EM8002: (Kitchener-Waterloo Record, March 29, 1980, page 7

Poorly phrased census questions costly

By REUEL AMDUR

One of the intriguing spectacles in social life in Canada today is the process by which the Parti Quebecois government has been working out its referendum question. Early in the term of the Levesque government, it became clear that any question asking Quebecers whether they were for or against independence from Canada would lead to a resounding vote against independence.

When it was clear that independence would be rejected by a strong majority, various polling organizations began looking at what the answers would be to various other kinds of questions related to the issue of the tie between Quebec and the rest of Canada. Some of these surveys were conducted by the PQ itself, while others were done by polling organizations for the media and other consumers of such information.

Levesque and his colleagues followed these polls with great care. They found that in order to approach their majority in the referendum they would have to steer clear of any suggestion of a vote for separation. Instead, they would need to talk about getting a mandate to negotiate with the rest of Canada. So the final question for the referendum speaks of a mandate for negotiation.

A yes or no on the question of independence is quite straightforward, but the meaning of receiving a mandate to negotiate sovereignty-association with the rest of Canada is very ambiguous. The question had to be asked that way because otherwise Levesque was sure to lose, but the meaning of the answer the voters of Quebec give will undoubtedly be the subject of considerable debate.

In surveys and referenda, the way in which the question is phrased will affect the answers received. The same question phrased in different ways will receive different answers. In many cases the difference in phrasing will lead to a different understanding of a question by the person answering. For that reason, it is important to be sure that questions are phrased in ways that get the desired information back.

Our largest Canadian survey is the census, and the same principles apply to questions it asks. In the 1976 census, there was a question about school attendance. The answers to this question showed a massive school dropout rate, with 40 per cent of youth 15 to 17 not attending school full time in Metro Toronto. Similarly very high dropout rates were reported in other locations as well. It turned out that the high dropout rate was phoney, a result of the way the question was worded.

Statistics Canada asked people if youth in the household were enrolled in "an educational institution." Many people did not understand that term to mean, simply, a school.

More recently, the Social Planning Council of Metro Toronto conducted a survey of young people in Metro. A respondent in that survey was asked if he was "enrolled as a full-time student in the month of April, 1979." The results showed a drop-out rate below 10 per cent.

Statistics Canada themselves learned of the error in the census question. In one bulletin, they alert users of census data that "the population 15 years and over attending school full-time may be underestimated by as much as 25 per cent of the published figure at the Canada level." The inaccuracy of the census on this question may vary from place to place to some extent, so that it is not clear how inaccurate the figure will be in any given location.

Perhaps the information is even worse than useless, because users of the data who reacted before seeing the government's commentaries on the results may have engaged in planning for problems in inappropriate ways.

But could the mistake have been prevented? Users of census information have had a constant battle with the federal government to prevent cutbacks. The government has attempted to shorten the census and limit the budget of Statistics Canada. As a result, the kind of careful exploration of different forms in which to put the question that Levesque and the pollsters in Quebec have gone through must have been neglected with the question on school attendance.

A poorly phrased census question is more expensive than the spadework needed to ensure that the question is asked correctly. When it is asked incorrectly, it then becomes necessary to issue bulletins telling users to ignore the data accumulated. If enough poorly phrased questions are asked, the census as a whole becomes subject to lack of confidence.

The article EM8002 reprinted above is used in Figure 8.8a of the STAT 220 Course Materials, in Chapter 4 of the STAT 231 Course Materials, in Figure 3.5a of the STAT 332 Course Materials and in Statistical Highlight #11.

1995-04-20