

# Computer Assisted Personal Interview (CAPI) Surveys

## Guide to Estimating Duration and Personnel Needs by Task

**Note:** *This guidance is designed to ensure that GAO policies on evidence and generally accepted government auditing standards are met. The guidance conforms to the generally accepted principles and practices of the appropriate disciplines. Statements that particular actions “should” be taken are practices that are expected to be followed, unless there are good reasons for not doing so. Before deviating from a practice expressed as a “should” statement, staff members must consult with an appropriate staff member in Applied Research and Methods (ARM) or a team specialist and must document the consultation.*

This paper is a guide to planning a GAO interviewer-administered survey using a questionnaire. While the questionnaires may take the form of paper-and-pencil data collection instruments (DCI) that are keypunched, this guide also considers the additional steps necessary to create a computerized DCI that presents questions and instructions to interviewers onscreen, and records answers to a data file. This guidance outlines the stages of the process of creating this type of Computer Assisted Personal Interview (CAPI), and, for each stage the tasks required to complete such a survey. It also provides estimates of the minimum, average, and maximum times needed to complete each task in developing and conducting a CAPI survey. These time estimates are affected by a number of factors, which are described in this document.

In general, the more complex the task, the longer it may take to complete. The time estimates provided in this document show the number of elapsed working days that the task may take. For some tasks, mostly those that we have some control, the elapsed time span may be less than what is shown, depending on team resources and other considerations. For example, a task such as editing a contact list may take less time if the list is divided up among multiple staff members. For tasks that we may have less control over, such as making telephone contact with the targeted respondent, we are limited in the extent to which we can speed up the process.

Many tasks can be done concurrently, which will shorten the overall timeframe. Contact an ARM/CDMA survey specialist or other design methodologist for more information.

### Stage 1 – Questionnaire Development and Pretesting

Tasks	Duration (in days) <i>Low &lt; Medium &gt; High</i>		Personnel Needs	
	Working estimates	Team estimates	Required staff	Assigned staff
1.1 Determine whether a survey (of any type) is the best way to obtain the necessary data.	1 < 2 > 3		Engagement staff and ARM/CDMA	
1.2 Develop researchable questions that will respond to the information requested by the requestor or mandated by the legislation.	1 < 10 > 21		Engagement staff and ARM/CDMA	
1.3 Determine whether a GAO-administered telephone survey is the best data collection method.	1 < 2 > 3		Engagement staff and ARM/CDMA	

1.4 Operationalize the research questions into a draft questionnaire containing actual questions, response categories, etc. that can undergo pretesting.	1 < 7 > 21		Engagement staff and ARM/CDMA	
1.5 Pretest and modify until all GAO parties agree that the questionnaire is ready for dissemination.	1 < 24 > 56		Engagement staff and ARM/CDMA	
1.6 Develop a detailed analysis plan and consider how results will be reported (i.e., e-supplement, table shells,etc.).	1 < 14 > 21		Engagement staff and ARM/CDMA	

**Stage 1 – Questionnaire Development and Pretesting – Factors that Affect Time Estimates**

Issue	Response	Result
1.a How complex is the subject matter covered by the survey?	<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low	Complexity ↑   Time ↑
1.b What is the level of subject matter expertise of GAO staff at start of engagement?	<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low	Expertise ↑   Time ↓
1.c Is there prior research/questionnaires on topic?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↓   If no - time ↑
1.d What is the risk level of engagement?	<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low	Risk ↑   Time ↑
1.e What is the level of non-survey workload required of assigned team staff?	<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low	Other workload ↑   Time ↑
1.f What is the length of questionnaire (number of questions/interview time)	<input type="checkbox"/> Long <input type="checkbox"/> Medium <input type="checkbox"/> Short	Length ↑   Time ↑
1.g Will pretest questionnaire be in CAPI form, or paper and pencil at first? <sup>1</sup>	<input type="checkbox"/> CAPI <input type="checkbox"/> Paper	If CAPI - time ↑   If no - time ↓
1.h How many pretests/questionnaire revisions will be needed to reach closure?	<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low	# of pretests ↑   Time ↑
1.i What is the number of distinct survey strata who may, for example, use different terminology thus requiring pretests are multiple physical/geographic locations?	<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low	# of strata ↑   Time ↑
1.j Other factors?		

<sup>1</sup> This choice will imply an opposite impact for Stage 4 – if the CAPI instrument is begun at this stage, it may reduce the length of time taken for this task later.

**Stage 2 – Determination of Survey Population/Sample**

Tasks	Duration (in days) <i>Low &lt; Medium &gt; High</i>		Personnel Needs	
	Working estimates	Team estimates	Required staff	Assigned staff
2.1 Determine the size and characteristics of the survey population.	1 < 21 > 35		Engagement staff and ARM/CDMA	
2.2 Obtain a listing of the survey population.	1 < 21 > 35		Engagement staff	
2.3 Decide on whether the entire survey population should be surveyed or whether a sample is appropriate.	1 < 3 > 5		Engagement staff and ARM/CDMA	

**Stage 2 – Determination of Survey Population/Sample – Factors that Affect Time Estimates**

Issue	Response	Result
2.a What is the size of the population/sample?	<input type="checkbox"/> Large <input type="checkbox"/> Medium <input type="checkbox"/> Small	Size ↑   Time ↑
2.b Does a listing of population exist in one place (e.g., single database)?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↓   If no - time ↑
2.c Does the listing of the population include telephone numbers and other contact information if mail or email advance contact (recommended) planned?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↓   If no - time ↑
2.d Is the listing of addresses in paper or electronic format?	<input type="checkbox"/> Electronic <input type="checkbox"/> paper	If electronic ↓   If paper - time ↑
2.e Is a separate screener survey necessary to identify those eligible for main survey?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
2.f Does the listing have to be supplemented with auxiliary information to measure nonresponse bias (if nonresponse is expected to be high)?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
2.g Other factors?		

**Stage 3 – Sampling (If no sample required, Skip to Stage 4)**

Tasks	Duration (in days) <i>Low &lt; Medium &gt; High</i>		Personnel Needs	
	Working estimates	Team estimates	Required staff	Assigned staff
3.1 Determine the level of generalizability required (i.e., the number of strata that needs to be separately reported on).	1 < 3 > 5		Engagement staff and ARM/CDMA	
3.2 Select the sample	1 < 7 > 14		Engagement staff and ARM/CDMA	

**Stage 3 – Sampling – Factors that Affect Time Estimates**

Issue	Response	Result
3.a Will the sample be stratified?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑    If no - time ↓
3.b How many strata will be required?	<input type="checkbox"/> Many <input type="checkbox"/> Few	Number of strata ↑    Time ↑
3.c Does the population listing from which the sample is to be selected from exist in electronic or paper format?	<input type="checkbox"/> Electronic <input type="checkbox"/> paper	If electronic - time ↓    If paper - time ↑
3.d Is it likely that the population listing has ineligible elements that will be discovered upon sampling, resulting in an adjustment to the sample design?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑    If no - time ↓
3.e Other factors?		

**Stage 4 – Advance fieldwork and interview preparation**

Tasks	Duration (in days) <i>Low &lt; Medium &gt; High</i>		Personnel Needs	
	Working estimates	Team estimates	Required staff	Assigned staff
4.1 Obtain mailing list of survey respondents.	1 < 14 > 28		Engagement staff	
4.2 Edit mailing list (i.e., ensure that addresses are complete and are properly formatted).	1 < 4 > 14		Engagement staff	
4.3 Produce all materials needed for advance email or letter mailout (e.g., address-merged letters, envelopes, postcards, labels, etc.)	1 < 7 > 14		ARM/CDMA and Shared Services	
4.4 Assemble initial contact materials and mail or email.	1 < 2 > 3		ARM/CDMA, Shared Services, Web Services	
4.5 Update and adjust sample frame and contact lists based on responses to advance contact.	3 < 5 > 7		Engagement staff and ARM/CDMA	
4.6 Prepare support materials – interviewer instructions, fieldwork management tools (logs, sample lists, etc.)	1 < 3 > 5		Engagement staff and ARM/CDMA	
4.7 Select, schedule, and train interviewers (moving to CAPI instrument upon development)	5 < 6 > 7		Engagement staff and ARM/CDMA	
4.8 Create CAPI instrument in appropriate platform <sup>2</sup> , test and revise, deploy/distribute final version, with final sample of cases	5 < 10 > 21		Engagement staff, ARM/CDMA, Web Services	

<sup>2</sup> CAPI instruments may be web-based, allowing for a central database to be used by interviewers in different locations to access and complete individual cases, or distributed as stand-alone databases each with individual cases (e.g., on individual GAO laptops) that will later be aggregated into a central database. For example, an instrument deployed to the internet would be accessible to multiple interviewers from any location with internet access. Copies of databases with electronic form templates, spreadsheets, or word processor-based forms for data input could be used for in-person interviewing by interviewers using laptops in the field.

**Stage 4 – Advance fieldwork and interview preparation – Factors that Affect Time Estimates**

Issue	Response	Result
4.a Does the mailing list exist in one place rather than having to be compiled from various sources?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↓    If no - time ↑
4.b Is the mailing list complete, up-to-date, and does it require minimal editing?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↓    If no - time ↑
4.c Is the mailing list in electronic or paper format?	<input type="checkbox"/> Electronic <input type="checkbox"/> paper	If electronic - time ↓    If paper - time ↑
4.d Is a pre-mailing or other advance contact planned (e.g., notification letter or advance call made prior to the initial survey mailout)? <sup>1</sup>	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑    If no - time ↓
4.e Will interviewers be working on standalone, distributed DCIs?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑    If no - time ↓
4.f What is workload of Web Services or Data Analysts supporting programming or deployment of CAPI instrument?	<input type="checkbox"/> High <input type="checkbox"/> Low	If high - time ↑    If low - time ↓
4.g Is there any kind of seasonal issue with respondent availability or level of respondent burden that may increase response time?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑    If no - time ↓
4.h Does the response process for each sampled case involve any of the following: a) multiple informants, b) directing the survey request to respondents other than the one initially to be contacted, or c) high level of review and approval before consent to interview is received?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑    If no - time ↓
4.g Other factors?		

<sup>1</sup> While this may add time early in the survey process, it often increases the initial response rate, thus reducing time spent on follow-up efforts.

**Stage 5 – Survey Fieldwork and Data Management**

Tasks	Duration (in days) <i>Low &lt; Medium &gt; High</i>		Personnel Needs	
	Working estimates	Team estimates	Required staff	Assigned staff
5.1 Conduct interviews – make initial calls, schedule callbacks, and follow up	14 < 21 > 35		Engagement staff	
5.2 Manage sample and interview records (aggregate from laptops or download from web server, as necessary, and change status or contact information in sample database(s) as necessary).	Simultaneous with conducting interviews.		ARM/CDMA staff and engagement staff	
5.3 Edit returned questionnaires for completion, consistency of responses, skip pattern errors, etc.	Done during pre-analysis data processing. Will add time depending on number of edits requiring manual review and extent of data problems encountered.		ARM/CDMA data analysis and survey staff and engagement staff to assist with edit decisions	
5.4 Code open-ended data as necessary: develop coding scheme, test reliability, code.	5 < 7 < 10		Engagement staff , ARM/CDMA staff	
5.5 Recontact respondents to clarify or verify responses or obtain additional information (If necessary/feasible).	Will add time depending on number of calls necessary, ability to recontact the respondent, etc.		Engagement staff	

**Stage 5 – Survey Fieldwork and Data Management – Factors that Affect Time Estimates**

Issue	Response	Result
5.a How many questions are included in the questionnaire?	<input type="checkbox"/> Many (over 50) <input type="checkbox"/> Moderate (25-50) <input type="checkbox"/> Few (Under 25)	If many - time ↑   If few - time ↓
5.b How complex is the question format (e.g., many skips, multi-part matrices, complex instructions, etc.)?	<input type="checkbox"/> Complex <input type="checkbox"/> Moderate <input type="checkbox"/> Simple	If complex - time ↑   If simple - time ↓
5.c Does the questionnaire contain numeric entry type questions (e.g., dollar amounts, percents, etc.) that require review and editing?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
5.d Does the questionnaire contain open-ended items that will require coding before data analysis?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
5.e Does fieldwork include in-person interviews requiring travel or other logistical complexities such as multiple phone calls to interview more than one respondent per case?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
5.f Is item nonresponse requiring follow-up or complexity of questions resulting in response errors likely to be an issue?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
5.f Other factors?		

**Stage 6 – Analysis**

Tasks	Duration (in days) <i>Low &lt; Medium &gt; High</i>		Personnel Needs	
	Working estimates	Team estimates	Required staff	Assigned staff
6.1 Write computer analysis program for analysis of survey data	1 < 7 > 14		ARM/CDMA Data Analyst	
6.2 (If sample used) Sampling weights calculated and applied	1 < 3 > 7		ARM/CDMA Statistician	
6.3 Appropriate descriptive and statistical procedures run and reviewed (Preliminary runs done as data returned from data entry. Final runs done when all data are available)	3 < 7 > 20		ARM/CDMA Data Analyst	
6.4 Nonresponse bias assessment and adjustment	2 < 7 > 14		ARM/CDMA Data Analyst	
6.5 (If necessary) Conduct data reliability assessment (May be necessary if quantitative data are being collected by the survey)	2 < 7 > 14		Engagement staff and ARM/CDMA	

**Stage 6 – Analysis – Factors that Affect Time Estimates**

Issue	Response	Result
6.a How many questions are included in the questionnaire?	<input type="checkbox"/> Many (over 50) <input type="checkbox"/> Moderate (25-50) <input type="checkbox"/> Few (Under 25)	If many - time ↑ If few - time ↓
6.b How complex is the sample?	<input type="checkbox"/> Simple random <input type="checkbox"/> Stratified sample	If simple random - time ↓ If stratified – time ↑
6.c What types of analyses are planned?	<input type="checkbox"/> Basic (e.g., frequencies) <input type="checkbox"/> Moderate (e.g., cross-tabs, correlations, etc.) <input type="checkbox"/> Complex (e.g., multiple regression, etc.)	If basic – time ↓ If complex – time ↑
6.d Is nonresponse likely to be high and a nonresponse bias analysis planned?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑ If no - time ↓
6.e. Is a data reliability assessment planned?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑ If no - time ↓
6.f Other factors?		

**Stage 7 – Report Preparation**

Tasks	Duration (in days) <i>Low &lt; Medium &gt; High</i>		Personnel Needs	
	Working estimates	Team estimates	Required staff	Assigned staff
7.1 Survey-based report text written and reviewed	1 < 14 > 28		Engagement staff and ARM/CDMA	
7.2 (If used) Survey-based graphics/tables developed	1 < 7 > 14		Engagement staff and graphics specialists	
7.3 (If used) Develop E-supplement (For information on e-Supplements, see “Guidance for Producing E-Supplements” in the EAGLE)	7 < 14 > 21		Engagement staff, ARM/CDMA, and Web Services staff	
7.4 Prepare OS&M and technical appendices (if used)	1 < 7 > 14		Engagement staff and ARM/CDMA	

**Stage 7 – Report Preparation – Factors that Affect Time Estimates**

Issue	Response	Result
7.a What type of report is this? (e.g., Chapter/Letter/Briefing/Testimony)	<input type="checkbox"/> Chapter report <input type="checkbox"/> Letter report <input type="checkbox"/> Briefing <input type="checkbox"/> Testimony	If chapter report – time ↑
7.b Will the report contain numerous or complex graphics depicting the survey results?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
7.c Will there be an e-supplement?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
7.d Will the report contain technical appendices?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
7.e Will the report contain confidence intervals or other sampling error disclosures for some or all estimates?	<input type="checkbox"/> Yes <input type="checkbox"/> No	If yes - time ↑   If no - time ↓
7.f Other factors?		