

Marion C. Appel
Inter/View Nederland B.V.
Amsterdam, The Netherlands

Costa Tchaoussoglou
S.U.M.M.O.*
Amsterdam, The Netherlands

DEVELOPING SUMMOSCANNER; WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

SUMMARY

Some participants of SUMMO are led to believe that a more simple, straightforward Recent Reading-method is preferable to the FRY-method now in use for SummoScanner. Their reason for so thinking is the small number of FRY cases, notwithstanding a large sample, for smaller magazines and for monthly magazines in particular.

In order to contribute to the decision on how to proceed in the future, more evidence was needed to make a fair and well-founded comparison of FRY and RR. By fair we mean under *ceteris paribus* conditions. In consequence we had to conduct some new experiments, since there was hardly any experience with RR in a CATI setting.

The main objectives of these experiments were to find out:

1. Whether Recent Reading is feasible within a the scope of a telephone interview
2. Whether Recent Reading measurement can be improved for its known defects: telescoping, forgetting of chance contacts, parallel reading and replicated reading
3. To what extent RR-AIR will differ from FRY-AIR

As to the first question the answer seems to be yes, provided the simple question *which of the following magazines did you read or look into in the past 7 (or 30) days?* is used.

* Samenwerkingsverband voor de Uitvoering van Multimedia Onderzoek, the Syndicated Media research organisation in the Netherlands since 1986, commissioner of SummoScanner and TGI.

DEVELOPING SUMMOSCANNER; WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

Secondly, checking the validity of the answer to this question seems to be necessary: several corrections were made after remembering the *most recent reading event*. One major finding seems to be that people are less sure of their answer, after we try to help them remember the exact day!

Also important in its effect on AIR measurement is the amount of parallel reading for some specific titles.

Also, first reading does make some difference, but is especially hard to establish for monthly magazines, where this seems to be of great consequence.

These corrections do have opposite tendencies, resulting in only small differences in AIR totals before and after validation. This conclusion is however extremely premature: the experiment was conducted on a very small scale. It must be pointed out, that results for particular titles can be very different depending on whether or not validation is carried out. Certainty in the answer to our third question, however, can only be acquired by carrying out the third stage of our study, which is planned for April 1991.

1. BACKGROUND

About a year after the first Readership Symposium in New Orleans in 1981, Maurice de Hond started the Inter/View Media Scanner. Readership figures were established on the basis of respondents reporting that they read or looked into a given issue of a title yesterday for the first time. There is some 'historical' logic in the fact that Inter/View began with what is called the FRY method.

Although the concept of FRY was not new in 1982, Maurice de Hond was simply combining the inspiration gained from the debates during the New Orleans Symposium with the fact that Inter/View had been operating a CATI Unit since the beginning of 1981.

In New Orleans, much time was spent on TTB (Through The Book) versus RR (Recent Reading). A remarkable contribution was made by Valentine Appel introducing the telescoping phenomenon in RR, although he supported his case with results for television programmes broadcasted weekly. Anyway, the assumption that the longer the period of recall, the less accurate the measurement of reading behaviour must be, was again confirmed. On the other hand, yesterday recall was used as a yardstick to validate or invalidate results obtained by either technique.

DEVELOPING SUMMOSCANNER; WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

Also, it will be quite clear that to obtain reliable results using the FRY method, rather large samples are necessary. This is no doubt the reason that nobody so far had put the promising FRY technique into practice. The possibility of doing 39.000 interviews on a yearly basis, (125 per day, six days a week) in the Netherlands removed that last obstacle.

Since the introduction in 1982 many changes have been made, especially since the time that the Media Scanner became the SummoScanner in 1986. Most of these changes have already been reported at earlier symposia or seminars.

Validation experiments of the First time Read Yesterday-method in the SummoScanner, accompanied by exacting scrutiny of sampling and weighting methods and fieldwork control, have now resulted in a national readership survey that can conform to fairly severe standards in media research, and which is suited to the needs of the Dutch media-field.

2. INTRODUCTION

Until last summer we were preoccupied with the problem of reading probabilities above one and the peculiar problem of Dutch TV-magazines. In the meantime, the validity and reliability of the results were discussed among SUMMO members and users. This is no surprise, since compromise is always necessary. The principal remaining problem is caused by the reduction in reach measurement to only *one day*, instead of a week or a month, as in the Recent Reading-method. This presents no problem to daily newspapers, rather more of a problem for weekly magazines, but a larger problem for monthly magazines. Readership calculation on the basis of one day ("yesterday") reduces the reliability of reach figures, since theoretically the sample is divided by 6 (the number of interviewing days within the publication interval) for weeklies and by 25 for monthlies. Reliability has been artificially improved by means of enlarging the calculation basis to yearly samples and by summing up figures of several titles (grouping similar publications together) for calculating probabilities.²

Still, grouping has a negative effect on the validity of individual reach figures, although the extent of this effect is not known. Especially for smaller weeklies and monthlies, grouping is necessary to meet the criterion of a minimum number of FRY's. Needless to say, by increasing the reliability, the validity of the figures decreases at the same time. The establishment of Utopia in Dutch readership research should however improve reliability, *without decreasing validity*. This is virtually impossible within the FRY method.

DEVELOPING SUMMOSCANNER; WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

The problem that is now facing SUMMO, i.e. the balance between validity and reliability, is no longer a purely technical one. The main reason for choosing FRY, as we did in 1985, is our belief in the greater validity compared to RR, thanks to the shorter recall period. The implication is, however, that the longer the issue period, the greater the number of readers must be to meet the minimum number of FRY's. In other words, monthlies are at a disadvantage, since they are much more likely to be grouped or not to be included at all, than weeklies with the same number of readers.

As an alternative the use of Recent Reading questions is being considered. This might however mean a decrease in validity caused by telescoping and forgetting, parallel reading and replicated reading, which are largely overcome or avoided within the FRY-method. On the other hand, advantages generated by CATI, such as a large national sample, computer assisted interviewing and interviewer control, must be maintained. It was therefore decided to do a study within the CATI system.

3. RESEARCH DESIGN

In tackling the use of Recent Reading within the SummoScanner we have defined two sets of problems: firstly those problems originating in the nature of telephone interviewing, secondly, the known problems of the Recent Reading method itself. Trying to split up these two sets, we developed an experiment in three stages.

The goal of the first stage was primarily to test the possibility and difficulty of using Recent Reading in a CATI interview. The second stage was intended to be the pilot, trying to establish a questionnaire that could in the final stage deliver some quantitative results. The first and second stages were carried out in the last three months of 1990. The third stage will take place in April 1991 on a larger scale (n=1000). The sampling procedure for all three stages of the experiment was similar to the usual SummoScanner procedure: random digit dialing. One person in the household was selected to be interviewed. The questions to be tested in each stage are being integrated into the usual SummoScanner questionnaire, aiming to simulate a real situation as near as possible. Therefore the wording of the ever-read hurdle is left unchanged. Of course some questions were left out, but only after the central question blocks had been completed. In this way it seems possible to draw conclusions concerning the consequences of using Recent Reading in CATI-interviews.

DEVELOPING SUMMOSCANNER;
WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

4. THE FIRST STAGE

To get some impression of the feasibility of a Recent Reading method by telephone, 25 interviews were carried out, each personally attended by one or two researchers. As in the actual SummoScanner the questionnaire was largely vertically designed, beginning with the *ever-read or look into*-question for all titles. Then the subscriptions and reading circle portfolios are established, as well as the purchasing of individual copies. The next questions in SummoScanner would be *titles read yesterday and first time reading*. For magazines these were replaced by questions establishing the *most recent reading event (excluding today) of each ever-read magazine*.

This question was asked for every title the respondent ever reads or looks into. This can go up to 30 or 40 titles, on average 16 magazines in this sample, which is not very different from usual SummoScanner findings.

In this stage no attention was paid to the validity of the answers given, merely the quality of the interview and the length of time the Recent Reading block takes, were taken into account.

Results of first stage

This simple experiment made it clear that *most recent reading* cannot be used in a telephone interview. On the one hand, people get irritated by having to answer this difficult question, especially when they read many titles, while on the other, this method of questioning lengthens the interview time substantially. On average the length of the RR-block was 5 minutes, as opposed to 2.4 minutes necessary for the (F)RY-questions (including the amount of time spent reading). Moreover, the variation in interviewing time was of course considerable: some respondents needed up to 8.5 minutes to remember all their most recent reading events.

The difficulty for some respondents to remember the day they last read a particular magazine is illustrated by the average amount of time per title which was in most cases between 0.1 min. to 0.66 min., one old lady even climbing to 0.97 min per title. Clearly, more highly educated people had fewer problems answering our questions. (It might be pointed out that more highly educated respondents seem to be less scrupulous in answering: they do not hesitate to make a (wild?) guess). Also some differences seem to be caused by age.

DEVELOPING SUMMOSCANNER; WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

We concluded after this stage that RR, using the last time read question, however beautiful it is from a theoretical point of view, was not feasible by telephone:

- * The length of the interview increases enormously for respondents who read more than 7 titles (which is true for 70 % of all respondents).
- * The degree of difficulty of the interview increases when open questions are added, thereby placing heavy reliance on people's memories. This is too large a strain within a telephone situation.
- * Also the understanding of the frequency question (number of issues read out of the *last* six) was often hindered because of the similarity in wording of the RR- and frequency question.

5. THE SECOND STAGE

For the next stages it was decided not to use the *most recent reading-question*, but first to adhere to the simple Recent Reading questions, establishing which titles have been looked into *within the most recent issue period (7 days or 30 days respectively)*.

In this way, the *yesterday-question* is replaced by the relevant *recent issue-period-question*. The only remaining factor influencing interview time is the greater strain people must put on their memories.

In the second stage of the experiment the question order was as follows: first monthlies, then bi-monthlies and weeklies, in an effort to overcome the known effect of fatiguing when the most difficult question is asked last.

Our goal in this stage was to establish the possibility of validating the answers to RR questions within the telephone-interview, in order to design an RR-questionnaire to be used in the third stage of our experiment.

Three problems were to be tackled in this stage:

- a. validating the answers close to the margins of the issue periods, correcting telescoping and forgetting of chance contacts.
- b. establishing first reading within the issue period
- c. establishing parallel reading within the issue period and first time reading for each copy read

**DEVELOPING SUMMOSCANNER;
WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?**

29 interviews were held in this second stage, 15 concentrating on weeklies and 14 concentrating on monthly magazines. This split was made to limit somewhat the time needed to ask probing questions in each separate interview.

Questionnaire

Having established Recent Reading the straightforward interviewing procedure was departed from and several probing questions were asked. These questions were CATI-programmed for the most part, but the interviewer had the freedom to add probing questions or leave some out if the answer was already clear. Each interview was attended by a researcher and was followed by a brief evaluation with the interviewer.

(a) The validation of Recent Reading was done in two steps: the first consisting of a *most recent reading*-question, which was not precoded, while the second step was to assist the respondent in remembering when the last occasion was, if the first answer was uncertain or marginal (just before or after 7 or 30 days).

(b) The measurement of first reading was accomplished by using the following probes: where was the last reading occasion, was it a chance contact, did one read or leaf through the copy, when was the copy received in the home, was this copy looked into on the day of receipt or was it left until later, if so, when was it taken up for the first time? In many cases people do glance through a newly arrived magazine, as was our experience in the FRY-experiments. This led us to concentrate our questions on the day of receipt, not considering the number of days on which the copy was picked up.

(c) The first time reading procedure was followed for each of the copies of a title read within the issue period.

DEVELOPING SUMMOSCANNER; WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

Results of second stage

It proved to be quite difficult to get people to remember the exact day and situation when they looked into a magazine. Of course this is not true of magazines the respondent is attached to and has some kind of habitual behavior towards³. Many contacts however are not remembered clearly, this obviously being more the case with monthlies than with weekly magazines. Often, corrections were made to the initial answers to the RR-questions.

The measurement of first reading also proved difficult: people do indeed look into the new copy as soon as it is physically present in the home, but especially in the case of monthly magazines the date of receipt is not clearly remembered. The problem is an easy one within the FRY-method (if received and looked into before yesterday, non-FRY is clear), but it turned out to be very hard thinking back a month or more. However, some cases were clear: had one received the latest issue say about a week ago, it could be assumed that the previous issue had to be received (and read) before the issue period. This was of some help.

Parallel reading was found in several cases, as was to be expected. Also, a large number of these cases turned out to be First Reading. This is clear when someone gets a pile of magazines once a month from a member of the family who subscribes. Obviously this is the case for women's weeklies and do-it-yourself magazines. This emphasises the fact that some titles can be severely under-reported in the RR method, e.g. Margriet was read in the issue period by 5 respondents, but in fact this would amount to 11 copies first time read (by 4 respondents!), an augmentation of AIR from 33 to 73%. Of course first time reading is even harder to measure in the case of (very) old copies: one might have looked at it at an earlier date but have completely forgotten about it.

In all, 12 out of 29 interviews were considered to be completely correct: no mistakes seem to have been made in reporting Recent Reading, nor was parallel reading or replicated reading encountered.

**DEVELOPING SUMMOSCANNER;
WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?**

Although figures in a small experiment such as this one are not of great consequence, we have summarized some outcomes as an indication of the extent of reach-corrections.

Some results of RR experiment (second stage) SummoScanner 1990

	n=15 weeklies	n=14 monthlies
	cases	cases
Gross readership	105	94
Average issue readership:		
initial: titles	53	36
definite: copies	80	49
Corrections RR:		
Pos.	5	1
Neg.	5	1
First time (including parallel reading and corrections)	59	39
Uncertain AIR	9	8
Certain and correct AIR	50	31

In 13 cases parallel reading was found concerning RTV-guides. This is to be expected, since the interviews took place on Monday, Tuesday and Thursday.

First reading was in almost all cases measured for only one of these issues.

Still, gross readership of weekly magazines turned out to be somewhat higher than initially measured (393 % instead of 353 %, including uncertain answers, which seem to be first time reading).

DEVELOPING SUMMOSCANNER;
WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

The same tendency occurs with monthly magazines: first time reading is slightly higher than the initial measurement of RR-AIR, (increase of gross readership from 257 % to 278 %) but 8 cases are not quite certain, which amounts to 20% of gross average readership relatively.

Of 14 interviews concentrating on monthly magazines, 7 revealed parallel reading of 11 titles altogether. This amounted to 3 extra instances of first time reading. It should be noticed that, in contrast to weeklies, several copies of monthly magazines were found to have been read in the home, by subscribers. Parallel reading of weeklies more often seems to occur at someone else's home.

Conclusion of second stage

When a respondent finds it difficult to answer a question, he is likely to modify it in such a way as to be able to answer it more easily.²

It is regrettable that respondents do not tell us in what way they modify those of our questions, which are difficult to answer. Sometimes it is more or less clear that people report on habits (*we usually glance through it on arrival. I always read Weekend, Story, or Prive when I am at the hairdresser's*). It is very difficult to get precise answers about something people do not tend to remember.

Listening to the recordings of the interviews we were led to believe that this is not a specific problem for telephone interviews, compared with face-to-face interviews³. Of course, during a telephone interview there is some limitation on the amount of time available, so there is not always sufficient time to avoid all possible mistakes. This seems to have consequences for some titles in particular, especially large Women's weeklies.

Most of the time, correction of the exact day a magazine was read occurred in the *most recent reading* question. Probably the cause of this is the fact that one has now an overall picture of the period and has time to reconsider the last contact. However, it does not seem feasible in the light of the first stage of this experiment, to use the *most recent reading*-question in a telephone survey.

It should be noted here that we cannot be sure on the point of telescoping: people do make corrections to earlier statements (occurred in 7 interviews), but it is not clear whether these mistakes should be labelled "telescoping". (In fact positive as well as negative corrections both amounted to 6 cases).

DEVELOPING SUMMOSCANNER;
WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

6. THE THIRD STAGE

In the final stage of the experiment a larger number of interviews will be completed, so as to be able to compare results of RR-questions with the actual SummoScanner results (FRY). In both cases CATI is used, sampling is identical and the questionnaire is, apart from the central readership questions, exactly the same.

At this point it has to be decided whether or not to try to use validations within the third stage, should this stage be carried out: the difference between initial AIR and the validated AIR overall is not very large. On the other hand, a great number of corrections do take place as a result of further probing. Will we be able to shut our eyes, in view of the uncertainties that will remain even after we have put so much energy into it?

Decisions have to be made concerning the following points:

- a. Will we add the *most recent reading*-question in order to improve the RR-measurement?
- b. Will we add probe questions for monthly magazines that were read in the home, concerning parallel reading? (Which would require technical solutions: how are we to count extra AIR from second or third copies!)
- c. Will we add probe questions for weekly magazines, which were read elsewhere, concerning parallel reading?

Since these points are still open at this moment we are not able to present a final questionnaire for the third stage of our experiment. Hopefully at our next meeting we will be able to be more definite as to how to bring Utopia into the real life SummoScanner.

At this very moment we are not at all sure whether the third stage will be carried through. Looking at the results of the first two stages, sceptics may argue that the advantages of a straightforward RR method, will never outweigh the foreseen loss of validity in replacing FRY with RR. To optimists and publishers of monthly magazines however, the same results may seem very promising. From a mere researcher's point of view the decision to renounce the third stage is arbitrary and unsatisfactory. So it is up to the SUMMO Board to cut this Gordian knot.

DEVELOPING SUMMOSCANNER;
WILL THE FUTURE FIND UTOPIA IN THE NETHERLANDS?

NOTES.

¹ De Hond, Maurice and Huzen, Walter (1983). New approach to readership surveys: The Mediascanner. Montreal Proceedings Readership Symposium.

Walstra, Bouke (1985), Validating the yesterday first-time-read method. Salzburg Proceedings.

Appel, Marion and Walstra, Bouke (1988). On validating the first time read yesterday method - A continuing story. Barcelona Proceedings.

² Van Niekerk, P.W. (1988). Reliability of the first time read yesterday method. Barcelona Proceedings

³ Van Vliet, J.I. and Tchaoussoglou, C.J. (1983). From far fetched ideal to practical usefulness: towards a better definition of readership. Montreal Proceedings.

⁴ Belson, W.A. (1981). The design and understanding of survey questions. Aldershot, U.K.

⁵ Brown, M. (1988). Developing and validating a measurement of "first reading", Barcelona Proceedings

Amsterdam, January 1991