

ESTIMATION OF PROXY, TIME AND CONTENT EFFECT FOR THE HEALTH AND ACTIVITY LIMITATION SURVEY

Éric R. Langlet, Statistics Canada
R.H.Coats building, 15R, Tunney's Pasture, Ottawa, Canada, K1A 0T6

Key Words: disability, false positives, false negatives, proxy effect, time effect, content effect

1. Overview of the problem

The household component of the 1991 Statistics Canada Health and Activity Limitation Survey (HALS) collects information on the nature and severity of disabilities, and information on the barriers which disabled persons face in the conduct of their daily activities. Two questionnaires are used, one for the adults aged 15 and over and one for the children aged 14 and under. This paper will focus on the adult population.

The 1991 survey selected a sample among individuals who reported a limitation in the Census ("yes" sample) and a sample from those individuals who did not ("no" sample). Even though the probability for someone in the "no" sample of being limited to the 1991 HALS is relatively small (the false negatives), individuals limited to HALS coming from the "no" sample represent more than 50% of the HALS target population. The sampling of the "no" population, although necessary for bias concerns, is however very costly given the size of this population. The sampling fraction in the 1991 HALS is about 10 times smaller in the "no" population compared to the "yes" population. This feature has a serious effect on the variance of the estimates. The false negatives can be due to three factors. First, the Census and the HALS definitions of disability are different (content effect). Second, there is a delay of 3 to 6 months between the Census and the HALS data collection (time effect). Third, the HALS selected respondent might be different than the household respondent who completed the Census questionnaire (proxy effect). It was never possible to distinguish between these three effects since the Census respondent could not be identified. Therefore, whether the selected respondent also completed the Census questionnaire is unknown.

For the 2001 Census, new screening questions more in line with the HALS limitation definition will be used. The degree of correspondence between the two definitions was measured in 1999 through a test involving respondents from the 1998 National Census Test (NCT). The correspondence between the two definitions was indeed greater with the new questions. The new questions are more inclusive and produce a lower false negative rate. However, it was not clear from this test whether the "no" sample could be eliminated due to the high disability rate observed in

the test compared to 1991. Because of this high disability rate, the percentage of false negatives was also high compared to 1991. Nonetheless, for practical reasons, it was decided to select only a "yes" sample for 2001. Moreover, the disability definition used for the new survey will be the one of the Census. The questions used in the new HALS survey will only be used as qualifiers of the disabled population as defined in the Census. Among other things, these qualifiers will define the nature and the severity of the disability. This approach, in addition of the obvious cost benefit, will be useful to develop other surveys since only two questions are now used to define the disabled population. This is a totally new approach that needs to be evaluated carefully, which is one of the purposes of the 2000 HALS pilot test.

In order to prepare the 2001 survey, a pilot test is currently being conducted. The test imitates the real survey process where, in the first phase, the Census questions are completed by a household respondent through self-enumeration for all household members. In the second phase of the pilot test, a sample of respondents from the first phase, including a "yes" and a "no" sample, will be selected and interviewed with the new HALS questionnaire. Since the second phase is conducted 4 to 5 months after the first phase and by possibly a different respondent than the one in the first phase, some individuals will move in and out of the target population. These movements can be due to either a time or proxy effect. The answers to the filter questions in the first phase can also be compared to the modified version of the Activity of Daily Living (ADL) questions in the second phase (the ADL questions were defining the 1991 HALS target population). Changes between these two sets of answers will be associated to either a content, proxy or time effect. It will also be possible to estimate what portion of the target population will be missed by not selecting a "no" sample and it will also be possible to provide some comparison points between the new HALS survey and the 1991 HALS.

2. The 1991 HALS survey

The 1991 HALS survey is a post-censal survey, that is a survey that uses Census microdata to identify the population of interest. A post-censal survey also integrates the survey operations with those of the Census and is conducted shortly after the Census to preserve the currency of the Census information. The post-censal survey approach has proved in the past to

be a cost-effective means of collecting data for a relatively small and scattered subgroup of the population and for producing small domain estimates. The Census microdata constitutes a good survey frame containing for each person a large amount of information that can be used in the sampling design and estimation methods. Other examples of post-censal surveys in Statistics Canada include the 1986 HALS and the 1991 Aboriginal Peoples Survey.

In the household component of the 1991 HALS, the survey questionnaire begins with a screening section where the nature, cause and severity of disability are defined. Other sections collect information on the impact of disability in the everyday life of the respondent in areas such as employment, education, transportation, social network, leisure, accommodation and finances.

The main problem of a post-censal survey is the identification of the population of interest through the Census questionnaire, namely the Census filter questions. This problem is summarized in the following table.

		Post-Censal survey screening	
		Screened-in	Screened-out
Census filters	positive	True positives	False positives
	negative	False negatives	True negatives

The Census, through a limited number of filter questions, allows the classification of respondents to “positive” and “negative” groups, also called “yes” and “no” populations. Ideally, most people classified as positive should belong to the target population and vice-versa. Then, during the post-censal survey, the interview starts with a set of screening questions, usually much more complex than the Census questions, which allow a more precise determination of whether the person belongs to the post-censal survey target population. The efficiency of a post-censal survey is a function of the concentration of the population in the table diagonal, which implies few false positives and few false negatives. The conformity of the two classifications depends mainly on the complexity of the concept measured. Typically, disability is a concept difficult to measure since it relies for a good part on the perception of the respondent.

For the 1991 HALS, Census questions on activity limitations and long-term disabilities are included on the “long form”, which is completed by a one in five households across Canada. These Census filter questions are used as a stratification factor improving the efficiency of identifying the target population in HALS. Any respondent who indicates a “yes” answer to at least one of the two filter questions will be considered as “positive” and all others as “negative”.

A sample from both groups is taken and selected persons are interviewed using the HALS questionnaire. An initial screening section identifies disabled persons who are then asked, in subsequent sections, questions concerning the barriers faced by disabled persons in their daily activities. Screening for the adults is carried out using a modified version of the Activities of Daily Living (ADL) developed by the World Health Organization in 1982 for physical disabilities. Mental disability screening questions are also included. Any respondent having a positive answer to one of the screening questions will be considered as disabled. The following table presents the relationship between the Census disabled population and the 1991 HALS disabled population.

Adult population	Post-Censal survey		Total
	Disabled	Not disabled	
Census Positive	80%	20%	10%
Negative	47%	3%	90%
	53%	97%	
Total	17%	83%	100%

The percentage of positives in the Census is 10% as compared to 17% for the HALS disabled population. Almost 80% of the adults classified as positive in the Census were classified as disabled in HALS (true positives). However, even if only about 10% of the persons classified as negative were actually disabled (false negatives), this group represents almost 53% of the HALS disabled population (high concentration of mildly disabled, however). Consequently, sampling only from the “yes” stratum would seriously bias the survey results. The sampling from the “no” stratum, although necessary for bias concerns, is however very costly. Given the very large size of the “no” population compared to the “yes” population, the sampling rate in the “no” is about 10 times smaller than in the “yes”. This feature, although necessary for cost-efficiency concerns, has a serious impact on the variance of the estimates.

3. The 1999 HALS test

In preparation for the 2001 HALS, new Census filter questions, more in line with the HALS definition of disability were considered. The former and new filter questions are presented in the Appendix. The first new question was a summary of the ADL questions and the second one was rephrased by using the term “activity reduction” as opposed to “activity limitation”. Moreover, for each question, a new category “sometimes” was added. These two new questions were found to be much more inclusive than the old ones and

also closer to the HALS definition of disability. It was therefore expected that these new questions would produce a higher rate of false positives but more importantly a smaller rate of false negatives. A substantial reduction of the false negatives could possibly eliminate the need to sample the “no” stratum. Also, even if using the new questions would produce a non-negligible false negative rate, it is possible that the Census questions would identify an important subset of the survey-disabled population, in particular individuals who are moderately and severely disabled. Typically, the group of mildly disabled individuals, which the Census may not identify, is of less interest to the users in general.

The expected increase of false positives with the new questions is not a big concern especially if no sampling is required in the “no stratum”. Under this scenario, substantial savings could be made by only increasing the “yes” sample appropriately. Even if sampling is still required in the “no stratum” for 2001, a higher degree of correspondence between the Census and the HALS screening questions could lead to a more efficient sampling plan. Under this second scenario, savings would be much less, however.

Results showed a stronger relationship between the new filter questions and the HALS screening questions compared to the old ones. Fewer false negatives and more false positives were observed with the new filter questions. The new filter questions screen-in more disabled individuals for all levels of severity and the milder the disability the more pronounced is this advantage. The two groups of positive respondents, that is the true positives and the false negatives, contain proportionally more mildly disabled individuals for the new filter questions than the old ones.

The results clearly showed the superiority of the new Census screening questions over the old ones in terms of selecting a larger portion of the target population and in terms of missing a portion of the target population which is less critical (the mild disabled population). However, the overall disability rate in this test was much higher than the one in 1991. As a result, the proportion of false negatives, even with the new filter questions, is much higher than in 1991. Therefore, this study did not permit to determine whether individuals with negative answers to the Census filter questions should be sampled or not. This substantial increase in the disability rate came mainly from the mildly disabled individuals. This group more than tripled compared to 1991.

A number of reasons can explain the large difference observed between the 1991 HALS and the 1999 test. The difference between the questionnaires is believed to be the major cause of this increase. The 1999 test only contained the screening questions to identify the disabled population, the type and severity

of the disability. The 1991 HALS included some follow-up questions for each difficulty reported as well as very extensive follow-up questions for screened-in individuals. These questions are known to be somewhat awkward to ask for individuals having very marginal limitations. The interviewer having gone through this process a few times might change his behavior for individuals appearing to be only marginally disabled. Hence, one or two initial “yes” answers by the respondent might be converted to “no” answers by the interviewer following clarifications with the respondent. It is not completely clear by doing this whether the interviewer reduces artificially the disability rate or whether the interviewer has a better understanding of the survey concepts after a few interviews and does the right thing thereafter. In other words, the “cost” of being screened-in (in terms of additional questions being asked) seems to affect the disability rate. This corresponds to an earlier finding by Binder and Morin (1988).

Among other differences between the two surveys, is the fact that the 1991 HALS was conducted by paper and pencil (PAPI) rather than by Computer Assisted Telephone Interviewing (CATI) in the 1999 test. Whether this fact could influence the disability rate is another question. Another element is the fact that the population covered by the 1998 NCT sites is obviously not representative of the Canadian population. Finally, as the population is aging and as the concept of activity limitation becomes more and more accepted in the community, there is a natural increase in the disability rate that is expected from year to year. For instance, there was an increase of 18% in the rate of positive responses to the Census filter questions between the 1991 and the 1996 Census. There is still, however, a major increase in disability rate even with age standardized rates.

Because of the very high disability rate in the 1999 test compared to 1991, a second phase was conducted. In this second phase, a sample of about 1,000 screened-in individuals in the first phase was selected. Since the increase in the disability rate came mainly from the “mild disabled”, this group was largely over-sampled in the second phase. Respondents selected were assigned the full 1991 HALS selection portion of the questionnaire. Because of the time constraint, this interview was conducted by telephone using a paper and pencil (PAPI) questionnaire. In this questionnaire, for each difficulty reported, respondents were asked at what age they first had the difficulty, what was the main condition or health problem which caused the difficulty and what was the cause of the condition. For the very mildly disabled individuals, it was suspected that asking the follow-up questions would cause some of them to realize that their difficulty was not serious enough to be reported. Individuals not reporting a

difficulty anymore in Phase 2 were asked a follow-up question on the reason for change compared to Phase 1. Since all questions in Phase 1 were also asked in Phase 2, this second phase can also be used to determine the stability of the responses.

Many changes from “disabled” in Phase 1 to “not disabled” in Phase 2 were observed. In fact the disability rate went down from 42% in Phase 1 to about 30% in Phase 2. This figure, however, assumes that all individuals who answered “no” in Phase 1 would still answer “no” in Phase 2. Among selected respondents in Phase 2, a few changes were observed in their patterns of “yes” and “no” to the different screening questions. Therefore, it is quite conceivable that a certain number of individuals saying “no” to all questions in Phase 1 would answer at least one “yes” in Phase 2. In order to estimate this number, a logistic regression predicting the answer “yes” or “no” in Phase 2 as a function of the disability score obtained in Phase 1 (the score is based on the patterns of “yes” and “no” to the different screening questions), the age of the respondent and also the number of “completely unable” (number of activities for which the respondent is completely unable to perform) was done. By projecting the model for individuals who were “no” in Phase 1 (score of 0), a number of individuals who would change from “no” in Phase 1 to “yes” in Phase 2 was obtained. Applying this model to all individuals in Phase 1, a disability rate around 34% was obtained.

Observation of the Phase 2 data collection revealed that in general, both the respondent and the interviewer had a better understanding of the questions. Four major reasons for changing from a “yes” in Phase 1 to a “no” in Phase 2 were observed. About 25% of them indicated that the difficulty was not serious enough to be reported. Another 25% mentioned that they never had a limitation. An additional 15% mentioned that they did not have a limitation at the last interview. These three reasons implicitly assume a misunderstanding from either the respondent or the interviewer in Phase 1. Finally, another 25% mentioned an improvement of the situation since Phase 1. This reason suggests that the difficulty reported in Phase 1 was not a difficulty that lasted or that was expected to last 6 months or more, as specified in the questionnaire. Since Phase 1 was conducted in April and Phase 2 in July, seasonal problems such as arthritis for instance, are subject to this type of change. Phase 1 was conducted in a CATI mode while Phase 2 was conducted by PAPI. Apparently, interviewers that have been exposed to CATI for a long period tend to prefer PAPI to CATI. The reason mentioned was the fact that interviewers had to use more their judgement with PAPI. With CATI, the complete flow of the questionnaire is decided for them. Is CATI more prone in some cases to

coding errors than PAPI? Without having evidence of this being true, this could be a possibility.

It is probable that, if the full 1991 HALS questionnaire had been administered to the respondent, the disability rate would have decreased even further. This test would have been, unfortunately, too costly. This test is however replaced by something similar in the 2000 HALS pilot test. More findings on the 1999 test can be found in Langlet (1999).

4. The 2000 HALS pilot test

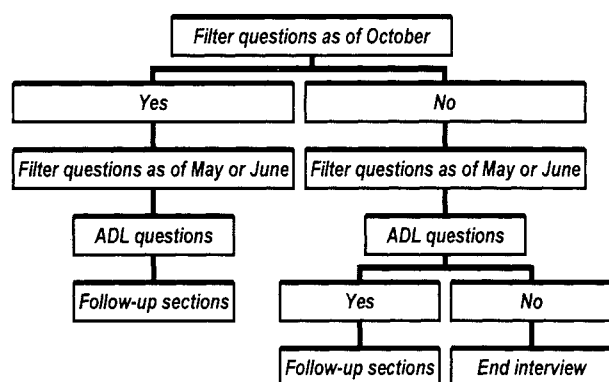
The 2000 HALS pilot test is conducted in two phases. The first phase was conducted in May and June as a Labor Force Survey (LFS) supplement. After the LFS interview, the household respondent was given a short questionnaire consisting of the two new Census filter questions to be completed for all household members.

The HALS pilot test is subdivided into two samples in the second phase conducted in October 2000. The larger portion of the sample (about 14,000 individuals) will be used to evaluate what is missed by not selecting a “no” sample in the Census in terms of numbers and characteristics as well as to evaluate the proxy, time and content effect for changes between both phases. This sample will have both Phase 1 respondents who answered “yes” and “no” to the filter questions. The second portion of the sample (about 1,000 individuals) will be used specifically to estimate parameters for the HALS 2001 sampling plan, namely the proportion of individuals who will answer “yes” to the filter questions in the Census and who will answer “no” to the same filter questions in HALS. This portion of the sample was given the same questionnaire flow as the one planned for 2001. This second sample was created because it is suspected that the questionnaire flow could have an incidence on the disability rate. This sample will only contain respondents who answered “yes” to the filter questions in Phase 1. Both samples will also be used to test the new 2001 HALS questionnaire. The rest of this paper will focus on the first sample.

4.1 Sample description and possible analyses

In the first phase of the pilot test, the household respondent in Phase 1 completed the Census filter questions for all household members. Then, in Phase 2, the same filter questions are asked to the selected respondent (with the October reference period). The same filter questions were also asked to the respondent as if the respondent were answering for Phase 1 (May or June reference period). These screening questions are asked again since the selected respondent in Phase 2 may not be the same as the household respondent in Phase 1. Therefore, their answers with respect to the filters in Phase 1 may differ. These additional filter

questions will be used to distinguish between the proxy and time effect. This new set of filter questions will be asked irrespectively of the answers given to the first set of filter questions. After the two sets of filter questions being asked, the interview will proceed with the ADL screening questions for the Phase 2 reference period (more precisely, a modified version of the 1991 screening questions), irrespectively of the answers provided to the two sets of filter questions. Respondents will have to complete the follow-up sections if a positive answer is given to either the Phase 2 filter questions or the ADL screening questions as of Phase 2. Since the follow-up sections refer to a current condition or health problem, individuals not being selected through the filter questions as of Phase 2 or the Phase 2 ADL questions will not be asked the follow-up sections. Schematically:



Considering all combinations of answers to the filter questions (“yes” or “no” for the three sets of filter questions, including Phase 1), the set of ADL questions (a “yes” to any of the ADL questions is a “yes”, otherwise it is a “no”) and all five combinations of possible respondents in Phase 1 and 2 with respect to proxy or non-proxy interview (proxy in Phase 1 and same proxy in Phase 2, proxy in Phase 1 and different proxy in Phase 2, proxy in Phase 1 and non-proxy in Phase 2, non-proxy in Phase 1 and proxy in Phase 2, non-proxy in Phase 1 and non-proxy in Phase 2), 80 combinations would be obtained. In this paper, the focus will be on the most frequent combinations with respect to proxy and non-proxy interviews, that is the 16 combinations for which a **proxy respondent is used in Phase 1 and the selected respondent is used in Phase 2** (non-proxy in Phase 2). This should be sufficient to provide an idea of the analyses that will be done.

Different analyses can be done on this pilot test. The first analysis compares filters in Phase 1 to filters in Phase 2. This analysis is done in section 4.2. In this analysis, given that only a “yes” sample will be used in

2001, the false positives represent a loss in the target population for 2001. In 2001, this net loss can be compensated by post-stratification of the survey weights to the Census totals (Phase 1 in the pilot test). The changes of answers in the two sets of filters between Phase 1 and 2 can be due to either a proxy or a time effect.

The second analysis compares the filters in Phase 1 to the ADL questions in Phase 2 (similarly to the 1991 survey). This is covered in section 4.3. In this analysis, the false negatives provide an estimate of the portion of the 1991 target population not covered in 2001. This estimate is in fact very crude since the HALS selection section containing the ADL questions is largely modified for 2001 compared to 1991. It will answer the following question: what disability rate would be observed in 2001 had a “no” sample also been selected as in 1991? The pilot test was also designed to compare the characteristics of this uncovered target population to the ones of the target population that will be covered (the true positives). Any change between phases, false positives or false negatives, can be due to either a content, proxy or time effect.

4. 2 Estimation of change between Phase 1 and Phase 2 filter questions

Changes between Phase 1 and Phase 2 filter questions are problematic for the main survey. This is because Phase 1 filter questions (the Census filter questions for the main survey) are used to select the sample but the target population is defined according to Phase 2 filter questions (filter questions as of October). For the main survey, only those with a “yes” to the Phase 1 filters will be selected. The fact that some of them will answer “no” in Phase 2 will result in a net loss of the target population. Also, had we interviewed the “no” in Phase 1, some of them would have answered “yes” in Phase 2 for the same filter questions. This gain in the target population will not be observed for the main survey but can be estimated from the test. Changes in the filter questions between both phases can be attributed to either a proxy or a time effect. These changes are highlighted in Table 1 with the corresponding effects.

Table 1: Changes between Phase 1 & Phase 2 filters

	Phase 1 filters	Phase 2 filters	Phase 2 filters as of Phase 1	Effect
False +	Y	N	Y	Time
	Y	N	N	Proxy
False -	N	Y	Y	Proxy
	N	Y	N	Time

The columns “Phase 1 filters” and “Phase 2 filters” refer to the answers to the filter questions in Phase 1

and 2 respectively. The column “Phase 1 filters in Phase 2” refers to the answers provided in Phase 2 to the filter questions as of Phase 1.

The **proxy effect** is defined in this paper as a change between both phases due to the fact that a proxy reported the information for the selected respondent in the first phase and the selected respondent reported the information in the second phase. Such a change can occur because a proxy respondent might have a different perception than the person concerned. This situation corresponds to the second and third row of Table 1. In these situations, a change is reported between Phase 1 and 2 but according to the selected respondent no change occurred. Therefore, excluding possible memory errors from the selected respondent, had this respondent also completed Phase 1, no change would have been observed.

The **time effect** is a legitimate change between the two phases due to the improvement or the deterioration of the condition of the selected respondent. This corresponds to the first and fourth row in Table 1. For these cases, a change is observed between both phases and both the proxy and the selected respondent give the same answer with respect to the Phase 1 filters.

The proxy effect can be altered by memory errors from the selected respondent. For rows 2 and 3 of Table 1, it is possible that the selected respondent did not report any change between both phases simply because he forgot what his situation was in Phase 1. Had he remembered properly, it is possible that he would have reported a change. This situation would now correspond to a time effect as opposed to a proxy effect. Hence, there is a portion of the proxy effect that is really due to a time effect. This can be estimated by the percentage of the times that the selected respondent contradicts himself with the answer provided in Phase 1 for situations in which the selected respondent was interviewed in both phases. An adjusted proxy and time effect can then be calculated accordingly.

It should be noted that this paper is mainly concerned with the estimation of change between both phases. However, the proxy effect also exists where there is no change between both phases. This situation occurs when both the proxy and the selected respondent give the same answer for the Phase 1 and 2 filter questions. However, in Phase 2, the selected respondent reports a change between the two phases. In this situation, excluding memory errors from the selected respondent, no change is observed between the two phases simply because the proxy has a different perception than the selected respondent. This component could also be included in the estimation of an overall proxy effect. This situation, however, is less problematic since it results in no gain and no loss in the target population. Therefore, this type of proxy effect will be ignored in this paper.

The false positives represent a loss in the target population due to a change between answers to filter questions in both phases. This can be used to estimate the loss expected in 2001. If a “no” sample was selected in 2001, the false negatives could theoretically compensate for that loss. If both totals were roughly the same, a post-stratification to the Census totals would compensate for that loss. A post-stratification to the Census totals will make the disability rate in HALS the same as the one in the Census within certain post-stratification classes.

Let's denote by P_{YN}^* and P_{NY}^* the memory adjusted proxy effect for “yes-no” and “no-yes” changes respectively and by T_{YN}^* and T_{NY}^* the memory adjusted time effect for “yes-no” and “no-yes” changes respectively. If $P_{YN}^* \approx P_{NY}^*$ and $T_{YN}^* \approx T_{NY}^*$ holds, the false positives will balance the false negatives.

However, it is conceivable that respondents have a tendency to omit details when they answer for someone else. This should be reflected by a higher rate of “no-yes” proxy effect than “yes-no” proxy effect for interviews where a proxy is used in the first phase and a non-proxy is used in the second phase ($P_{NY}^* > P_{YN}^*$). On the other hand, for interviews where a non-proxy is used in the first phase and a proxy is used in the second phase (not shown in this paper), the opposite relation should hold. However, this situation is relatively rare since an attempt is always made to contact the selected respondent in the second phase.

With respect to time effect, if there is any such thing as a seasonal effect for disability, the disability rate in October could be higher than the one in May or June. Arthritis problems for instance, are subject to seasonality. However, this would probably be more obvious if Phase 2 was conducted in January for instance. This means that, with respect to time effect, more “no-yes” combinations than “yes-no” combinations could be observed ($T_{NY}^* > T_{YN}^*$).

As a result, more false negatives than false positives should be observed. This means, for the main survey, that even a post-stratification to the Census totals would be an insufficient adjustment to reflect the true disability rate at the time of the survey. If these assumptions are true, a one-phase survey conducted in October by non-proxy interviews would likely produce a higher disability rate than a survey conducted in May or June by proxy interviews most of the time. The pilot test will be able to test this hypothesis. The post-stratification to the Census totals will certainly be better, however, than not adjusting at all for the loss due to the false positives.

4.3 Estimation of change between Phase 1 filter questions and Phase 2 ADL

The changes between the Phase 1 filter questions and the Phase 2 ADL questions are of special interest. In the 1991 HALS, a false negative or a false positive could be due to either a proxy, time or content effect (Census filter questions vs. ADL questions). It was never possible to distinguish between these three effects. The pilot test can be used to obtain some measures of these effects. The changes of answers between the Phase 1 filters and the Phase 2 ADL questions are shown in Table 2. The Phase 2 ADL questions refer to the modification of the 1991 HALS screening questions, which were defining the 1991 HALS target population.

Table 2: Changes between Phase 1 filters & Phase 2 ADL questions

	Phase1 filters	Phase2 filters	Phase2 filters as of Phase1	Phase 2 ADL	Effect
False +	Y	Y	Y	N	content
	Y	Y	N	N	content +proxy
	Y	N	Y	N	time
	Y	N	N	N	proxy
	N	Y	Y	Y	proxy
False -	N	Y	N	Y	time
	N	N	Y	Y	content +proxy
	N	N	N	Y	content

The **content effect** can be easily identified by a change between the Phase 2 filters and the Phase 2 ADL questions (columns 2 and 4). These two columns refer to the same time period and are completed by the same respondent (therefore excluding possible time and proxy effect).

A **“pure” content effect** is observed when, in addition, both the selected and the proxy respondents are in agreement with respect to the Phase 1 filters (rows 1 and 8). Since this table shows only changes between the Phase 1 filters and the Phase 2 ADL, this implies that answers to all three sets of filters coincide.

A **content effect combined with a proxy effect** is observed when, in addition to a difference in columns 2 and 4, there is disagreement between the proxy and selected respondent with respect to the filters in Phase 1 (rows 2 and 7). In this circumstance, excluding memory errors, no change between the Phase 1 filters and the Phase 2 ADL would have been observed had the selected respondent also completed Phase 1.

A **time effect** is observed when there is a change between the Phase 1 and 2 filters according to both the proxy and the selected respondent, and when the

answer to the Phase 2 filters coincide with the one given to the Phase 2 ADL (rows 3 and 6).

A **“pure” proxy effect** occurs when a change is observed between the Phase 1 and 2 filters but not according to the selected respondent and when both the answers to the Phase 2 filters and the Phase 2 ADL correspond (rows 4 and 5). Again, omitting possible memory problems from the selected respondent, no change between the Phase 1 filters and Phase 2 ADL would have been observed, had the selected respondent also completed Phase 1.

Again, the counts of proxy effect could be adjusted for possible memory problems. That is, a portion of the proxy effect (Y-N-N-N and N-Y-Y-Y combinations) should be moved to a time effect (Y-N-Y-N and N-Y-N-Y combinations) according to the portion of the cases where the selected respondent contradicts himself when interviewed in both phases. Similarly, a portion of the combined proxy-content effect should be moved to a pure content effect. That is, a portion of the Y-Y-N-N and the N-N-Y-Y combinations in Table 2 should have been Y-Y-Y-N and N-N-N-Y combinations, had the selected respondent remembered properly his Phase 1 situation.

An additional effect, not directly measured in this pilot test, causing changes between both phases is the context effect. Indeed, the answers to the filter questions could be sensitive to the context of the survey. This result was found in Binder and Morin (1988). The filter questions in Phase 1 are asked directly after the LFS interview. First, it is not clear whether answers to these filters would be similar to the answers provided to the same filters in the middle of the Census questionnaire. Second, it is quite possible that people would be more likely to give a positive answer to these filters in the context of an activity limitation survey than in the LFS or the Census. An indication of this happening could be the observation of more “no-yes” than “yes-no” contradictions to the Phase 1 filters when the selected respondent is interviewed in both phases.

5. Conclusion

The main problem of a post-censal survey is the identification of the population of interest through the Census questionnaire. Since the amount of questions on the Census form to identify the disabled population is usually relatively limited, differences in the populations identified will always persist with a detailed activity limitation survey such as the 1991 HALS. This fact implies almost inevitably the sampling of the Census “no” population. In fact, usually the most detailed the survey questionnaire, the most people with disability will be identified. The sampling of the “no population” is however very costly given the relative sample size required in the “no population” compared to the “yes

population”, as seen in the 1991 HALS (more than three times higher).

The 1999 HALS test attempted to use new Census filter questions identifying a larger number of individuals limited in their activities and questions more in line with the concepts of the 1991 HALS. These new filter questions were expected to produce a larger proportion of false positives but more importantly a smaller proportion of false negatives. The results clearly showed the superiority of the new Census filter questions over the old ones in terms of selecting a larger portion of the target population and in terms of missing a portion of the target population which is less critical (the mildly disabled population). However, because of the very high rate of false negatives, this study did not permit to determine whether individuals with negative answers to the Census screening questions should be sampled or not. Nonetheless, for practical reasons, it was decided to select only a “yes” sample for 2001 and to use the Census definition of disability in HALS. This is a totally different approach than the one used in 1991. In fact, the approach is so different that the 2001 HALS will have a different survey name. The impact of this new approach is one of the purposes of the 2000 HALS pilot test.

In this pilot test, two samples were designed. The larger one will be used to measure the different components of change between the two phases. Comparing changes in terms of answers to the filter questions in both phases, it will be possible to distinguish between proxy and time effect. Knowing who completed each phase and re-asking in Phase 2 the filter questions for both the Phase 2 and the Phase 1 reference periods will make this decomposition possible. Comparing changes in terms of answers to the filter questions in the first phase and ADL questions in the second phase, it will be possible to distinguish between content effect (Census filter questions vs. ADL questions), the proxy and time effect. Before this pilot test, it was never possible in HALS to distinguish if a false positive or a false negative was due to one of these three factors. The pilot test will permit to estimate these different components of change. This sample will also be used to compare the characteristics of the false negatives and true positives, the two former groups forming the 1991 HALS target population. These analyses will permit to estimate the size and characteristics of the target population not covered by selecting only a “yes” sample. Some comparison points between the 2001 HALS and the 1991 HALS will be therefore possible.

The second sample, which will only contain adults who answered “yes” in the first phase, will be subject to a similar questionnaire flow than the one expected for 2001. This sample is used mainly to estimate sampling

parameters for 2001. In particular, a precise estimate of the proportion of false positives needs to be obtained for 2001. Using only the first sample for this estimation could be problematic due the modification of the questionnaire flow for this sample compared to the planned questionnaire for 2001. It is believed that the interviewer effect could alter the answers to the filter questions in Phase 2 when comparing one sample to the other. The pilot test will be able to measure this effect, if present.

An effect not directly measured in the pilot test is the context effect. It is quite conceivable that respondents would be more likely to answer “yes” to filter questions in the context of an activity limitation survey than in the context of the Labor Force Survey. More “no-yes” than “yes-no” contradictions to the Phase 1 filters when the selected respondent is interviewed in both phases would tend to confirm this assumption.

Since the data collection of Phase 2 is currently being conducted, no results are available yet. These results should be available in the spring of 2001.

References

Binder D. A. and Morin J. -P. (1988). Use of questions on activities of daily living to screen for disabled persons in a household survey. *The Canadian Journal of Statistics*, Vol. 16, Supplement, 143-156.

Langlet, E.R (1999). *Comparison of Screening Strategies for the Health and Activity Limitation Survey*. Proceedings of the Section on Survey Research Methods, Vol. II, American Statistical Association, 560-565.

Appendix

Former version of the Census filter questions

ACTIVITY LIMITATIONS	At home? 03 <input type="radio"/> No, not limited 04 <input type="radio"/> Yes limited
18. Is this person limited in the kind or amount of activity that he/she can do because of a long-term physical condition, mental condition or health problem: See Guide	At school or at work? 05 <input type="radio"/> No, not limited 06 <input type="radio"/> Yes, limited 07 <input type="radio"/> Not applicable In other activities, e.g. transportation to or from work, leisure time activities? 08 <input type="radio"/> No, not limited 09 <input type="radio"/> Yes limited
19. Does this person have any long-term disabilities or handicaps? See Guide	10 <input type="radio"/> No 11 <input type="radio"/> Yes

New version of the Census filter questions

ACTIVITIES OF DAYLY LIVING 20. Does this person have any difficulty hearing, seeing, communicating, walking, climbing stairs, bending, learning or doing any similar activities?	01 <input type="radio"/> Yes, sometimes 02 <input type="radio"/> Yes, often 03 <input type="radio"/> No
21. Does a physical condition or mental condition or health problem reduce the amount or the kind of activity this person can do: (a) at home?	04 <input type="radio"/> Yes, sometimes 05 <input type="radio"/> Yes, often 06 <input type="radio"/> No
(b) at work or at school?	07 <input type="radio"/> Yes, sometimes 08 <input type="radio"/> Yes, often 09 <input type="radio"/> No
(c) in other activities, for example, transportation or leisure?	11 <input type="radio"/> Yes, sometimes 12 <input type="radio"/> Yes, often 13 <input type="radio"/> No