CSES Module 5: Planning Committee

Report of

Subcommittee *New Technology and Data Collection Guidelines*[1]

Rüdiger Schmitt-Beck (Chair)
Alberto Simpser
Markus Wagner
Masahiro Yamada

September 2016

Contents:

[1] The subcommittee is indebted to Michael Bosnjak (GESIS Mannheim) for helpful advice. Rüdiger Schmitt-Beck gratefully acknowledges the hospitality of the Research School of Social Sciences of the Australian National University Canberra which allowed him to concentrate on this work during a Visiting Fellowship.

(0) Executive Summary

Up until module 4 CSES has been operating under methodological recommendations designed in the mid-1990s. In view of rapid technological change in data collection methods during the past two decades the subcommittee has been assigned the task to evaluate new technologies, survey research methods, and modes of data collection and to provide recommendations as to which should be considered acceptable for use in CSES election studies. Specifically, the subcommittee was asked to consider revisions to the CSES Guidelines for Data Collection.

To address these tasks the subcommittee first inquired into **how data collection within CSES has developed** since module 1. For module 1 itself, an extensive analysis had been provided by Dave Howell and Karen Long Jusko[2]. This analysis was updated, including CSES modules 2, 3 as well as the 2015 advance release of module 4 which consisted of 17 election studies. The lead question was whether long-term trends concerning methods of CSES data collection could be identified that might possibly affect the quality of CSES data in general, and the future feasibility of CSES in particular. To that end, the most important ones of Howell and Jusko's analyses were replicated, based on the information provided in individual CSES studies' design reports.[3] The results of this analysis are documented in section (1) of this report. Data for CSES module 1 are taken directly from Howell and Jusko (2009). The outcomes of this analysis can be summarized as follows:

1. **Mode of interviewing**:
   Pure face-to-face surveying is still the standard mode of CSES, but appears to become less dominant (from about 70% down to 60%). Pure PAPI/mail-back surveying appears outdated. Pure CATI interviewing may continue as a marginally relevant mode, no trend could be identified. Mixed-mode approaches are clearly gaining in importance. More recently they have always included a Web component, typically as alternative to one or several traditional modes. At the very least this means that survey designs are becoming more complex, so that documentation and quality assessments become more challenging. As of yet there were no pure Web surveys within CSES.
2. **Study context of surveys**:
   During the past modules CSES practice has increasingly followed the guidelines, and post-election single-wave surveys have become the almost universal standard. Against this background it is advisable not to consider moving away from post-election interviewing as preferred (or even exclusive) mode of conducting CSES studies. By moving to pre-election interviewing or allowing it as an alternative the methodological integrity of CSES – which has become stronger over time in this regard – would be damaged, and over-time comparability of data impaired.
3. **Timing of surveys**:
   a. CSES studies typically start quickly after elections, and as it appears increasingly so. The guidelines are thus observed overall quite well. Where this is not the case (all modules of CSES include studies that started several months after the election) probably little can be done, since funding problems are the most likely reason for this.
   b. The share of surveys ending only long after Election Day increased over time.
   c. Likewise, the average duration of fieldwork increased. This is an ambivalent development: It is worrisome to the extent that it may for various reasons impair the

[2] David Howell & Karen Long Jusko (2009). Methodological Challenges: Resarch Opportunities and Questions for the Future. In: Hans-Dieter Klingemann (ed), *The Comparative Study of Electoral Systems*, Oxford: OUP.
[3] The subcommittee is greatly indebted to Yioryos Nardis and Dave Howell (University of Michigan) for conducting these analyses.

validity of the data, but if it is due to more thorough and careful fieldwork it may lead to improved data quality.

4. **Response rates:**
   CSES does not escape the overall trend of declining response rates in survey research, and needs to keep an eye at this development.
5. **Numbers of completed interviews**:
   The currently recommended minimum number of cases (1,000) has been reached by almost all studies of CSES modules 1 - 4. But many studies also would have met a higher target of N = 1,500, especially in CSES module 4 (2015 advance release) where the average N was higher than in previous rounds.

From these findings various **recommendations** were derived by the subcommittee. They concern the Guidelines for Data Collection (cf. section (2) of this report), the information to be collected in the Design Reports of the individual CSES surveys (section (3)) and paradata that additionally should be added to the CSES datasets (section (4)).

The most important recommendations concern the issue of survey **modes** that are to be considered acceptable for use in CSES election studies. The subcommittee recommends to maintain face-to-face interviewing as primary standard, but to allow Web interviewing, alongside paper-and-pencil/mailback and telephone interviewing, as acceptable secondary mode of data collection, if it helps to increase response rates and/or compensate for undercoverage. However, the subcommittee does not recommend at this stage to give up the requirement for all surveys included in CSES to offer adequate coverage of the target population and to be based on random sampling at all stages.

As long as these requirements are fulfilled it should also be possible to include Web surveys that are based on access panels. This means that in order to conform to the recommended revision of CSES Guidelines access panels need to be recruited offline by random sampling, and participants without access to the Internet are to be provided the necessary equipment to participate in Web surveys. At present only a small number of countries already has or is developing research infrastructures that may meet these requirements.

Opening CSES for Web surveys based on non-probability sampling seems not advisable at present since such surveys have repeatedly been shown to generate biased point estimates as well as trend estimates (that also cannot be remedied by weighting). Relationships between variables often appear not to deviate significantly between probability and non-probability samples, but it is not clear to which extent this can reliably be expected to be the case in individual surveys, as there are also studies that did show differences[4]. Moreover, even if within a country a strategy for approximating findings from non-probability surveys to those of surveys based on random-sampling would be developed, its traveling capacity to other countries (with different conditions for survey research) would need to be additionally ascertained in the context of an internationally comparative study like CSES. For Web surveys based on access panels it is further recommended to demand access panels to be rather fresh.

It is further recommended to allow the inclusion of mixed-mode surveys that include a Web component into CSES if they are a feasible way to increase response rates and/or compensate

[4] Cf., e.g., Josh Pasek (2016). When will Nonprobability Surveys Mirror Probability Surveys? Considering Types of Inference and Weighting Strategies as Criteria for Correspondence. *International Journal of Public Opinion Research* 28:269-291.

for undercoverage. However, also as part of a mixed-mode design the Web component needs to conform to the requirements outlined in the previous paragraph. Moreover, mixed-mode approaches should seek to minimize variation of modes within countries (across elections) and within surveys (across modes), as well as within modes (e.g., Web questionnaires should vary as little as possible when accessed on different devices).

Importantly, opening up data collection within CSES for Web surveys as stand-alone mode or within mixed-mode designs requires very detailed additional information that is to be provided in the Design Reports. Moreover, it is recommended to add a number of administrative variables to the CSES datasets in order to allow researchers to take possible effects of mode variations into account (at both the study and respondent level) when planning and conducting their analyses.

Further recommendations concern the **timing of surveys** and the **numbers of completed interviews** that CSES should seek to achieve. It is recommend to start interviewing no later than six months after an election and to increase the standard number of cases expected of CSES studies to N = 1,500 (which conforms to the standard of the ISSP). Most recent studies reach this goal anyway. Studies that cannot reach this aim, for instance for funding reasons, should not be excluded. But raising the project`s collective ambition could (and should) lead to an overall increased quality and utility of the CSES data.

(1) Report on development of data collection from CSES modules 1 to 4

Data Collection beyond CSES 1: An Update

(prepared for Seattle conference of CSES Planning Committee)

Rüdiger Schmitt-Beck
(schmitt-beck@uni-mannheim.de)

23 September, 2015

Purpose

- Update of analysis of CSES 1 methods presented in

  David Howell & Karen Long Jusko (2009). Methodological Challenges: Resarch Opportunities and Questions for the Future. In: Hans-Dieter Klingemann (ed), The Comparative Study of Electoral Systems, Oxford: OUP

  to include CSES 2 - 4 (2 and 3 complete, 4 advance release of 2015 with 17 studies): Are there long-term trends concerning methods of CSES data collection possibly affecting its quality in general, and future feasibility in particular?
- Basis is replication of all tables of this article from info provided in individual CSES studies` design reports. I am extremely grateful to Yioryos Nardis (University of Michigan) for preparing these tables. Only a fraction of the massive amount of information offered by their analyses can be used in the following. Data for CSES 1 is taken directly from Howell & Long`s article.
- Unless not indicated otherwise, the unit of analysis is individual CSES studies; n of cases differ due to missing data.

Topics

1. Modes of interviewing
2. Study context of surveys
3. Timing of surveys: start, end, duration of fieldwork
4. Response rates
5. Numbers of completed interviews

1. Modes of interviewing (%)

CSES	1	2	3	4
F2F	69	71	72	59
Mail/self-compl	15	7	4	0
Phone	10	10	16	12
Mix	5	12	8	23
N (studies)	**39**	**41**	**50**	**17**
F2F	65	69	71	59
Mail/self-compl	18	8	5	0
Phone	12	10	15	12
Mix	6	13	10	23
N (countries)	**34**	**39**	**41**	**17**

Note: There were no mode changes within countries and CSES rounds.

1. Modes of interviewing

- At levels of both studies and countries F2F dominates across all CSES rounds (as it should, according to guidelines). Shares are quite stable from rounds 1 to 3, but seem to decline in round 4 (but note that this is only based on advance release, not on all CSES 4 studies).

- Shares of (pure) mail/self-completion studies quite low from the beginning, and declining over time. No studies of this type in CSES 4 (advance release).

- CATI: shares always low, no clear trend over time.

- Considerable increase of mixed-mode studies in CSES 4.
- Mixed-mode studies encompass all kinds of mode combinations. Internet for the first time used within a mixed-mode study in one case of CSES 3 (DK 2007) . Three out of the five mixed-mode studies of CSES 4 include online interviewing, combined with either mail-back (Australia 2013, NZ 2011), or F2F (Greece 2012).

1. Modes of interviewing

- **A few reflections on Internet-based interviewing in CSES:** In its recommendation for the revised guidelines the Subcommittee New Technologies has suggested that Internet interviewing should be allowed within CSES if it is – in line with the guidelines` established philosophy – based on random sampling at all stages. For various scenarios of relevance for CSES this has different implications.

  - 1st scenario: Single-mode study that collects data by means of Internet-based interviewing
    As of yet no such study within CSES. Since such a study will typically be based on an access panel the sampling of cases out of the access panel as well as in particular sampling for the access panel itself should be random, and it should assure adequate coverage of the entire population. That means that sampling for the access panel has to be conducted offline, and respondents without access to the Internet have to be provided with the necessary equipment or be interviewed by means of other modes (e.g., the German GESIS panel uses self-completion/mail-back for panel members without Internet access). This implies that at present only few countries have or are developing research infrastructures that may meet these requirements.
  - 2nd scenario: Mixed-mode study that combines sub-samples interviewed via the Internet with sub-samples interviewed by other modes
    Denmark 2007 and Greece 2012 are examples for this. For the Internet-based subsamples the logic outlined for the 1st scenario should apply. This implies that very detailed information on sampling needs be be given in the Design Report. In the two cases mentioned it is difficult to tell whether they meet these requirements. (The Internet segment of DK 2007 is based on an access panel, but it does not become fully clear whether that is based on a random sample. The Internet segment of GR 2012 is based on original RDD sampling, but it is not clear how sampled cases were treated that had no access to the Internet.)
  - 3rd scenario: Mixed-mode study that offers Internet interviewing to randomly selected respondents as optional alternative to one or several other modes
    Examples are Australia 2013 and New Zealand 2012: unproblematic according to revised guidelines. But needs proper documentation and should include markers in dataset.

2. Study context of surveys (%)

CSES	1	2	3	4
Post-election	77	85	92	94
Pre-post panel	23	10	8	6
Between rounds		15		
N	**39**	**41**	**50**	**17**

2. Study context of surveys

- There is a clear trend towards an increasing dominance of pure post-election studies in CSES (in accordance to guidelines). Since CSES 3 more than 90% of all studies conducted as single-wave post-election surveys.

3.1 Timing of surveys: Start of fieldwork (days after election)

CSES	1	2	3	4
Within 2 weeks	64%	60	64	82
Within 1 month	20%	15	10	0
Later	15%	25	26	18
N	**38**	**40**	**50**	**17**
Min	*0*	*0*	*1*	*1*
Median	*6*	*6*	*6*	*4*
Mean	*23*	*29*	*35*	*30*
Max	*325*	*211*	*219*	*200*

3.1 Timing of surveys: Start of fieldwork

- In line with the guidelines` request for a timely start of fieldwork most studies begin data collection quite early – most within two weeks after the election, some more at least later in the first month following the election.
- The median starting date has moved closer to election day and the share of early starters that got going within 2 weeks after the election has increased somewhat in the 4th round, but that might be a selection effect having to do with which studies were included in the advance release.
- What appears problematic is the fact that a small but significant share of studies enter into fieldwork only quite a while after the election. In some cases this amounts to several months, so that the validity of these data as election studies can be questioned. This phenomenon appears in all rounds of CSES. It might have improved somewhat in CSES 4, but that remains to be seen, once the full dataset will have been released.

3.2 Timing of surveys: End of fieldwork (days after election)

CSES	1	2	3	4
Within 1 month	42%	37	30	12
Within 2 months	24%	25	26	41
Later	34%	37	44	47
N	**38**	**40**	**50**	**17**
Min	*9*	*5*	*12*	*18*
Median	*45*	*45*	*52*	*52*
Mean	*73*	*76*	*81*	*88*
Max	*362*	*348*	*310*	*269*

3.2 Timing of surveys: End of fieldwork

- Assuming that the validity of election-related survey data suffers if interviewing takes place long after election day, it appears worrisome that the share of studies that reach the end of fieldwork only quite late seems to have risen continuously from CSES 1 to 4.

3.3 Timing of surveys: Duration of fieldwork (days)

CSES	1	2	3	4
Up to 1 month	53%	47	52	29
1-2 months	21%	30	24	35
Longer	26%	22	24	35
N	**38**	**40**	**50**	**17**
Min	*3*	*1*	*3*	*6*
Median	*27*	*37*	*25*	*49*
Mean	*50*	*46*	*46*	*58*
Max	*218*	*191*	*204*	*129*

3.3 Timing of surveys: Duration of fieldwork

- In CSES 1 – 3 about half of all studies were finished within one month, the rest took longer. In CSES 4 the average duration of fieldwork seems to have increased. The share of studies that were finished within one month declined considerably, whereas an increased share of studies took longer than even two months.
- Interestingly this is the case although the share of F2F surveys has been smaller in the 4th round, and although, in contrast to CSES 1 – 3, there were no outliers where fieldwork took as long as 200 days.
- The increasing time for surveys to be completed may be seen as problematic, because it implies that more interviews were conducted a long time after election day. It may also indicate rising difficulties in fieldwork, notably with regard to reaching and contacting respondents. On the other hand, it may also point to overall more thorough and careful, and thus improved fieldwork, potentially leading to better data.
- Overall, it appears worrisome that several studies, especially in round 3, but to a lesser extent also in other rounds, were finished extremely quickly, within a week or less. France 2002 is an extreme case, fieldwork was completed within a single day. This raises concern about coverage bias due to insufficient contact protocols; hard to reach segments of the populations may by systematically missing in these studies. (Importantly, mode does not explain whether a study was conducted within a week or took longer: with only two exceptions (France 2002 and Israel 2003) all `one-week` studies were conducted F2F!)

4. Response rates (%)

CSES	1	2	3	4
Min	28,0	13,8	2,6	27,6
Median	62,2	55,4	48,7	38,2
Mean	61,2	54,7	48,8	44,6
Max	94,8	81,8	99,3	98,4
SD	17,2	15,3	20,1	19,7
N	**29**	**29**	**43**	**13**

4. Response rates

- This is pehaps the strongest message emerging from the longitudinal perspective at CSES 1 – 4: CSES is strongly affected by the well-known trend of declining response rates.
- While all rounds included some studies with extremely high response rates (over 80%, in three rounds even close to 100%) or very low response rates (below 30%, in two rounds even below 15%), average response rates declined continuously, and considerably (median -24%, mean -17 %).
- It may be worth considering whether extremely high reported response rates, say above 80 or even 90 percent, can indeed be accurate, and how more accurate rates could be obtained.

5. Numbers of completed interviews (N)

CSES	1	2	3	4
Min	674	582	815	967
Median	1,525	1,496	1,419	1,853
Mean	1,600	1,567	1,603	1,942
Max	4,080	2,514	4,495	4,391
N < 1,500	49%	51%	54%	29%
N	**39**	**29**	**50**	**17**

5. Numbers of completed interviews

- The minimum number of cases recommended by the current guidelines, is 1,000. The New Technologies Subcommittee has proposed considering the possibility of raising the standard for desirable minimum numbers of cases to 1,500 (following the example of ISSP).
- In all rounds of CSES most studies reached the treshold of 1,000 cases. However, there were always also studies with lower numbers of cases.
- From CSES 1 to 3 always about half of all studies reached N = 1,500 or higher. In CSES 4 this share increased to about 70%.

Conclusions

1. **Mode of interviewing**:
   Pure F2F still standard mode, but appears to become less dominant. Pure PAPI/mail-back appears outdated. Pure CATI may continue as a marginally relevant mode. Mixed-mode approaches clearly gaining in importance. More recently they have always included a Web component, typically as alternative to one or several traditional modes. At the very least this means that survey designs are becoming more complex, so that documentation and quality assessments become more challenging.

2. **Study context of surveys**:
   During the past rounds CSES practice has increasingly followed the guidelines, and post-election single-wave surveys have become the almost universal standard. Against this background it is probably not a good idea to give up post-election interviewing as preferred (or even exclusive) mode of conducting CSES studies. By moving to pre-election interviewing or allowing it as an alternative the methodological integrity of CSES – which has become stronger over time in this regard – would be damaged, and over-time comparability of data impaired.

Conclusions

3. **Timing of surveys**:
   1. CSES studies typically start quickly after elections, and as it appears increasingly so. The guidelines are thus observed overall quite well. Where this is not the case (all rounds of CSES include studies that started several months after the election) probably little can be done, since funding problems are the most likely reason for this.
   2. The share of surveys ending only long after election day increased over time.
   3. Likewise, the average duration of fieldwork increased. This is an ambivalent development: It is worrisome to the extent that it may for various reasons impair the validity of the data, but if it is due to more thorough and careful fieldwork it may lead to improved data quality. What does appear worrisome, however, is that in each round some studies were finished extremely quickly, within a week or less. It seems unlikely that such speedy surveys can deliver good coverage of the target populations. It may be necessary to emphasize the importance of ambitious contact protocols.

Conclusions

4. **Response rates:**
   CSES does not escape the overall trend of declining response rates, and needs to keep an eye at this development. In addition, it seems advisable to give special attention to studies with very low or extremely high response rates, both with regard to data quality, and by being sensitive to the possibility that especially very high reported response rates might not be accurate.

5. **Numbers of completed interviews**:
   The currently recommended minimum number of cases (1,000) has been reached by almost all studies of CSES 1 - 4. But many studies also would have met a higher target of N = 1,500, especially in CSES 4 (advance relesase) where the average N was higher than in previous rounds. Implementing the suggested new `1,500 rule` may thus lead to problems, but it should be considered anyway. Studies that cannot reach this aim, for instance for funding reasons, should not be excluded. Still, raising the project`s collective ambition could (and should) lead to an overall increased quality and utility of the CSES data.

(2) Recommended changes for Guidelines

COLLABORATOR INSTRUCTIONS FOR THE ADMINISTRATION OF THE CSES QUESTIONNAIRE:

(1) Following these collaborator instructions, this document is comprised of three sections:

))) CSES MODULE 4 QUESTIONNAIRE: ADMINISTRATIVE VARIABLES

The "Administrative Variables" section is a list of common administrative variables that, if possible, should be provided at the time data are deposited with the CSES Secretariat.

))) CSES MODULE 4 QUESTIONNAIRE: CSES MODULE

This is the CSES Module itself, a common module of survey questions for researchers to include in their national post-election survey. The CSES Module is intended to be administered exactly as it is specified in this document.

))) CSES MODULE 4 QUESTIONNAIRE: DEMOGRAPHIC VARIABLES

Collaborators are asked to provide data on background (demographic) characteristics of respondents, coded to an agreed upon set of standards as indicated in this section. There is great international variation in the ways that collaborators will go about soliciting information on the background characteristics of their respondents. The objective here is not standardization of the way collaborators ask these background questions, but instead, standardization to a common, cross-national scheme for coding each variable.

(2) The CSES Module is intended to be administered in its entirety as a single, uninterrupted block of questions, unless noted otherwise for particular questions. In most cases, the CSES Module is included as part of a larger study. For reliable comparisons to be made, it is important that any additional items investigators may wish to include do not interrupt the CSES Module.

(3) The CSES module should be administered as a post-election interview.

(4) Where the CSES module is included in a larger study, to ensure that question-ordering effects are minimized, it is most preferable for the CSES Module to be administered at the beginning of the survey instrument. Where this is not possible, collaborators should be sensitive to the effects questions asked immediately prior to the module may have.

(5) NOTES often precede the question TEXT, and provide instructions for

the administration of the item. Where no question TEXT is provided, collaborators should provide documentation of the question used.

(6) Showcards may be helpful for the administration of some questions. For this reason, a Respondent Booklet is available for download from the CSES website. The Respondent Booklet contains showcards for select questions. It is indicated in the NOTES when a showcard is available for a question.

(7) The response options that should be read to the respondent are contained in the body of the question TEXT.

(8) Where lower-case words appear in brackets [ ] collaborators should select the words that are most appropriate.

For example:
[party/presidential candidate]
...indicates that either the word "party" or the phrase "presidential candidate" should be read, but not both.

(9) Where upper-case words appear in brackets [ ] collaborators should substitute the words that are most appropriate.

For example:
[COUNTRY]
...should be replaced with the name of the country where the election was held (perhaps "Canada" or "the Philippines").

Another example:
[NUMBER OF YEARS BETWEEN THE PREVIOUS AND THE PRESENT ELECTION OR CHANGE IN GOVERNMENT]
...should be replaced with a number that indicates the amount of years that have passed between the previous election and either the current election or recent change in government.

(10) Phrases that appear in parentheses ( ) contain words that are optional - that collaborators (or their interviewers) can decide to read or not read to respondents as needed.

(11) Words in question text that are in upper-case but NOT within brackets [ ] should be emphasized by the interviewer when reading the question text.

For example, the word "COUNTRY" would be emphasized in the following question when the interviewer reads the question to the respondent:

What COUNTRY do you live in?

But in this next example, the interviewer does not emphasize the word "[COUNTRY]". Instead, this is an instruction for the collaborator to substitute the name of the respondent's country into the question text (for more information, see the eighth Collaborator Instruction above):

How long have you lived in [COUNTRY]?

(12) Interviewer instructions are available for some questions. These interviewer instructions, labeled as HELP, are intended to provide advice to the interviewers to assist in administering the question. It is also useful to discuss the interviewer instructions as part of interviewer training. The interviewer instructions, where available, appear after the question TEXT. In interviewer-administered surveys, interviewer instructions should be available to the interviewer, but not to the respondent. For example, in a computer-assisted interview, the interviewer instructions might appear on the screen in a special color, and interviewers trained to make use of those instructions as necessary, but the interviewer should NOT read the interviewer instructions to the respondent.

(13) Some response options are followed by an arrow (->) and a skip pattern instruction. If the respondent selects that response option, the skip pattern instruction after the arrow is to be executed.

(14) Respondents who volunteer the response "DON'T KNOW" (or who have REFUSED to answer a question) should be coded as such. Interviewers should accept these responses and should NOT probe for additional information or force a respondent to use one of the response options provided in the text of the question.

(15) Special care should be taken in the administration of the Vote Choice items (Q5 and Q6 question series).

Wording for the Q5 and Q6 question series, which is to record vote choice in the elections, should follow national standards.

Collaborators are invited to compare their own national instrument with other instruments of countries that are part of the CSES and look for convergence where this is possible.

For Q6 (previous election), ask about the previous national election of the same type (whether legislative or presidential). For countries where more than one institution is being currently elected on the same day (e.g. president and legislature), please

consider asking about the previous lower house election if votes have been recorded for the current lower house election.

For Q5 (current election), for countries where more than one institution is elected on the same day (e.g., president and legislature) using different votes, please ensure that all votes are supplied. Consider including all national elections having been held within three months before or after the study's data collection period.

Please ensure all vote choices are supplied as separate variables in the dataset that you deposit.

For countries where voters have two votes for the same institution (e.g. parallel and mixed member proportional systems; double ballot systems), please ensure that both/all votes are supplied.

For countries using preferential systems (e.g., STV, AV) please provide first and second preference vote.

(16) For questions asking about parties, collaborators should be advised that they may add one or several party blocs to a list of individual parties if they feel that it will be difficult for respondents to recognize individual parties.

(17) Collaborators in the Comparative Study of Electoral Systems shall adhere to the following standards of data quality:

a. Mode of interviewing: Interviews should be conducted -face-to-face, unless local circumstances dictate that telephone, Web or mail surveys will produce higher quality data. Mixed-mode surveys are acceptable to increase response rates and/or compensate for undercoverage associated with particular survey modes. In cases of mode variation as well as in cases of within-mode variation (e.g. adaption of Web surveys to multiple devices) presentation of questionnaires to respondents should be as similar as possible. All variation within surveys shall be documented in detail, and technical information on survey mode and, where appropriate, device used shall be identified in the data set for each respondent. National studies should seek to keep mode changes to a minimum to maximize comparability of their data sets across modules.

b. Timing of interviewing: We strongly recommend that collaborators in the Comparative Study of Electoral Systems conduct their interviews in the weeks following their national election. Interviewing should not start later than six months after the election. Out of concern for data quality, data collection should be completed in as timely a fashion as possible. In the event of a runoff election, interviewing shall be conducted after the first round election. The date of interview shall be provided for each respondent.

c. Placement of module in post-election questionnaire: The questionnaire module should be asked as a single, uninterrupted block of questions. We leave it to each collaborator to select an appropriate location for the module in their national survey instrument. Collaborators should take steps to ensure

that questions asked immediately prior to the questionnaire module do not contaminate the initial questions in the module. Collaborators are also free to select an appropriate place in their survey instrument to ask the turnout, vote choice, and demographic questions.

d. Population to be sampled: National samples should be drawn from all age-eligible citizens. No sampling frames with systematic undercoverage of significant population groups (such as citizens without access to the Internet) are acceptable. When non-citizens (or other non-eligible respondents) are included in the sample, a variable should be provided to permit the identification of those non-eligible respondents. When a collaborator samples from those persons who appear on voter registration lists, the collaborator should quantify the estimated degree of discrepancy between this population and the population of all age-eligible citizens. Studies based on panels or access panels are acceptable if dictated by local circumstances. In such cases the collaborator should seek to minimize the time lag between initial sampling and the CSES survey and quantify the estimated degree of discrepancy to the population of all age-eligible citizens and provide weights. Details about initial sampling must be documented.

e. Sampling procedures: We strongly encourage the use of random samples, with random sampling procedures used at all stages of the sampling process. Collaborators should provide detailed documentation of their sampling practices in all stages.

f. Sample Size: We strongly recommend that 1,500 age-eligible respondents, and under no circumstances ~~that no~~ fewer than 1,000 age-eligible respondents be interviewed.

g. Interviewer training: Collaborators should pre-test their survey instrument and should train interviewers in the administration of the questionnaire. The Planning Committee will provide each collaborator with documentation that clarifies the purposes and objectives of each item and with rules with respect to probing "don't know" responses.

h. Field practices: Collaborators should make every effort to ensure a high response rate. Investigators should be diligent in their effort to reach respondents not interviewed on the initial contact with the household and should be diligent in their effort to convert respondents who initially refuse to participate in the study. Data on the number of contact attempts, the number of contacts with sample persons, and special persuasion or conversion efforts undertaken should be coded for each respondent.

i. Strategies for translation (and back-translation): Each collaborator should translate the questionnaire module into their native language(s). To ensure the equivalence of the translation, collaborators shall perform an independent re-translation of the questionnaire back into English. Collaborators engaged in translation of the questionnaire module into the same language (e.g., Spanish, French, English, German, and Portuguese) should collaborate on the translation.

(3) Recommended changes for Design Report

Comparative Study of Electoral Systems (CSES)
Module 54: Design Report (Sample Design and Data Collection Report)

~~September 10, 2012~~

Country: Date of Election: Prepared by: Date of Preparation:

NOTES TO COLLABORATORS:

- Where brackets [ ] appear, answer by placing an "X" within the appropriate bracket or brackets.
- If more space is needed to answer any question, please lengthen the document as necessary.

Collaborator(s):

Collaborators are the contact persons for election studies that appear in the CSES dataset - they are not necessarily the parties who collected the data. These collaborators and their contact information will be listed on the CSES website.

Name: Title: Organization: Address: Telephone: Fax: E-Mail: Website:	Name: Title: Organization: Address: Telephone: Fax: E-Mail: Website:
Name: Title: Organization: Address: Telephone: Fax: E-Mail: Website:	Name: Title: Organization: Address: Telephone: Fax: E-Mail: Website:

Data Collection Organization:

Organization that conducted the survey field work/data collection:

Organization: Address: Telephone: Fax: E-Mail: Website:

Funding Organization(s):

Organization(s) that funded the data collection:

Organization: Address: Telephone: Fax: E-Mail: Website:
Organization: Address: Telephone: Fax: E-Mail: Website:
Organization: Address: Telephone: Fax: E-Mail: Website:

Archiving Organization

If appropriate, please indicate the primary location where the full, original election study dataset (not just the CSES portion) will be archived:

Organization: Address: Telephone: Fax: E-Mail: Website:

Please indicate the date when the study is expected to be available at this archive:

Study Design

1. Timing of the study that the CSES Module was included in:

[ ] Post-Election Study (started up until 6 months after election)
[ ] Post-Election Study (started more than 6 months after election)
[ ] Pre-Election/Post-Election Panel Study
[ ] Between Rounds
~~[ ] Between Rounds~~

2a. Date Post-Election Interviewing Began:

2b. Date Post-Election Interviewing Ended:

3a. Mode of interviewing for the post-election survey in which the CSES Module appeared: (If multiple modes were used, please mark all that apply.)

[ ] In person, face-to-face: questionnaire on paper
[ ] In person, face-to-face: electronic questionnaire (CAPI)

[ ] Telephone
[ ] Mail or self-completion supplement
[ ] Internet

3b. Was there a mode change *within* interviews (e.g. self-completion elements within questionnaire)?

[ ] No
[ ] Yes, please provide details:

4a. Was the survey part of a panel study?

[ ] Yes
[ ] No

4b. If the survey was part of a panel study, please describe the design of the panel study, including the date at which interviewing for each prior wave began and ended:

4c. If the survey was entirely or partly conducted via the Internet, please indicate whether it was based on an access panel (i.e., respondents were selected from a group of pre-screened panelists):

[ ] Yes
[ ] No

4d. If the survey was based on an Internet access panel, please describe the access panel (company, population [does it include persons without initial access to the Internet and how are they interviewed], method of recruiting members, ~~size~~total size of access panel, method of selecting survey respondents from the panel):

Translation

Please provide copies of questionnaires in all languages used as part of the election study deposit. For questionnaires in a language other than English, please also provide a version of each translated back into English. Note: Questions are based on those developed for the ISSP.

5. Was the questionnaire translated?

[ ] Yes, translated by member(s) of research team
[ ] Yes, by translation bureau
[ ] Yes, by specially trained translator(s)
[ ] No, not translated

6. Please list all languages used for the fielded module:

7a. If the questionnaire was translated, was the translated questionnaire assessed/checked or evaluated?

[ ] Yes, by group discussion
[ ] Yes, an expert checked it
[ ] Yes, by back translation
[ ] Other; please specify: _____
[ ] No
[ ] Not applicable

7b. If the questionnaire was translated, was the questionnaire pre-tested?

[ ] Yes
[ ] No
[ ] Not applicable

7c. If the questionnaire was translated, were there any questions which caused problems when translating?

[ ] Yes
[ ] No
[ ] Not applicable

7d. If the questionnaire was translated, please provide a list of all questions which caused problems when translating. For each question listed, describe what problems were encountered and how they were solved:

Sample Design and Sampling Procedures

8. Please describe the population that your sample is meant to be representative of:

Eligibility Requirements

9a. Must a person be a certain age to be interviewed?

[ ] Yes
[ ] No

If yes, what ages could be interviewed?

9b. Must a person be a citizen to be interviewed?

[ ] Yes
[ ] No

9c. Must a person be registered to vote to be interviewed?

[ ] Yes
[ ] No

9d. Please list any other interviewing requirements or filters used:

Sample Frame

10a. Were any regions of the country excluded from the sample frame?

[ ] Yes
[ ] No

If yes, what percent of the total eligible population did this exclude from the sample frame? _____ %

If yes, please explain:

10b. Were institutionalized persons excluded from the sample?

[ ] Yes
[ ] No

If yes, what percent of the total eligible population did this exclude from the sample frame? _____ %

If yes, please explain:

10c. Were military personnel excluded from the sample?

[ ] Yes
[ ] No

If yes, what percent of the total eligible population did this exclude from the sample frame? _____ %

If yes, please explain:

10d. If interviews were conducted by telephone, what is the estimated percentage of households without a phone? _____ %

Please explain:

10e. If interviews were conducted by telephone, were unlisted telephone numbers included in the population sampled?

[ ] Yes
[ ] No

If no, what percent of the total eligible population did this exclude from the sample frame? _____ %

10f. If interviews were conducted via the Internet, what is the estimated percentage of households without access to the Internet? _____ %

10g. If interviews were conducted via the Internet, were provisions taken to include members of the population without access to the Internet, and if so, which?

[ ] Yes
[ ] No

If yes, please explain:

If no, what percent of the total eligible population did this exclude from the sample frame? _____ %

10h~~f~~. Were other persons excluded from the sample frame?

[ ] Yes
[ ] No

If yes, what percent of the total eligible population did this exclude from the sample frame? _____ %

If yes, please explain:

10~~ig~~. Please estimate the total percentage of the eligible population excluded from the sample frame: _____ %

Sample Selection Procedures

11. Please describe, in your own words, how the sample for the study was selected. If the survey is part of a panel study and/or based on an ~~i~~Internet access panel, please also describe the original sample, from the beginning of the study.

12a. What were the primary sampling units?

12b. How were the primary sampling units selected?

12c. Were the primary sampling units randomly selected?

[ ] Yes
[ ] No

Please explain how the units were randomly selected. If the units were not randomly selected, please provide a justification for why the units were not randomly selected.

13. Were there further stages of selection?

[ ] Yes
[ ] No

13a. If there were further stages of selection, what were the sampling units at each of the additional stages?

13b. If there were further stages of selection, how were the sampling units selected at each of the additional stages?

13c. If there were further stages of selection, were units at each of these stages randomly selected?

[ ] Yes
[ ] No

Please explain how the units were randomly selected. If the units were not randomly selected, please provide a justification for why the units were not randomly selected.

14a. How were individual respondents identified and selected in the final stage?

14b. Could more than one respondent be interviewed from a single household?

[ ] Yes
[ ] No

If yes, please explain:

15. Did the sample design include clustering at any stage?

[ ] Yes
[ ] No

If yes, please describe:

16. Did the sample design include stratification?
Definition: Stratification involves the division of the population of interest according to certain characteristics (for instance: geographic, political, or demographic). Random selection then occurs within each of the groups that result.

[ ] Yes
[ ] No

If yes, please describe (please include the list of characteristics used for stratification, and in case of multi-stage selection processes the stage(s) at which stratification occurred):

17. Was quota sampling used at any stage of selection?

[ ] Yes
[ ] No

If yes, please describe:

18. Was substitution of individuals permitted at any stage of the selection process or during fieldwork?

[ ] Yes
[ ] No

If yes, please describe:

19. Under what circumstances was a household designated non-sample? Please check all that apply:

[ ] Non-residential sample point
[ ] All members of household are ineligible
[ ] Housing unit is vacant
[ ] No answer at housing unit after _____ callbacks
[ ] Other (Please explain):

20. Were non-sample replacement methods used?

[ ] Yes
[ ] No

Please describe:

21a. For surveys conducted by telephone, was the sample a random digit dial (RDD) sample?
[ ] Yes
[ ] No

21b. For surveys conducted by telephone, was the sample a listed sample?
[ ] Yes
[ ] No

21c. For surveys conducted by telephone, was the sample a dual frame sample?
[ ] Yes
[ ] No

If yes, what % list frame_____ and what % RDD_____

22. For surveys conducted by mail, was the sample a listed sample?
[ ] Yes
[ ] No

Please describe:

23. For surveys conducted on the Internet, did ~~any~~ respondents at any stage self-select into the survey?
[ ] Yes
[ ] No

Please explain:

Incentives

24a. Prior to the study, was a letter sent to the respondent?

[ ] Yes
[ ] No

(If yes, please provide a copy of the letter.)

24b. Prior to the study, was a payment sent to the respondent?

[ ] Yes
[ ] No

If yes, please describe (including amount of payment):

24c. Prior to the study, was a token gift sent to the respondent?

[ ] Yes
[ ] No

If yes, please describe:

24d. Did respondent receive an additional payment after their participation? (Do not include any payment made prior to the study.)

[ ] Yes
[ ] No

If yes, please describe (including amount of payment):

24e. Were any other incentives used?

[ ] Yes
[ ] No

If yes, please describe:

Interviewers

25. Please describe the interviewers (e.g., age, level of education, years of experience):

26. Please provide a description of interviewer training, if possible differentiating between general and study-specific components:~~:~~

26a. Content, structuration and time used for general training of interviewers:

26b. Content, structuration and time used for training interviewers in the specifics of the study within which CSES was run:

Contacts

27a. What was the average number of contact attempts made per household, for the entire sample?

27b. For households where contact was made, what was the average number of contact attempts prior to first contact?

27c. During the field period, how many contacts were made with the household before declaring it a **non-sample?**

28d. During the field period, how many contacts were made with the household before declaring it a **non-interview?**

28e. During the field period, what were the maximum number of days over which a household was contacted?

28f. During the field period, did interviewers vary the time of day at which they re-contacted the household?

[ ] Yes
[ ] No

If yes, please describe:

Refusal Conversion

29a. Were efforts made to persuade respondents who were reluctant to be interviewed?

[ ] Yes
[ ] No

Please describe:

29b. Were respondents who were reluctant to be interviewed sent a letter persuading them to take part?

[ ] Yes
[ ] No
(If yes, please provide a copy of the letter or letters.)

If yes, please describe:

29c. Was payment offered to respondents who were reluctant to take part?

[ ] Yes
[ ] No

If yes, how much?

29d. Were respondents who were reluctant to take part turned over to a more experienced interviewer?

[ ] Yes
[ ] No

29e. What was the maximum number of re-contacts used to persuade respondents to be interviewed?

29f. Were any other methods used to persuade respondents reluctant to be interviewed to take part?

[ ] Yes
[ ] No

If yes, please describe:

Interview/Survey Verification

Definition: Interview/survey verification is the process of verifying that an interview was conducted and that the survey was administered to the correct respondent, for quality control purposes.

30. Was interview/survey verification used?

[ ] Yes
[ ] No

If yes, please describe the method(s) used:

If yes, please indicate the percent of completed surveys that were verified: _____ %

Response Rate

Note: If multiple modes of interviewing were used for the post-election survey in which the CSES Module appeared, please repeat the following questions as appropriate for each of the modes used.

31. What was the response rate of the survey that the CSES Module appeared in? Please show your calculations. (If the CSES Module appeared in a panel study, please report the response rate of the first wave of the study, even if the CSES Module did not appear in that wave.)

32. Please provide the following statistics for the survey that the CSES Module appeared in. (If the CSES Module appeared in a panel study, please report the statistics for the first wave of the study, even if the CSES Module did not appear in that wave.)

A. Total number of households in sample: _____

B. Number of valid households: _____
C. Number of invalid (non-sample) households: _____
D. Number of households of unknown validity: _____

_____

E. Number of completed interviews: _____
F. Number of partial interviews: _____
G. Number of refusals and break-offs: _____
H. Number non-contact (never contacted): _____
I. Other non-response: _____

The sum of B+C+D should equal the value of A. If not, please describe why:

If statistic D (number of households of unknown validity) has a value greater than zero (0), please estimate the proportion of households of unknown validity that are valid:

The sum of E+F+G+H+I should equal the value of B. If not, please describe why:

If statistic I has a value greater that zero (0), please describe what cases fall into this category:

33. If the CSES Module appeared in a panel study, how many waves were conducted prior to the wave that included the CSES Module?

34. If the CSES Module appeared in a panel study, what was the total panel attrition between the first wave of the study and the wave that included the CSES Module? Please show your calculations.

35. If the CSES Module appeared in a panel study, please provide the number of completed interviews for the wave that included the CSES Module:

36. If the CSES Module appeared in a panel study, please provide the following statistics for panel attrition by age and education. In each cell, indicate the percent of all completed interviews in each category for the indicated wave.

Age	First wave of study	Wave that included CSES
18-25	%	%
26-40	%	%
41-64	%	%
65 and over	%	%

Education	First wave of study	Wave that included CSES
None	%	%
Incomplete primary	%	%
Primary completed	%	%
Incomplete secondary	%	%
Secondary completed	%	%
Post-Secondary Trade/Vocational	%	&
University incomplete	%	%
University degree	%	%

Post-Survey Adjustment Weights

37. Are weights necessary to make the sample representative of the populated being studied?
[ ] Yes
[ ] No

If yes, please explain:

38. Are weights included in the data file?
[ ] Yes
[ ] No

39. If weights are included in the data file, please describe in detail how the weights were constructed:

40a. If weights are included in the data file, are the weights designed to compensate for disproportionate probability of selection?
[ ] Yes
[ ] No

If yes, please describe:

40b. If weights are included in the data file, are the weights designed to match known demographic characteristics of the population?
[ ] Yes
[ ] No

If yes, please describe:

40c. If weights are included in the data file, are the weights designed to correct for non-response?
[ ] Yes
[ ] No

If yes, please describe:

40d. If weights are included in the data file, are the weights designed to correct to the official election results?

[ ] Yes
[ ] No

If yes, please describe:

41. Comparison of Completed Interviews to Population (please provide as percentages of the total):

Characteristic	Population Estimates	Completed Interviews: Unweighted Distribution	Completed Interviews: Weighted Distribution
Age			
18-25	%	%	%
26-40	%	%	%
41-64	%	%	%
65 and over	%	%	%
Education			
None	%	%	%
Incomplete Primary	%	%	%
Primary Completed	%	%	%
Incomplete Secondary	%	%	%
Secondary Completed	%	%	%
Post-Secondary Trade/ Vocational	%	%	%
University Incomplete	%	%	%
University Degree	%	%	%
Gender			
Male	%	%	%
Female	%	%	%

42. Please indicate the source of the population estimates in the prior question. English language sources are especially helpful. Include website links or contact information if applicable.

(4) Recommended changes for administrative variables

Recommendations concerning admin variables included in the CSES datasets
(Statement of New Technologies Subcommittee, 10/20/15)

The following admin variables are currently included in the CSES data set (cf. http://www.cses.org/datacenter/module4/data/cses4_codebook_part2_variables.txt):

D1022 >>> STUDY TIMING (post-election, panel, etc)
D1023 >>> MODE OF INTERVIEW (see below)
D1024 >>> A02 INTERVIEWER ID WITHIN ELECTION STUDY
D1025 >>> A03 INTERVIEWER GENDER
D1026 >>> A04a DATE QUESTIONNAIRE ADMINISTERED - MONTH
D1027 >>> A04b DATE QUESTIONNAIRE ADMINISTERED - DAY
D1028 >>> A04c DATE QUESTIONNAIRE ADMINISTERED - YEAR
D1029 >>> DAYS INTERVIEW CONDUCTED POST ELECTION
D1030 >>> A06 LANGUAGE OF QUESTIONNAIRE ADMINISTRATION

Interviewers' gender deserves special mention, because in many CSES studies this information is actually missing. In 8 countries of CSES 4 (Australia, Czech Republic, France, Greece, Ireland, Japan, Mexico, New Zealand) these data are completely missing. In Israel and the U.S. they are partly missing. We would like to emphasize the importance of this information and urge CSES teams to try to obtain this information and include them in their data sets (although realizing that in some countries this information may be difficult to obtain from survey companies). We should be conscious of interviewer's effect, especially when we ask questions concerning gender norms.

Furthermore we suggest the following additions:

D1026 >>> DATE QUESTIONNAIRE ADMINISTERED - MONTH
D1027 >>> DATE QUESTIONNAIRE ADMINISTERED - DAY
D1028 >>> DATE QUESTIONNAIRE ADMINISTERED - YEAR

CSES studies differ widely with regard to the duration of fieldwork. Information on the dates when fieldwork commenced and ended is given in the comments on these variables. We suggest to add two additional variables, at the study level, indicating (a) **how many days after the election fieldwork of the study began**, and (b) **how many days fieldwork took overall**. Although data users could of course generate these variables themselves on the basis of the information given in the documentation this additional service could decrease the threshold for taking such information into account when analyzing CSES data. That way studies could easily be distinguished by whether they immediately followed an election or were conducted only much later in the electoral cycle, even if only as a potentially interesting control (assuming that respondents may think differently about an election if it lies only a few days or several months in the past). Moreover, combined with variable D1029 it can be easily determined when within the period of fieldwork a person was interviewed which. This may be of special interest for studies with a long field period. This might be used as a proxy for (at least roughly) distinguishing between respondents that are hard and easy to reach, for instance. On that basis studies with very short field periods, that presumably include mainly easy to reach respondents, could be compared in a (somewhat) more controlled manner to studies with more extensive fieldwork.

D1023 >>> MODE OF INTERVIEW

This variable should be **replaced by the following set of new variables** to meet the increasing complexity of study designs:

D1023A: Set of multiple response variables, indicating **which mode(s) were used in study**.

```
#1. IN PERSON, FACE-TO-FACE, PAPER QUESTIONNAIRE
#2. IN PERSON, FACE-TO-FACE, ELECTRONIC QUESTIONNAIRE (CAPI)
#3. TELEPHONE
#4. MAIL SELF-COMPLETION QUESTIONNAIRE
#5. INTERNET SELF-COMPLETION QUESTIONNAIRE

#9. MISSING
```

D1023B: Set of multiple response variables, indicating **for each respondent how he or she was interviewed**.

```
#1. IN PERSON, FACE-TO-FACE, PAPER QUESTIONNAIRE
#2. IN PERSON, FACE-TO-FACE, ELECTRONIC QUESTIONNAIRE (CAPI)
#3. TELEPHONE
#4. MAIL SELF-COMPLETION QUESTIONNAIRE
#5. INTERNET SELF-COMPLETION QUESTIONNAIRE

#9. MISSING
```

D1023C: A further variable should indicate whether the respondent **self-selected** into the mode he/she was interviewed in. This only applies to single-mode surveys, i.e. where D1023A has only one value for the survey.

```
1. RESPONDENT SELF-SELECTED INTO MODE
2. RESPONDENT WAS ASSIGNED TO MODE, NO SELF-SELECTION

9. DOES NOT APPLY
```

Moreover, wherever possible (especially when interviews are computerized this should be easy) **response time** should be included at the level of individual respondents. The easiest way to do this is to simply take the duration of each single interview. This will not allow comparisons across studies, but at least within studies (for instance to identify respondents that satisficed in Web interviews). It would of course be preferably to get precise readings of the duration of the CSES block, excluding demographics which may vary nationally, and non-CSES questions in cases where CSES is included in larger studies. But that is probably not feasible. In any case, however, an additional variable should be added indicating whether the study was conducted as a **stand-alone CSES study or as part of a larger study**.

In sum, this leads to the following list of suggested additional admin variables:

Study-level variables:

(1) Time until start of fieldwork (number of days after election)
(2) Duration of fieldwork overall (number of days)
(3) CSES run as stand-alone study or as part of larger study (dummy)
(4) Mode/s (multiple response)

Respondent-level variables:

(5) Response time (duration of interview in seconds)
(6a) Mode(s) of interview (multiple response)
(6b) Self-selection into mode (dummy)