THE IMPORTANCE OF BEING CONSTANT
The effects of questionnaire overload

Wim van der Noort, Inter/View, Costa Tchaoussoglou, S.U.M.M.O.

Synopsis

In the course of 1993, the Dutch national readership survey, the SummoScanner showed a decline in the readership levels which did not seem to be in line with circulation figures.
After a thorough examination the conclusion was reached that a substantial increase in the size of the questionnaire had affected the readership figures.
In this paper we will describe how we came to the conclusion, that there are indeed limits in what we can ask from the respondents. We will especially focus on the critical role the interviewer plays.
We postulated that the increase in the questionnaire has led to an increased pressure on the interviewer to keep the interview within acceptable time limits for the respondent. In this way extensions in the second half of the interview can affect the measurement of printmedia in the first part of the interview.
This hypothesis was tested with a special test using an abbreviated SummoScanner run parallel to the normal SummoScanner.

In a second paper "Back to before yesterday", also presented at this Readership Symposium in Berlin, we will deal with the way S.U.M.M.O. tackles this problem of questionnaire overload and the consequences it has for the method of measuring readership.

Introduction

The SummoScanner is a continuous telephone survey (CATI). Except on sundays, interviews are conducted daily 615 every week, 32,000 a year. The survey is primarily focused on the measuring of printmedia. But in addition television, radio, cinema are also measured and also some target group information about groceries and cars.

The total interview time is on average 25 minutes, but can take much longer, depending on the reading and viewing behaviour of the respondent.

The key figures that we measure for the printmedia are:

- ever read
- source of copy
- (first time) read yesterday (FRY)
- time spent reading yesterday
- recent or older issue
- reading frequency

Now at the end of 1993 some worry was expressed from the side of the publishers, that the readership figures did not develop in line with the circulation developments. In itself this is not an unusual reaction. Real and imagined discrepancies between AIR and circulation data keep media researchers going. Nevertheless, S.U.M.M.O decided to undertake a thorough analysis of the decreasing readership figures.

With over a 100 titles to analyze, it's not easy to present a clear picture, since the readership developments in relation to the market developments do in fact differ a lot for different media groups and titles. But in general the following conclusions were drawn:

The overall gradual decline in total readership and reading frequency could be attributed to demographic and cultural changes like more individualisation, less free time and changing reading habits.
Over and above there was also a drop found in the reading probabilities. The reading probabilities in the SummoScanner are measured by the FRY-questions and are calculated and attributed for 14 different groups of titles to the reading frequency scores. This drop in the reading probabilities didn't seem plausible and not in line with findings in other surveys. Extensive checks of the software and the performance of the fieldwork showed nothing unusual and offered no explanation.
So in the end, only one possible explanation remained: the expansion of the SummoScanner in recent years had possibly led to an overloaded questionnaire, affecting the readership levels.

'Fortunately' the new low reading probabilities were not used yet in the calculation of the average issue readership, since S.U.M.M.O. was just working out a new system of attributing reading probabilities. (See C. Tchaoussoglou, 'Future Perfect: uniform fixed probabilities', San Francisco, 1993).

So the board of S.U.M.M.O decided for the short term to fix the probabilities on the level of 1992. They also asked the Technical Subcommittee to provide more substantiation of the hypothesis of questionnaire overload, before taking measures.

The hypothesis of questionnaire overload

In recent years the questionnaire of the SummoScanner had gradually, but substantially expanded. Part of this expansion is a sort of natural process. Print titles come and go, but more come than go and of course the SummoScanner had to adopt to the changing market and to incorporate new titles.
Over and above a rather drastic change (and extension) also took place in the measurement of television, the number of radio channels grew and there were also some nice new questions about cars, income and the refusal of direct non mail.
We always thought - perhaps somewhat naively - that changes in the second half of the interview could not interfere with the measurement of the printmedia in the first half of the interview. For the respondent doesn't know what is to come. But the interviewer does know. So what we saw in our analyses was that in spite of the extensions, the interview time remained more or less constant at 25 minutes. That raised the hypothesis, that there seemed to be a sort of psychological barrier with regard to the total interview time. Since interviewers are paid for worked hours and not per completed interview, there is nothing to be gained by rushing through the interview. But interviewers do feel the pressure to complete the interview within a time that is acceptable to the respondent. In this way a longer questionnaire can lead to a more hasty interview and so affect the results of the printmedia. This was the hypothesis we wanted to submit to a test.

Test Reduced Scanner

During 6 weeks in September and October 1994 we did a split-run test. In this period half of the SummoScanner sample was interviewed with the standard S.U.M.M.O questionnaire. For the other half of the sample a reduced version of the questionnaire was used. The first part of the questionnaire, in which the printmedia are handled, was in both versions exactly the same. Left out in the reduced Scanner were in particular parts of the television- and radio-measurement and some other questions. The total reduction of the questionnaire amounted to 15%.

The interviewers were divided in two groups, matched on the criteria sex, age, interview experience and quality performance.
The written instruction was for both groups the same and did not reveal the real aim of the experiment.

With this test we hoped to find an answer to the following questions:

- Will a relative small reduction of the questionnaire lead to less time pressure, and so to a better quality interview.

- Will a shorter interview produce higher readership figures for the print media.

In the next sections we will present some of the main findings. We will look at differences in interviewing time between the two versions, differences in measurement results and also the judgement of the interviewer will be taken into account.

Interview time

A nice feature of telephone surveys is the fact that we can register exactly how much interviewing time on average is spent on the different parts of the questionnaire. In table 1 you find an overview and a comparison between the SummoScanner and the reduced version of the SummoScanner.

Table 1: Time spent in minutes

	SummoScanner min.	Reduced Scanner min.	Change %
Print	10.67	10.91	+1.0
TV, radio	7.70	4.70	-12.2
Other (unchanged)	2.99	3.12	+0.5
Other (changed)	3.29	2.87	-1.7
Total	24.65	21.60	-12.4

As can be seen from this table, almost half of the interview time of the SummoScanner is spent on printmedia, 10.67 minutes. Next, the questions on radio and television also take much time, 7.7 minutes. As a result of the reduction of the radio and television questions, the time spent on radio and television was exactly three minutes less in the Reduced Scanner, 4.7 minutes. Percented on the total interview time of 24.65 minutes, this means a reduction of 12.2%. So together with the reduction of 1.7 in the other parts of the questionnaire, the realised time reduction due to the changes was 13.9%.

It's obvious, that the time saved by reducing the questionnaire is not automatically filled in by taking more time for the remaining parts of the interview. Only slightly more time was spent on printmedia in the Reduced Scanner, 10.91 minutes, which means a relative increase of 2% compared to the 10.67 minutes of the SummoScanner.
Percented on the total interview time of the SummoScanner the time reduction of 13.9% is compensated by 1.5% more time for the printmedia and the other parts of the questionnaire, that remained unchanged.
So, by and large the reduction resulted for say 90% in a shorter interview time and only for 10% in a less hasty interview.

Readership data

Did the shortened version of the SummoScanner produce higher readership figures? Table 2 gives the answer. And the answer simply is no.

Table 2: Comparison of the readership scores in the SummoScanner and the Reduced Scanner

Average number of titles	SummoScanner n=1850	Reduced Scanner n=1862
Ever read	14.21	13.99
Read yesterday	1.53	1.40
First time yesterday (FRY)	0.63	0.55
Subscriptions	2.51	2.46
Single copies	3.52	3.56
Average reading frequency	2.39	2.30

The differences in total reach, the number of titles ever read, are not significantly different between the two versions of the SummoScanner. The number of titles read yesterday (for the first time) and the average reading frequency - the basis for the average issue reach- tend to be even lower in the Reduced Scanner. It's obvious, that our hypothesis wasn't sustained by these results.

So if there are any effects to be expected from (slightly) reducing the questionnaire they do not come immediately and automatically.

Additionally we also analyzed the results of the last two weeks of the test, to see whether there is perhaps some learning effect, i.e. that the interviewers need some time to gain experience with the new reduced version of the SummoScanner, before changing their interviewing habits. As we can see in table 3 there are some indications, that this can be the case.

Table 3:   **Last two weeks of the test period compared with the whole period**

	SummoScanner		Reduced Scanner	
	wk 34/39 average	wk 38/39 index	wk 34/39 average	wk 38/39 index
Ever read	14.2	101	14.0	103
Read yesterday	1.5	101	1.4	*111
First time yesterday (FRY)	0.6	103	0.6	*118
Subscriptions	2.5	103	2.5	104
Single copies	3.5	102	3.6	103
Average reading frequency	2.4	105	2.3	104

Especially with regard to the results that determine the level of the reading probabilities (read yesterday for the first time), we see in the Reduced Scanner a marked increase in the overall level in the last two weeks of the test compared with the whole test period.

Unfortunately the test didn't extend over a longer period, to see whether this trend would have been continued.

Judgement of the interviewer

As part of our test we also asked the opinion of the men and women "in the field": our interviewers who are of course a key factor in the quality performance of the survey.
During the last two days of the fieldwork the interviewer filled in a checklist at the end of each interview. This checklist contained questions about the easiness of the interview, whether there was any time pressure felt and how the interviewer perceived the reliability of the answers given by the respondent.

In Table 4 we see that the majority of the SummoScanner interviews are seen by the interviewers as 'very easy' (33%) or 'easy' to do .

Table 4:   **Easiness of the interview**

	SummoScanner n=100 %	Reduced Scanner n=102 %
Very easy	33	37
Easy	43	47
Not so easy	18	14
Not at all easy	6	2

We also see that in 24% of the cases the interview was not so easy or even not easy at all. In the Reduced Scanner fewer 'problems' are reported by the interviewer. Here 'only' 16% of the interviews are considered as not easy to do.

A comparable difference between the SummoScanner and the Reduced Scanner comes forward when we ask the interviewer whether there was any time pressure felt during the interview (see table 5).

Table 5:   **Was there any time pressure and when was the time pressure felt for the first time**

	SummoScanner n=100 %	Reduced Scanner n=102 %
Yes, time pressure	26	19
First felt:		
Direct from start	31	30
During printmedia	35	35
After printmedia	34	35

In the SummoScanner, in 26% of the cases, time pressure was felt. This is more often the case than in the Reduced Scanner, where time pressure was felt in 19% of the interviews.

We also asked the interviewer when the time pressure was first felt. Here we see that in one third of the cases with time pressure, the interview suffers from this pressure right from the start. Other typical moments in the interview where respondents tend to show impatience, are at the end of the print section when the reading frequency questions are being asked ('is there still more to be asked about my reading behaviour?') and during the rather extensive television questions.

We see here a confirmation of our hypothesis, that (adding to) the length of the interview can affect the whole interview and not just the end of it, when some respondents indeed can experience the feeling, that their stamina is being tested too much. The interviewer can anticipate the possibility of time pressure. When trying to persuade the respondent to participate in the interview, the often successful phrase 'it won't take long' is unfortunately not applicable to the SummoScanner. Total interview time often plays a role in the hesitations of the respondent to participate. There are no explicit Instructions for the interviewer what to say to the respondent about the interview time to be expected. They will try to avoid being very explicit, saying something like that it differs a lot (what in fact it does). But when they do give an indication of the time the interview will take, it's in the nature of things to be a bit optimistic. In that case an interviewer will feel from the start more pressure to bring the interview to a good end in order to avoid the terrible waste of time when a respondent terminates the interview before the end of it, which always means a complete loss of the interview. This happens in about 2% of the calls.

Finally we show you in table 6 how the interviewers judged the reliability of the interview they had just conducted.

Table 6:   **Perceived reliability of the total interview**

	SummoScanner %	Reduced Scanner %
Very reliable	29	44
Reliable	65	48
Not so reliable	6	8

It's pleasing to see that both in the SummoScanner and in the Reduced Scanner the answers of only a small minority of the respondents (6% to 8%) are judged as 'not so reliable' by the interviewers.
A marked difference however can be seen between the categories 'very reliable' and 'reliable'. The interviews of the Reduced Scanner were far more often credited with the predicate 'very reliable' (44%) than in the case of the normal SummoScanner (29%). When asked to specify this for the different parts of the questionnaire the perceived reliability differed the most for the television questions, which in fact were in the Reduced Scanner far more easy to handle, leaving out the part in which the respondent has to state which specific television programs he/she did see yesterday. But also with regard to the printmedia questions, that were the same in both versions, the Reduced Scanner was judged as more reliable than the SummoScanner.

Conclusion

So what could we conclude from the results of this test? Frankly, we had not been able to really 'prove' that overburdening the questionnaire was the cause of the decline in the reading probabilities. And if it was, we could now see that this process could not be immediately reversed by leaving some of the questions out. Reduction of the interview did not result in a direct effect on the readership figures.
But the results of the test did give support to the common sense feeling, that you can't go on adding to the questionnaire, without paying a price. We did find indications that reduction of the questionnaire means less haste and more quality in the interview, which in the longer run also would have it's effect on the readership levels.
So it convinced the members of S.U.M.M.O. that the decline in the reading probabilities of printmedia could be partly attributed to changes in the survey circumstances and that measures were necessary to guarantee the stability of the SummoScanner, to protect it's use as currency for the printmedia.

How it went on

The dilemma S.U.M.M.O. was confronted with, was not an easy one to solve. The wish to protect the measurement of the printmedia from questionnaire overload seemed fair enough. But at the same time S.U.M.M.O. highly values the multi media approach of the survey. After all, the extension of the SummoScanner is a reflection of the growing information needs of the members, which S.U.M.M.O wants to keep fulfilling. Our paper 'Back to before yesterday' deals with the way S.U.M.M.O. tackled this dilemma.

There remains here one last point to deal with: what happened to the reading probabilities? How did the readership levels develop in the second half of 1994 and the first half of 1995?

The reading probabilities based on FRY did in fact recover from the 'deep fall' that they had been through. So in our last publications we could again calibrate our reading probabilities on the actual measured FRY data and still at the same time produce stable or even slightly rising readership levels. How is this possible? One thing is sure, we didn't reduce the questionnaire. What did happen was, that the whole process of scrutinizing the SummoScanner seemed to have had an effect in itself on the interview. For one thing, extra sessions had been held with the interviewers, to hear their experiences and points of view, and also to motivate them by telling more about the backgrounds and the use of the SummoScanner and the importance of accurate reach figures. But also the strings of control were pulled more tight. In addition to the constant supervising and monitoring of interviews, analyses were made of the average measurement outcomes of each interviewer. The differences that were shown between interviewers provided us with a tool for a more directed quality control.

As long as the information content of the SummoScanner questionnaire remains on the same 'high density' level, the challenge now is to maintain these extra efforts of the fieldwork performance.
One thing which we have learned from this whole experience is: the importance of being constant.