – as the panel project progresses, in particular if there are high numbers of waves, it may be worth stopping using unconditional incentives for those that have (for example) not taken part in the previous four waves. For a study with a smaller budget, conditional incentives may be preferable for all cases.

Consideration should also be given to the type of incentives offered and the mode of administration. Monetary incentives are typically more effective than non-monetary ones, with a value of c. €5 being typical for European probability-based panels for a 20 to 30 minute survey (Blom et al. 2016), substantially higher than non-probability panels, but broadly in-line with larger longitudinal studies. This is partly a reflection of logistical barriers preventing the provision of incentive values of less than €5. Evidence from CRONOS suggests that a one-off unconditional incentive designed to cover multiple waves may be as effective as the same total value of incentives administered unconditionally across waves. Other approaches include offering payments to charity, or entry in prize draws – though these tend to be considered less effective. Using these approaches would allow for the average value to be reduced below €5. Alternatively, as a web panel, a web-based incentive system may be appropriate, or a points-based one as used in commercial research panels. This would provide additional cost savings as incentives would not need to be mailed, and there would be no cost for unclaimed incentives. However, there is not clear evidence on how this would work on a probability panel, where surveys may be less frequent and we are trying to avoid a transactional relationship with panel members.

Weighting

Non-response bias can occur at multiple stages of the panel study: 1. recruitment to the ‘parent’ study; 2. recruitment to the panel; 3. Requests to leave the panel; 4. non-response to a panel study wave. Weights should be computed to adjust the participating sample at each panel survey wave to be representative of the target population for analysis. One of the advantages of the ‘piggy-back’ recruitment approach is that there is extensive information available for both responders and non-responders at stages 2, 3 and 4, allowing for the effective modelling of non-response¹⁰.

The variables and categories selected should be those that are likely to predict biases in the sample, are key variables that are likely to be used in analysis, or related to outcomes of interest. As this blueprint is for a study without a specific substantive focus, we would recommend focusing on commonly used demographic variables such as sex, age, and education, and variables that might predict non-response such as internet use and civic engagement¹¹. As there are multiple stages of non-response, it may be sensible to model & adjust for non-response at stages 2, 3 and 4 separately,

¹⁰ See, for example, CRONOS user guide, p12-13:

https://www.europeansocialsurvey.org/docs/cronos/CRONOS_user_guide_e01_1.pdf

¹¹ For example, the CRONOS weights used data on whether voted in the last general election, whether born in the country, gender, age, marital status, self-reported urbanity, highest level of education, whether in paid work, and region: https://www.europeansocialsurvey.org/docs/cronos/CRONOS_user_guide_e01_1.pdf
www.seriss.eu

although it may be that the predictors are similar at each stage, and that stage 3 will be quite small, so worth combining them.

While it would be optimal to review the design of the weights at each wave to reflect the changing survey content and potential changes in bias, given the lack of specific research questions we would recommend standardising the weighting approach across waves to improve comparability and reduce time and costs.

While the focus of this study is cross-sectional research, longitudinal analysis is possible. However, with a large number of waves, the number of potential combinations of waves (and therefore different weights) needed for longitudinal analysis is very high, so we would not recommend the study provides longitudinal weights as a standard deliverable.

Ethical considerations

The ethical principles for a cross-national panel study should be the same as any social research survey: respecting the autonomy of participants, protecting them from harm, and maximising the benefits of the research. However, there are a number of specific considerations that relate to this panel study design that may not apply in other cross-sectional or longitudinal studies.

The frequency of data collection and contact for a panel study may be higher than other studies, and care should be taken not to over-burden or harass panel members and to be clear up-front what participation involves and that it is voluntary. A high number of waves, each associated with an incentive, may add up to relatively large amounts of money (although not necessarily unprecedentedly so compared to other studies with large monetary incentives for single waves), which for some people could represent financial coercion to participate.

The frequency of data collection, and its longitudinal nature may also increase the disclosure risk. If researchers are able to link individuals' data together across survey waves then people that may not be identifiable in a single dataset may become identifiable (e.g. through a change in age group, or region). Consideration should be given to the level of access given to longitudinal data, and the variables included whilst trying to maximise compliance with the FAIR principles.

Finally, the ethical implications of a web-only design should be considered. As well there being a potential risk of impacting the quality of the sample, it systematically excludes groups of people (in particular 'vulnerable' groups) from the opportunity to having their opinions and experiences represented. This may be particularly ethically problematic should the research then be used, for example, to inform decision making that affects them.

6. Organisational structure

To achieve high input-harmonisation in order to enable cross-national comparability, it is recommended to use a centralised survey management approach for the online panel¹². In addition to using the same centralised web survey platform to conduct online surveys in all participating countries, the project management should reflect the centralised organisational structure, with the central team playing an intensive project design and project coordination role and the national teams facilitating the local implementation of the panel in their country.

¹² This section is based on the experience with CRONOS and following the recommendations of the data archive of the project, the Norwegian Centre for Research Data.

The key task of the central team is to manage the project in close collaboration with national teams, to advise and supervise national teams, to design and operate the survey centrally, and to act as a link to other project stakeholders. In terms of survey design and survey operations, typical tasks for the central team include questionnaire design, drafting of various generic documents (e.g. respondents correspondence), survey documentation, programming, pretesting and sending out of surveys and online respondent communication, fieldwork monitoring and quality control. A data management team responsible for data extraction, cleaning, curation, documentation, and publishing can be an integrated part of the central team institution or based at a separate data archive (e.g. the archive of the parent survey).

National teams are mainly responsible for getting funding for the panel and panel recruitment in their country via the parent survey, translation of surveys and relevant source documents in their languages, setting up and maintaining a local helpline for respondents, updating panellists contact information, administering incentives, sending out offline correspondence to panellists (e.g. postal letters), and, if offliners are included via an offline mode, conducting and/or monitoring panel fieldwork in alternative mode in their country. In countries where face-to-face fieldwork for the parent survey is conducted by a different organisation than the national team organisation that will be operating the online panel, national teams will be responsible for liaising with the fieldwork organisation for issues related to panel recruitment. National teams will also be depositing and extracting various country-specific information to/from the central system. Due to the cyclical nature of an online panel, a syntax-based workflow can be developed for some of these tasks.

To get the best out of the web mode advantages in a cross-national setting, efficient use of technology and smooth communication and information flow between involved project partners as well as with IT systems are crucial. An optimal IT infrastructure is likely to consist of various tools that enable national teams to carry out their tasks on their own using the central system. The tools should also enable automation of labour processes where possible. As academic online panels, especially cross-national panels, are usually relatively complex as compared to commercial surveys, available commercial online survey products might be not flexible enough to deal with the cross-national project complexity without further customisation. It should be carefully decided which tools are the most appropriate for the project's aims. In case of long-term projects, it might be more optimal if the tools to run the cross-national online panel (especially the sample management system) are developed and maintained by a project stakeholder in a close collaboration with the central team (whereby some elements can be still outsourced to commercial providers). This would enable a setting where the technology is adjusted to suit the needs of the project rather than adjusting the needs of the project to the limitations of available technological infrastructure. It is also important that the selected online panel system is able to communicate with other external systems relevant for the project (e.g. web survey platform, database where all the information are stored). Further, it is crucial for the central team to avoid possible complications for the central infrastructure resulting from local issues in participating countries where standardisation of certain procedure is not possible. Before implementation of alternative solutions, suggested local optimisation should be thoroughly assessed by the central team in terms of its possible implications for different areas of the central infrastructure.

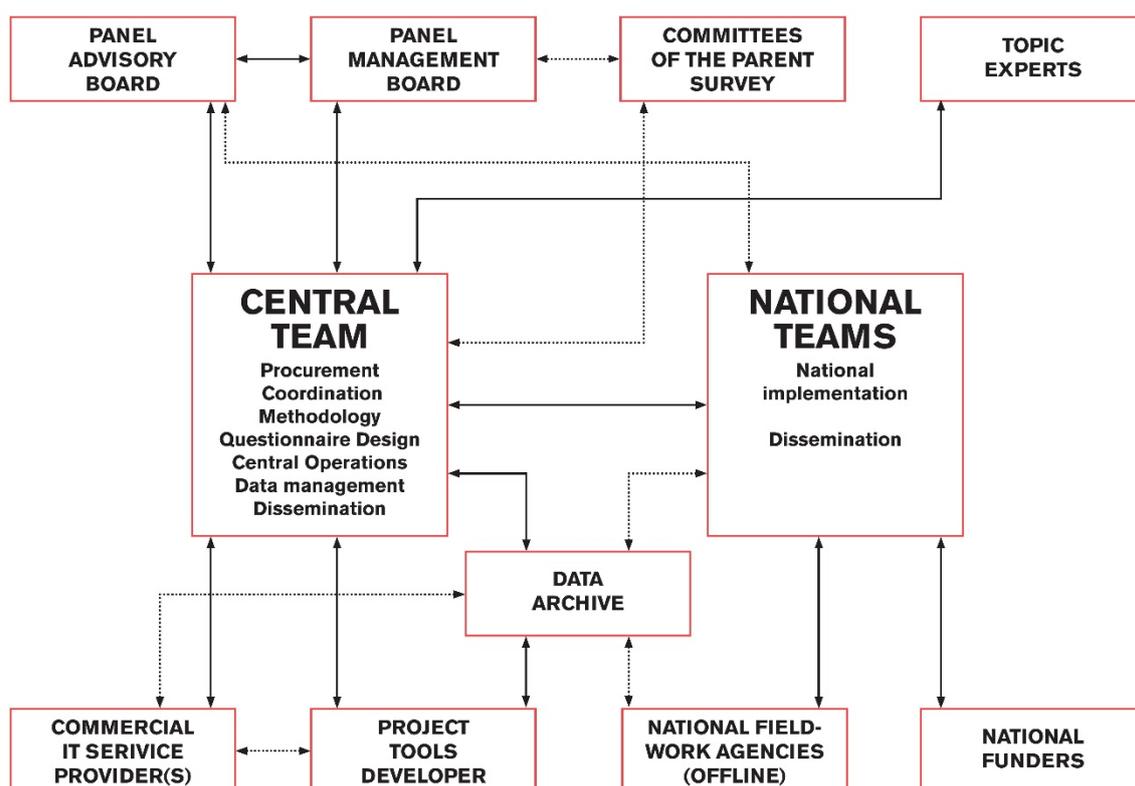
Building of an Advisory Board consisting of online survey methodology experts is recommended with regular meeting to guide the methodological and operational work of the central team, to discuss panel progress at different stages of its life cycle as well as to enable exchange of best practices and

knowhow, to discuss potentially critical issues, implemented and planned strategies as well as design and examine outcomes of methodological experiments.

Depending on the scale of the project and whether it becomes an independent unit or an integrated subproject of the parent study, further committees (such as ethics committee, steering committee, finance committee) might be needed to run the project. If the online panel is conducted as an integrated part of the parent study, the online panel can “piggy-back” on existing committees of the parent survey. In absence of such committees, a panel management board responsible for taking or advising on major decisions could be an alternative.

In addition, collaborations with different topic experts as well as survey researchers and practitioners can be fruitful for developing innovative ideas for questions and experiments to be implemented in a cross-national online panel (e.g. through public calls for questionnaire and experiment proposals). In addition, collaboration with other large cross-national surveys and established national probability-based online panels are encouraged to explore possibilities of infrastructure and tools sharing for specific purposes, and to enable design of joint experiments and cross-survey exchange of experience. The following diagram demonstrates a possible organisational set up for a cross-national online panel.

Figure 1 – Online panel organisational set up



Full or dotted arrows indicate the intensity of interaction: full arrows indicate more intense interactions

7. Data outputs

General principles

The culmination and purpose of the data collection processes are the data outputs, and the fundamental value of the project is determined by the utility of these to the research community. Data outputs should aim to maximise their utility by ensuring that they are archived following FAIR

principles: that they are findable, accessible, interoperable, and re-usable. They must also respect the reassurances provided to respondents about confidentiality. Datasets should be checked, cleaned and systematically named/labelled before archiving in accessible formats alongside documentation of the data and study: fieldwork summaries, questionnaire specifications, derived variables, weighting variables, etc. This documentation, like the data themselves, should follow FAIR principles and be accessible and useful to researchers, including guidance on how to use the data and specific variables (e.g. weighting or sampling variables, or specific derived variables).

Consideration should also be given to ensuring the data outputs protect panel members. All data outputs should be reviewed for disclosure risk with national teams playing a key role as disclosure risk differs significantly according to size of country. All contact details used for panel management should be removed from the published data, variables top-coded¹³, etc. Additional consideration should be given as to whether longitudinal data may increase the risk of disclosure. Some forms of data may require additional curation – for example the linking of geodemographic data or administrative records and may need to be made available under special license.

Data included in outputs

The content of the data outputs can be categorised into three groups: sample data that is available for all panel members, paradata that is available for all panel members invited to a particular wave, and survey data that is available for all panel members that participate in a survey wave. In combination, these data should provide all that is required for the substantive analysis of survey data, as well as transparency and analysis of the methodological elements.

Table 1: data outputs

Data type	Source	Available for	Number of sets	Examples of data included
Sample data	Parent survey Sample management system	All panel members	One per panel cohort	Survey data from parent survey, sampling variables for parent survey, contact preferences, whether left the panel, experimental groups, etc.
Paradata	Sample management system Panel survey systems Panel comms systems	All panel members invited to participate in a panel survey	One per survey wave	Fieldwork paradata: date of interview, outcome code, etc. Survey paradata: page level timings data, session number, user-agent string, answer changes, etc. Comms paradata: experimental groups, email open/click-through rates, etc.
Survey data	Panel survey	All panel members that participate in a panel survey	One per survey wave	Answers to survey questions, routing variables, survey weights, derived variables, etc.

¹³ For example: when asked ‘how many people live in your household’, all those over 5 might be grouped together into a ‘top-code’ of ‘6+’, to reduce the risk of disclosure.

Timing of releases

Releases should aim to be as prompt following data collection as possible, with a ‘rolling release’ of data following each fieldwork wave. To minimise the time between the conclusion of fieldwork and the release of data, processes should be designed to be as standardised and efficient as possible:

- Standardisation and documentation of the systems, processes and conventions used for the management, checking, cleaning, etc. of data across waves should reduce processing time and errors
- Standardisation of the weighting approach should reduce the time required for creating these for specific waves
- Early delivery of data during fieldwork should be possible with web questionnaire systems, allowing for early checks/set up of syntax for the processing of data and development of weights

With these principles followed, and adequate resourcing, it should be possible to process data and make it available for researchers within two months of fieldwork completion. However, an additional limitation may be the parent study. Large-scale face-to-face studies tend to take longer to process the data, and it is possible that the data collection from the first substantive wave of the panel would be completed before the data for the parent study are ready. With parent study data being necessary for the computation of weights/provision of additional variables for analysis, it may be necessary to delay the (full) delivery of the early waves of panel survey data.

8. Impact and dissemination

In order to maximise the value of the cross-national panel study, part of its design should consider strategies for dissemination of findings and data to maximise the study’s impact, and therefore value.

Connections to parent study

The use of a ‘piggy-back’ approach will mean that the panel study will have an associated parent study. It would be efficient to make as much use as possible of any established dissemination infrastructure related to the parent study such as a study web page or social media presence. This would not only be cost-effective, but it is likely that the target audiences for the outputs of the two studies will overlap substantially meaning that any outputs would likely be reaching the right people. However, some care should be taken to distinguish the two.

Data archiving and documentation

A fundamental element of the impact and dissemination should be through data archiving and documentation (discussed in Section 7). The data should be made as accessible and easy to use as possible, with a key impact activity the encouraging of use by researchers external to the project team and beyond those from outside the team who propose and design questions included on the panel.

Reporting

As well as producing data outputs and encouraging others to use them, the study team should produce some written outputs independently as part of the project. These may make the data more accessible and engaging, especially to non-data users, and allow for more direct impact. However, this should only take place after the data has been released publicly to ensure that everyone has equal access to the data.

Core/repeat data collection

Some of the survey content will be ‘core’ measures of interest agreed by the study team. Alongside the archiving of the data we would recommend the production of a short descriptive report and set of data tables/charts covering the ‘headline’ figures from data collected at that wave – for example simple population estimates for each variable for each country and, for repeat measures, if/how that has changed from the previous wave. However, displaying information across time with multiple countries can be challenging and a highly time-consuming task in a cross-national survey.

The purpose of these should not be to go into lots of detail or aim to answer complex research questions, and should avoid stepping on the toes of those looking to do more detailed substantive research based on the findings. Rather, it aims to raise awareness of the data that are available, and to provide simple figures for those unable to run their own analysis.

Ad-hoc data collection

This form of report would be less appropriate for data collected on an ‘ad-hoc’ basis. However, in these instances, to ensure that the data collection is worthwhile, one of the conditions of carrying questions for specific research projects could be that those researchers provide a short summary report covering initial findings from the data to be published as a study output.

Methodological work

A report covering the potential methodological findings should also be published alongside the data. This should not be a detailed report, but rather a simple descriptive output to help to raise awareness of the data and ensure that the methodological work is conducted and published in some form.

Online impact and events & conferences

As with many studies, a key mode of dissemination for the study would be online. The study should have its own study page (potentially within the existing website infrastructure of the parent study), which people can use to find out key information, access data and documentation, and find examples of how the data have been used. A study social media presence may also be useful for ensuring that the study, or research that uses data from the study, finds as large an audience as possible.

However, as well as finding a broad audience, it is important that the study also engages with a ‘core’ audience in a more in-depth fashion, and offline events may be more appropriate for this. Given the potential volume of data collected, and therefore research conducted using it, the study should be able to provide sufficient content for short events/conferences aimed at engaging with and promoting research conducted using the panel. If possible, these should be themed (e.g. a methods event, an event focused on a specific substantive area, etc.) to ensure that attendees with broader interests are more likely to attend, and include training sessions for people looking to start using the data.

9. Costs

In this section, we consider the likely costs of running a cross-national web panel that recruits off the back of an established face-to-face survey with up to 800 panellists in 12 countries and 7 waves. We first consider the number of person months the central team should seek funding for, and then the estimated average cost per interview minute (CPIM) related to the national costs. Costs that were covered by the parent survey (e.g. sampling) are not considered here, as they would not be encountered when running a panel based on the ‘piggy backing’ model presented in this Blueprint. For more information about the costs of the CRONOS web panel, including a detailed list of costs incurred by the central and local teams, see SERISS deliverable D7.6.

The costs of a cross-national web panel can be split into central costs and local costs.

In order to field 7 waves in 12 countries over 2 years, it is estimated that the central team would need approximately 350 person months. This would cover all of the tasks of the central team although some costs would need to be budgeted on top of this, for example the cost of any provision for offliners.

We calculated the cost per interview minute (CPIM). The CPIM calculation involves dividing the total costs¹⁴ for that country by the estimated total number of interviews across all 7 waves (to give the cost per interview), and then dividing this by the mean expected interview duration.

It is estimated that a cross-national web panel fielded for 7 waves would cost each country, on average, €1.13 per interview minute. Looking separately at lower/medium-wealth and higher-wealth countries¹⁵, in lower/medium-wealth countries we estimate the cost to be approximately €0.84 per minute, and in higher-wealth countries €1.42 per minute. Below, we consider the elements that are behind in these figures.

Types of costs incurred

Central costs

In the preparation stage, costs incurred by the central team are likely to include: the preparation of generic documents (including protocols, instructions, manuals and specifications), the recruitment interview, panellist communications and any open calls for questionnaire content proposals; purchase and set-up of any communication and data processing software; liaison with project partners and National Coordinating (NC) teams; meetings with the Advisory Board, questionnaire design teams, project partners and NC teams; training for NC teams and programmers; budget management; any procurement associated with provision for offliners; and, preparations for any methodological experiments.

During implementation, the central team is likely to incur the cost of: interviewer training supervision; recruitment monitoring; questionnaire design, and programming and testing the source questionnaire; responding to translation queries; monitoring participation; data processing and managing experiments; and, reporting and dissemination activities.

Local costs

In the preparation stage, the costs incurred by participating countries are likely to include: the translation and adaptation of generic documents, the production of country-specific documents (e.g. the project flyer), and the printing and distribution of documents that are to be shared in hardcopy; the production of a national website for the panel; setting up the national helpline; attending interviewer training and providing a training extension at the parent-survey interviewer training; and, identifying, purchasing and distributing suitable incentives.

During the implementation stage of the panel, country teams are likely to incur the costs of: interviewer supervision and support, recruitment monitoring, uploading recruitment data, and organising interviewer briefings during recruitment to discuss the recruitment process and best strategies; translating questionnaires and transferring the questionnaires to the web survey platform;

¹⁴ The 'total costs' can be either the actual cost or an estimate.

¹⁵ We ordered the countries according to Eurostat's 2018 figures for GDP per capita (Eurostat, 2019). The CPIM figures showed a clear division at around the mid-point in the GDP range (€30,000), where there was a slightly larger gap in GDP than between most of the other countries.

translating, programming and testing participant communications, and distributing these to panellists; processing responses to any open-ended questions; maintaining the national helpline; meeting with the central team (plus travel costs); and, reporting and dissemination activities.

Additional costs will be incurred at both stages if including offliners. For example, if providing hardware and connectivity for offliners to participate in web mode these costs may include: sourcing devices and internet service providers, producing documentation and manuals, the administration of internet provision, setting up the devices with the relevant software, setting up email accounts for delivery of email communications and electronic incentives, briefing interviewers regarding delivering devices and showing panellists how to use them, and delivery of devices to the offliners.

10. Funding and governance

A key decision to be made is over governance and ownership of the panel. As this is a follow-up panel to a face-to-face 'parent' study the optimal model would be that the organisation running the parent study takes responsibility for the panel too. However, if only a subset of those member countries are taking part in the panel then clear governance arrangements within the parent organisation will be required. Alternatively, a new organisation might take over the panel (either as a legal entity or as an informal consortium) and clear agreements between the parent survey and the panel organisation would need to be put into place (particularly regarding data protection, consent and linkage to the data from the parent study). In general, a model that keeps the parent survey organisation in charge of the panel is recommended as it simplifies the overall organisation and makes efforts such as returning to the parent for sample refreshment more straightforward. It also helps to ensure the longer-term sustainability of the panel by keeping it tied closely to the parent.

There are four funding models that are most likely to be used based on existing experience of Research Infrastructure funding in Europe for infrastructures such as SHARE ERIC, CESSDA ERIC and ESS ERIC:

- 1) A single central funder pays all the costs. It might be possible to find a single funder like the European Commission to cover all costs for a panel which was the model in the CRONOS panel and for SHARE in its early years.
- 2) A single funder pays the central costs but local costs are met by a national funder. This was the early model for the ESS prior to it becoming an ERIC (European Research Infrastructure Consortium).
- 3) The participating countries share the central costs according to an agreed formula but local costs are met by a national funder. This is the current model for ESS ERIC. The model most often used is based on nominal GDP (again sometimes with a larger contribution from the country housing the HQ).
- 4) The participating countries share the central costs and the local costs according to an agreed formula. In other words, the overall costs at local and central level are added together and that total is divided according to the agreed funding formula.

If country specific waves are being included, then there might need to be a separate arrangement depending on how these draw on central resources (e.g. programming) as well as those at national level.

The most important issue for a panel such as this is to ensure there is funding for a sustained period when the panel is expected to operate. So participating countries would need to make a clear commitment to fund all of the waves anticipated up until a possible renewal / refreshment date. It would be crucial to ensure guaranteed funding and not allow this to be dependent in *ad hoc* grant

applications or uncertain funding sources. Since countries would primarily be taking part in order to facilitate cross-country comparisons there would need to be a clear guarantee that all countries taking part could meet their financial and operational obligations for participation.

REFERENCES

Blom, A. G., Bosnjak, M., Cornilleau, A., Cousteaux, A. S., Das, M., Douhou, S. & Krieger, U. (2016). A Comparison of Four Probability-Based Online and Mixed-Mode Panels in Europe. *Social Science Computer Review*, 34(1), pp. 8-25. doi: 10.1177/0894439315574825

Bosnjak, M., Neubarth, W., Couper, M. P., Bandilla, W., & Kaczmirek, L. (2008). Prenotification in web-based access panel surveys: The influence of mobile text messaging versus e-mail on response rates and sample composition. *Social Science Computer Review*, 26(2), 213-223.

Bottoni, G. and Sommer, E. (2019) Best strategies to recruit and maintain web panels. Deliverable 7.5 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221.

Callegaro, M., Baker, R. P., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (Eds.). (2014). *Online panel research: A data quality perspective*. John Wiley & Sons.

Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). *Web survey methodology*. Sage.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public opinion quarterly*, 64(4), 464-494.

Durkheim, E. (1964). *The rules of sociological method*, 8th edition, New York: The Free Press.

Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet research*, 15(2), 195-219.

Jessop, C. (2018). The NatCen Panel: Developing an open probability-based mixed-mode panel in Great Britain. *Social Research Practice*, 6, 2-14.

Lynn, P. (2014). Longer interviews may not affect subsequent survey participation propensity. *Public opinion quarterly*, 78(2), 500-509.

Schaurer, Ines, Bernd Weiß, Kai Willem Weyandt, and Michael Blohm. 2019. "Exclusive Recruitment Interview vs. Piggybacking: Comparison of Two Recruitment Strategies for a Probability-Based Self-Administered Panel." 8th Conference of the European Survey Research Organisation (ESRA), 18.07.2019.

Schmidt, W. C. (1997). World-Wide Web survey research: Benefits, potential problems, and solutions. *Behavior research methods, instruments, & computers*, 29(2), 274-279.

Singer, E., Van Hoewyk, J., Gebler, N., & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics*, 15(2), 217.

Sommer, E. (2017) Summary of existing web panel strategies. Deliverable 7.1 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Available at: www.seriss.eu/resources/deliverables.

Sue, V. M., & Ritter, L. A. (2012). *Conducting online surveys*. Sage.

Survey Research Center (2016). Guidelines for Best Practice in Cross-Cultural Surveys. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. Retrieved July 5 2019, from <http://www.ccsr.isr.umich.edu/>

Villar, A., Sommer, E., Finnøy, D., Gaia, A., Berzelak, N. and Bottoni, G. (2018) CROss-National Online Survey (CRONOS) panel data and documentation user guide. London: ESS ERIC

Villar, A. and Sommer, E. (2017). Web recruitment design plans and experimental testing. Deliverable 7.3 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221