

Questions to help buyers of online samples.



Contents

 Purpose and Scope	3
 Company profile	4
 Sample sources and Recruitment	5
 Sampling and Project management	9
 Data quality and Validation	12
 Policies and Compliance	14
 Metrics	18
 Project team & Sounding board	19
 Glossary	20

Purpose and Scope

This set of questions offers a framework for buyers to use when evaluating the offerings of different online sample providers. It updates and replaces the 2012 ESOMAR publication, **28 Questions to Help Buyers of Online Samples**.

The questions identify the key issues to consider, introduce consistent terminology, explain why each question should be asked, and note the issues buyers should expect to be covered in an answer. The intended use of these questions is that they form a basis for a conversation between buyer and sample provider, rather than simply being used as a checklist to compare offerings across providers.

The questions do not cover B2B samples, nor do they attempt to cover specific requirements for different types of research such as pricing, new product development, ad testing etc.


When online access panels were first introduced in the 1990s, the model was relatively simple: a buyer provided sampling specifications to a panel owner who drew a sample (from that panel). Over the intervening 25 years, online sample selection has changed in two fundamental ways. First, buyers can now access a a broader set of sources that now includes participants in loyalty programmes and rewards communities within “Get Paid To” sites, customer lists, intercepts from offer walls, affiliate networks, social media, and other platforms, as well as traditional panels that may or may not be owned by the provider. Second, buyers have the option to access these sources directly via self-service tools, rather than relying on a sample provider to generate the sample on their behalf.

There have been other important changes as well. Online research has become truly global and mobile devices have become a common data collection platform. The use of online samples has broadened beyond surveys to include qual/quant applications, communities, passive data collection, and so on. Concerns about privacy and data protection have led to a much-changed regulatory environment that imposes new requirements on both sample buyers and sample providers. Quality assurance techniques have become increasingly sophisticated. As a consequence, the number of issues that buyers must consider when choosing a sample provider has increased substantially.


Finally, sample quality is an essential component of all research but it alone does not guarantee reliable, actionable results. While not covered in this document, we note that a well thought out research design; a clear definition of the target population; a questionnaire that is both easy for participants to complete and accurately measures key variables; and a well-designed analysis plan are also essential.

Company Profile


- 1. What experience does your company have in providing online samples for market research? How long have you been providing this service? Do you also provide similar services for other uses such as direct marketing? If so, what proportion of your work is for market research?**

 This answer might help you to form an opinion about the relevant experience of the sample provider as well as potential biases that might result from other uses such as being paid to watch ads or receiving a high volume of marketing messages.

- 2. Do you have staff with responsibility for developing and monitoring the performance of the sampling algorithms and related automated functions who also have knowledge and experience in this area? What sort of training in sampling techniques do you provide to your frontline staff?**

 It is important to know if the provider's offerings have been designed by and are monitored by staff with knowledge of basic principles of sampling. This may be useful at the sample design stage as well as during fulfilment when quotas become difficult to fill or when weighting may be required. Ditto for any frontline staff who may serve as your main point of contact with the sample provider.

- 3. What other services do you offer? Do you cover sample-only, or do you offer a broad range of data collection and analysis services?**

 Depending on your company's capabilities, you may wish to work with a one-stop shop that can host your survey, produce basic tabulations, code open ends, and so on. There may be time and cost savings with this approach.

Sample sources and Recruitment

What this section is about

Answers to the questions in this section will help you understand the types of sample available from different sample providers in the market and the sources they rely on. This will help you evaluate the quality of the sample being offered, whether it is suitable for measuring change over time, and whether there are any specific constraints you need to consider when using it. It will also allow you to understand whether the sample provider is drawing the sample from its own sources or aggregating sources from other providers. We recommend that you first identify the sample types being offered and then ask the relevant questions for all sources.

Broadly speaking, there are two models of sample sources and recruitment:


Panels

These are databases of potential participants who declare that they will cooperate for future data collection if selected, generally in exchange for a reward/incentive. This includes traditional access panels, co-branded panels, or opt-in databases of individuals who agreed to complete research projects and also undertake other non-market research activities (watch ads, download an app, complete marketing offers, etc, also known as loyalty programmes, or rewards communities within GPT (Get paid to) sites.) Loyalty card and subscription databases are included here if there is a continuous relationship with members who understand the commitment asked of them.


Intercepts

This includes intercepts from offer walls, affiliate networks, social media or other platforms to drive traffic to a survey. Intercept is an approach where potential participants are asked to take a survey for a reward while they are engaged in another activity such as playing an online game, reading news, or some other online activity. Intercepted participants may be previously unknown to the sample provider or may have been pre-identified and profiled through a prior survey experience.

4. Using the broad classifications above, from what sources of online sample do you derive participants?

 Sample providers may deliver sample from a single source, such as their own proprietary panel, or other panels. Or they may leverage a range of technologies and platforms to aggregate/blend participants from a combination of sample sources. Some providers may do both. Clarity about the sources being used will help you to understand what type of sample is being offered. This answer might differ from country to country and from project to project.


- 5. Which of these sources are proprietary or exclusive and what is the percent share of each in the total sample provided to a buyer?** (Assume proprietary to mean that the sample provider owns the asset. Assume exclusive to mean that the sample provider has an exclusive agreement to manage/provide access to sample originally collected by another entity.)

 This question will help you to understand whether the vendor is ‘running’ the source or ‘marketing’ the source. Running the source implies a closer relationship with panellists and a deeper knowledge of recruitment techniques. This may also help you to understand whether the sample is exclusively available from this provider.

- 6. What recruitment channels are you using for each of the sources you have described? Is the recruitment process ‘open to all’ or by invitation only? Are you using probabilistic methods? Are you using affiliate networks and referral programs and in what proportions? How does your use of these channels vary by geography?**

 Understanding the method of recruitment and whether the recruitment is by invitation only will help you to understand the quality of the sample and how it may be used.

- 7. What form of validation do you use in recruitment to ensure that participants are real, unique, and are who they say they are?** Describe this both in terms of the practical steps you take within your own organisation and the technologies you are using. Please try to be as specific and quantify as much as you can.

 Understanding the level of recruitment validation undertaken by the sample provider will help you to mitigate effects of fraud in your projects. Working with providers who have fully developed strategies and are using up to date detection technologies is recommended.

8. What brand (domain) and/or app are you using with proprietary sources? Summarise, by source, the proportion of sample accessing surveys by mobile app, email or other specified means.

By understanding the domain/app and method the sample provider is using with members, you will gain an indication of the extent of activity with those members and the quality of their relationship with the sample.


9. Which model(s) do you offer to deliver sample? Managed service, self-serve, or API integration?

Sample provision is offered through three main channels: managed service, self-serve, and API (Application Programming Interface) integrations. In a self-serve model, buyers are given access to a platform which they can use to specify the audience they want to access, and manage all the steps of a research project, from sample design to launch to fieldwork management to closing. In a managed service model, sample providers will provide that service. API integrations are the mechanics which allow sample providers, buyers and data collection platforms to automate some aspects of the process.

10. If offering intercepts, or providing access to more than one source, what level of transparency do you offer over the composition of your sample (sample sources, sample providers included in the blend). Do you let buyers control which sources of sample to include in their projects, and if so how? Do you have any integration mechanisms with third-party sources offered?

It is well documented that different sources can produce different results. Consistency in source blending can be vital for tracking studies or other inter-survey comparisons. The use of a single, narrow source, such as a single supermarket's loyalty scheme, may result in unintended bias.

11. Of the sample sources you have available, how would you describe the suitability of each for different research applications? For example, Is there sample suitable for product testing or other recruit/recall situations where the buyer may need to go back again to the same sample? Is the sample suitable for shorter or longer questionnaires? For mobile-only or desktop-only questionnaires? Is it suitable to recruit for communities? For online focus groups?

 By understanding the constraints of the sample being offered, you can understand if the actual sample available from the provider meets your particular research needs and changes any of the answers given previously to this section.

Sampling And Project Management

What this section is about

Answers to the questions in this section will help you understand the processes and procedures that are undertaken to provide you with a sample of participants for your survey. You should understand what biases may be inherent in, or as a result of, the approaches taken and the likely severity of those biases.


12. Briefly describe your overall process from invitation to survey completion. What steps do you take to achieve a sample that “looks like” the target population? What demographic quota controls, if any, do you recommend?

☞ The sampling process (i.e., how individuals are selected or allocated from the sample sources) may affect how random the sample is from within the sources proposed. Quota controls are commonly used to make samples look like the target population and, if done without thought, may be less than optimal for your particular project.


13. What profiling information do you hold on at least 80% of your panel members plus any intercepts known to you through prior contact? How does this differ by the sources you offer? How often is each of those data points updated? Can you supply these data points as appends to the data set? Do you collect this profiling information directly or is it supplied by a third party?

☞ Targeting samples based on pre-existing profiles increases efficiency. Some bias may result depending on the precise questions asked, when they were asked, and to how many people. Appending existing information reduces the burden on the panellists in the survey itself.


14. What information do you need about a project in order to provide an estimate of feasibility? What, if anything, do you do to give upper or lower boundaries around these estimates?

 A sample provider failing to meet your sample requirements may require use of additional sample providers, adding time and complexity to the project. Trackers should be assessed in the light of any exclusion periods you may want to introduce that will reduce the available sample for subsequent waves.


15. What do you do if the project proves impossible for you to complete in field? Do you inform the sample buyer as to who you would use to complete the project? In such circumstances, how do you maintain and certify third party sources/sub-contractors?

 There may be good reasons why certain sample providers should not be used. For example; the provider may not have experience of operating in the geography relevant to your project.


16. Do you employ a survey router or any yield management techniques? If yes, please describe how you go about allocating participants to surveys. How are potential participants asked to participate in a study? Please specify how this is done for each of the sources you offer.

 Biases of varying severity may arise from prioritization in the order in which surveys are presented to participants or the methods used to allocate a participant to one of the various surveys for which they may appear to qualify.


17. Do you set limits on the amount of time a participant can be in the router before they qualify for a survey?

 An excessive amount of time spent in a router answering screening questions may cause a participant to become fatigued, potentially impacting data quality.


18. What information about a project is given to potential participants before they choose whether to take the survey or not? How does this differ by the sources you offer?

 The information about the survey (and associated rewards) may influence the type of people who agree to take part, creating the potential for bias.


19. Do you allow participants to choose a survey from a selection of available surveys? If so, what are they told about each survey that helps them to make that choice?

 The level of detail and the nature of the information given about a project may influence who responds, creating the potential for bias.


20. What ability do you have to increase (or decrease) incentives being offered to potential participants (or sub-groups of participants) during the course of a survey? If so, can this be flagged at the participant level in the dataset?

 The reward or incentive system may have an impact on the reasons people participate in a specific project and these effects can result in bias in the sample.

21. Do you measure participant satisfaction at the individual project level? If so, can you provide normative data for similar projects (by length, by type, by subject, by target group)?

 Participant satisfaction may be an indicator of willingness to take future surveys. Participant reactions to your survey from self-reported feedback or from an analysis of the points where participants drop out of the survey may enhance your understanding of the survey results and lead to improvements in questionnaire design for future surveys.

22. Do you provide a debrief report about a project after it has completed? If yes, can you provide an example?

 You should expect a full sample debrief report. Sample providers should be able to list the standard reports and metrics that they make available.

Data quality and Validation

What this section is about

This section focuses on the quality of the in-survey data. In-survey data quality includes project level data validity and representativeness, survey-taking behaviours, sample blends, participant characteristics, and project level data health and audit practices.

23. How often can the same individual participate in a survey?

How does this vary across your sample sources?

What is the mean and maximum amount of time a person may have already been taking surveys before they entered this survey? How do you manage this?

Answers to this question may alert you to about the potential for bias due to the participation of professional participants, simply survey fatigue, or category bias.


24. What data do you maintain on individual participants such as recent participation history, date(s) of entry, source/channel, etc? Are you able to supply buyers with a project analysis of such individual level data? Are you able to append such data points to your participant records?

You may wish to append data that enables you to analyse and trend data to look for potential biases based on participation levels, sources, tenure, and other data the provider may hold.


25. Please describe your procedures for confirmation of participant identity at the project level. Please describe these procedures as they are implemented at the point of entry to a survey or router.

Given the widely acknowledged risk of fraud in online research, buyers should understand identity and fraud controls, not just at recruitment, but at the point of survey entry. It is essential that there be measures in place to ensure that participants are who they say they are and that the member or email account has not been hacked, is not a duplicate with other accounts from other channels or panels, and whether or not the account is shared by other members of the household.


26. How do you manage source consistency and blend at the project level? With regard to trackers, how do you ensure that the nature and composition of sample sources remain the same over time? Do you have reports on blends and sources that can be provided to buyers? Can source be appended to the participant data records?

 Participant source is a known contributor to data representativeness. Knowing all the sources used for the project, especially for tracking and longitudinal research, and that the proportions from each source are known and reportable over time, will allow you to understand any population biases that might exist.

27. Please describe your participant/member quality tracking, along with any health metrics you maintain on members/participants, and how those metrics are used to invite, track, quarantine, and block people from entering the platform, router, or a survey. What processes do you have in place to compare profiled and known data to in-survey responses?

 Buyers and providers often work together to track individual survey response quality, so buyers should understand what data the provider uses to confirm survey answers, block or remove a member, and how to enable that information exchange.

28. For work where you program, host, and deliver the survey data, what processes do you have in place to reduce or eliminate undesired in-survey behaviours, such as (a) random responding, (b) illogical or inconsistent responding, (c) overuse of item non-response (e.g., “Don’ t Know”) (d) inaccurate or inconsistent responding, (e) incomplete responding, or (f) too rapid survey completion?

 Data cleansing methods are often built into survey programs and platforms. Some of those methods are set up to automatically remove responses, while others are optional or manual. Understanding what tools will be used will aid buyers in understanding how much cleaning they should plan to do once they receive the final dataset, and what biases might be introduced by automated cleaning practices.

Policies And Compliance

What this section is about

Sample providers, buyers, and their clients are subject to data protection and related information security requirements imposed by data protection laws and regulations. In addition, they may be subject to laws and regulations that may impact incentives paid to participants.


These laws and regulations vary by jurisdiction with different laws and regulations applying in different countries or states within countries, and are generally interpreted based on where the participant resides.

Applicable data protection laws and regulations include, but are not limited to: the Act on the Protection of Personal Information or APPI (Japan); the Australian Privacy Act (Australia); the California Consumer Protection Act or CCPA (state of California in the United States); the Children's Online Privacy Protection Act or COPPA (United States); the Data Protection Act (United Kingdom); amendments regarding data localisation requirements to the Data Protection Act (Russian Federation); the General Data Protection Law (Brazil); the EU General Data Protection Regulation or EU-GDPR (EU/EEA); the Health Insurance Portability and Accountability Act or HIPAA (United States); the Graham-Leach Bliley Act or GLBA (United States); and PIPEDA (Canada). AB 2257 (the state of California in the United States) is an example of law and regulation related to employment which may impact incentives paid to participants.


Information security frameworks and standards include, but are not limited to COBIT, HITRUST, ISO 27001, the NIST Cybersecurity Framework and SOC 2.

Answers to the questions in this section can help you understand the data protection, information security and compliance policies, procedures and practices that a sample provider has implemented.

29. Please provide the link to your participant privacy notice (sometimes referred to as a privacy policy) as well as a summary of the key concepts it addresses. (Note: If your company uses different privacy notices for different products or services, please provide an example relevant to the products or services covered in your response to this question).

 A privacy notice is required by various data protection laws and regulations as well as other laws and regulations as well some market research industry codes. A privacy notice discloses information about the personal data that a sample provider collects and processes and the way that that personal data is used, disclosed, and managed. A review of a sample provider's privacy notice can help you understand their procedures and practices related to personal data and the degree to which they comply with applicable laws, regulations, and industry codes.

30. How do you comply with key data protection laws and regulations that apply in the various jurisdictions in which you operate? How do you address requirements regarding consent or other legal bases for the processing personal data? How do you address requirements for data breach response, cross-border transfer, and data retention? Have you appointed a data protection officer?

 As noted above, buyers and sample providers are subject to data protection and related information security requirements imposed by data protection laws and regulations, other laws and regulations as well as clients. Understanding a sample provider's compliance position with these laws and regulations is essential.

31. How can participants provide, manage and revise consent for the processing of their personal data? What support channels do you provide for participants?

In your response, please address the sample sources you wholly own, as well as those owned by other parties to whom you provide access.

Consent for the collection and processing of personal data has long been required by market research industry codes. It is also explicitly required by some data protection laws and regulations. Some data protection laws and regulations, including EU-GDPR and CCPA as examples, also provide for access rights for participants to correct, update, or delete their data. Implementation of a participant support channel is also required by ISO 20252 (ISO 20252:2019: Market, Opinion and Social Research, Including Insights and Data Analytics - Vocabulary and Service Requirements).


32. How do you track and comply with other applicable laws and regulations, such as those that might impact the incentives paid to participants?

As stated above, buyers and sample providers are subject to laws and regulations such as those that may impact incentives paid to participants.


33. What is your approach to collecting and processing the personal data of children and young people? Do you adhere to standards and guidelines provided by ESOMAR or GRBN member associations? How do you comply with applicable data protection laws and regulations?

Some data protection laws and regulations (for example COPPA and EU-GDPR) impose specific requirements with the respect to the collection and processing of the personal data of children and young people. These requirements include specific age definitions as well as a requirement for verifiable parental consent. See the ESOMAR & GRBN Guideline on Research and Data Analytics with Children, Young People, and Other Vulnerable Individuals for further discussion.


34. Do you implement “data protection by design” (sometimes referred to as “privacy by design”) in your systems and processes? If so, please describe how.

 Data protection by design” (which may also be referred to as “privacy by design”) is an approach that requires the consideration of privacy and data protection issues at the design phase of any system, service, product or process and then throughout the lifecycle. Understanding a sample providers use or lack of use of “data protection by design” can help you understand its data protection compliance posture.

35. What are the key elements of your information security compliance program? Please specify the framework(s) or auditing procedure(s) you comply with or certify to. Does your program include an asset-based risk assessment and internal audit process?

 Information security frameworks such as ISO 27001 or SOC 2 are accepted and recognized frameworks for information security compliance. Understanding which framework(s) a sample provider uses or if a sample provider doesn't use such a framework can help you understand the sample provider's information security compliance posture.

36. Do you certify to or comply with a quality framework such as ISO 20252?

 ISO 20252 is an international quality standard recognised by many market research industry associations. In addition to requirement for a system to manage research processes, it explicitly addresses requirements for data protection and information security compliance.

Metrics

What this section is about

This section lists common sample and data health metrics. Reviewing metrics periodically can serve as the basis for a conversation with sample providers about consistency and reliability, as well as whether the sample is appropriate for the population and business question being examined. Unexpected or unexplained shifts in metrics may also indicate the potential for bias or error. While not all of these metrics are required and there are no benchmarks on the “right answers,” providing transparency over time will create a meaningful dialogue about quality and utility.

37. Which of the following are you able to provide to buyers, in aggregate and by country and source? Please include a link or attach a file of a sample report for each of the metrics you use.

01. Average qualifying or completion rate, trended by month
02. Percent of paid completes rejected per month/project, trended by month
03. Percent of members/accounts removed/quarantined, trended by month
04. Percent of paid completes from 0-3 months tenure, trended by month
05. Percent of paid completes from smartphones, trended by month
06. Percent of paid completes from owned/branded member relationships versus intercept participants, trended by month
07. Average number of dispositions (survey attempts, screenouts, and completes) per member, trended by month (potentially by cohort)
08. Average number of paid completes per member, trended by month (potentially by cohort)
09. Active unique participants in the last 30 days
10. Active unique 18-24 male participants in the last 30 days
11. Maximum feasibility in a specific country with nat rep quotas, seven days in field, 100% incidence, 10-minute interview
12. Percent of quotas that reached full quota at time of delivery, trended by month

Project team and Sounding board

Project Team:

Reg Baker, ESOMAR

Pete Cape, Dynata

Melanie Courtright, Insights Association

Peter Milla, Peter Milla Consulting

Judith Passingham, ESOMAR

Administrative Support:

Joke Ruwen-Stuursma, ESOMAR

Sounding Board:

Rob Berger, Maru/Blu

Adam Birss, Rakutan Insight

Mike Cooke, ESOMAR

Olivier de Gaudemar, Consultant

Jonathan Deitch, Cint

Philippe Guilbert, Syntec Conseil

Jon Puleston, Kantar

Efrain Ribeiro, Consultant

Mary Beth Weber, CASE

Denise Saldana Davalos, Colgate-Palmolive

Glossary

For the purpose of this document these terms have the following specific meanings:

Affiliate partner (or Affiliate network) means a network of communities with which a sample provider has a relationship to direct intercept traffic to their surveys.

API (application programming interface) means a set of definitions and protocols for building software applications capable of accessing and exchanging data.

Blending means the practice of combining multiple, heterogeneous sample sources with the aim of achieving a more consistent or more representative sample.

Children means individuals for whom permission to participate in research must be obtained from a parent, legal guardian, or responsible adult. Definitions of the age of a child vary substantially and are set by national laws and self-regulatory codes. In the absence of a national definition, a child is defined as being 12 and under and a “young person” as aged 13 to 17.

Completion rate means the number of participants who fully complete the survey divided by the number of participants who start the survey.

Consent means freely given and informed indication of agreement by a person to the collection and processing of his/her personal data. Note that the specific requirements for consent will vary by jurisdiction.

Exclusion means excluding a potential participant from a research project based on their previous participation in a research project involving the same or similar product/service category and/or methodology.

Fraudulent participant means a participant who deliberately misrepresents their identity, profiling information, or responses, including organisations that use bots to impersonate participants.

Health metrics means measures of quantitative assessment commonly used for comparing and tracking performance or production over time. In this context, health metrics refers to quantitative data used to track stability or changes in the sample a provider offers, and the metrics suggested are based on data that has been previously known to impact quality over time.

Loyalty programme means an arrangement in which customers of a company (or group of companies) are rewarded for purchases made with these companies. Rewards are normally given in a currency that can be spent at those companies (or their chosen partners).

Paid completes means interviews/surveys that are delivered and accepted by a client, are included in the final dataset, and for which the sample provider receives payment.

Panel member (or simply member) means an individual recruited from a documented source who has provided profile data and appropriate information for validation of identity, given explicit consent to participate in research according to the terms and conditions of panel membership, and has not opted out.

Participant (sometimes call a participant or data subject) means a person or organisation from whom or about whom data is collected for research.

Personal data (sometimes referred to as personally identifiable information or PII) means any information relating to a natural living person that can be used to identify an individual, for example by reference to direct identifiers (such as a name, specific geographic location, telephone number, picture, sound, or video recording) or indirectly by reference to an individual's physical, physiological, mental, economic, cultural or social characteristics.

Profiling information means descriptive characteristics of a panel member.

Quarantined members means individuals who have broken some set of quality assessment protocols that result in them being either temporarily or permanently suspended from participating in future research activities with the company that quarantines them.

Referral program means a process whereby a panel offers its existing panellists the opportunity to gain rewards by referring family, friends and colleagues (or visitors of their site) to join the panel.

Representativeness means the degree to which a sample reflects the target population being studied. A representative sample is one in which the distribution of important characteristics is approximately the same as in the target population.

Rewards community (within Get Paid To (or GPT) sites) means databases or panels of individuals who may undertake non-research activities (watch ads, download an app, complete marketing offers etc) usually in exchange for a reward, but who also agree to take part in research projects.

Router means an online software application that screens incoming research participants and then uses those results to assign participants to one of multiple available research projects. A router can also offer participants additional screeners and surveys after screener qualification failure or survey completion.

Sample provider means a service provider responsible for the provision and management of online samples from relevant sources including panels, intercepts, email lists, etc.

Survey attempts means the number of times the same individual clicked a link or entered into a survey environment in an attempt to complete a survey.

Third Party Sources means sources that the sample provider does not directly run or control.

Yield management means a variable allocation strategy through which outcomes are maximised by matching supply with demand.

ESOMAR

Atlas Arena, Hoogoorddreef 5
1101 BA Amsterdam
The Netherlands
www.esomar.org