Gunda Opfer
GETAS Gesellschaft für angewandte Sozial-psychologie mbH
Bremen, West Germany

2.3   THE EFFECTS OF SHORTENING THE PERIOD OF RECALL

THE IDEA OF SHORTENING THE PERIOD OF RECALL

Any approach to coverage measurement through interviewing is dependent on the respondent's recall ability. So it seems obvious that the results will become more and more reliable the shorter and the more recent is the time period referred to in the interview.

The best known and most frequently discussed example is the so-called 'reader yesterday'. Depending on whether one starts from the method 'Recent Reading' (RR) or 'Through-the-Book' (TTB) different versions of questioning have been applied.

(a) For several years the questionnaire of the German AG.MA included the possibility of answering 'yesterday' to the question: "When did you read in or look through any issue of this magazine for the last time?"

(b) In their contributions to the Symposium in Montreal, the authors Eadie/Lysaker (1983) and Douglas/Lysaker (1983) reported several experiments using the method 'yesterday recall' with the following sequence of questions (the introductory and explanatory sentences are skipped here):

- Did you happen to read any magazines yesterday? (filter)

- For each of these magazines* please tell me whether or not you happened to read or look into any issue of that magazine yesterday.

- How many issues of this magazine did you read or look into yesterday?

- Was it for the first time?

---

* The interviews were done by telephone; the titles were read to the respondents.

(c) The ARF-Certitude Tests, 1 to 3 (1978 and 1979), used a version of the 'yesterday recall' which worked with the presentation of original issues of the magazines; so it followed the idea of TTB. It also included the question whether the issue was read yesterday for the first time.

The version (a) has also been part of other readership surveys working with the method RR (eg the German LA-MED, readership analysis of medical magazines [1983]). This version (a) is not designed to contribute to coverage measurement, but it is used to calculate the number of reading days. In order to produce coverage values from RR-questions one has to include the question: "How many issues of this magazine did you read or look into yesterday for the first time?" such as the version (b) did.

In our opinion an additional argument against version (a) is: this type of question does not exhaust the advantage of the shorter period of recall. In order to do this it is necessary that the respondent really has the opportunity to think intensively of what has happened yesterday. This is the case in the versions (b) and (c) but hardly in version (a).

On the other hand version (c) was designed in the ARF-Certitude Tests in order to measure overclaim and underclaim. This TTB-yesterday recall method does not seem suitable for measuring coverages because for this purpose the interviews had to be equally distributed from the day when the issue occurs up to the last day of its life.

For these reasons we concentrate our attention on version (b) and its variants which follows the principle of the method 'Recent Reading'.

Such variants may consist in choosing different formulations (as has been tried out by the authors quoted above)

as well as in selecting different periods of time: each shorter period promises to produce more reliable results (condition: the shortened period starts from the present).

So for example it would be possible to ask for the reading acts during the last three days (see eg Tom Corlett, 1964), or - more extreme - during the last eight hours, during the last hour or even for the reading acts at this moment!

As in many situations, it turns out to be worthwhile to pursue the idea to the extreme in order to point out its advantages as well as its weaknesses. For it may be presumed that the advantage of the respondent's more reliable answer has to be offset by disadvantages regarding other aspects.

This paper will deal with assessing the balance between advantages and disadvantages by reviewing existing experience and by some theoretical considerations.

OUTLINE OF OUR PROCEDURE

Our work on the topic 'shortening the period of recall' led us to two advantages of the method. However, we found three disadvantages. This paper will discuss them as far as possible not only qualitatively but also quantitatively. It will demonstrate the gaps which would have to be filled by experimental work in order to decide which may be the optimum length for the period of recall.

To give an outline of the procedure we first list these advantages and disadvantages which are, in our opinion, connected with shortening the period of recall. See **Charts 1-5.**

CHART 1

Advantage 1 ⟶ *The shorter* and the more recent the time period referred to in the interview - the better and *the more reliable* is the respondent's ability of recall.

CHART 2

Advantage 2 ⟶ *The shorter* and the more recent the time period referred to in the interview - *the better is the possibility of validating* the respondent's answer.

CHART 3

The shorter the period referred to in the interview - *the more important* ⟵ Disadvantage 1 is it to ask for *'first-time reading'*.

CHART 4

The shorter the period referred to in the interview - the *higher is the* ⟵ Disadvantage 2 *data tolerance range* (produced when calculating the coverage values).

CHART 5

The shorter the
period referred to
in the interview -
the *more difficult*
is it to realise   ←—— Disadvantage 3
the demand of
representatively
distributing the
interviews to
these periods.

PERIODS OF TIME SHORTER THAN ONE DAY CAN HARDLY BE USED TO MEASURE COVERAGE VALUES

'The shorter and the more recent the
time period referred to in the
interview the better and the more
reliable is the respondent's ability of
recall': The high plausibility of this
assertion may lead to the assumption
that there is no need for a proof.

However, regarding the problem more in
detail, several restrictions occur:

- *Without any doubt* a person is able to
tell whether he is reading '*at this
moment*' in any issue of title X.

- The case will become *less certain*, if
the person is asked in addition whether
it is *for the first time* that he reads
in this issue.

- So it turns out to be necessary to
look at our 'first advantage' in
combination with the 'first
disadvantage': The shorter the period
referred to in the interview the more
important is the information of first-
time reading.

It is a well known fact, that without
this additional information the 'reader
yesterday' would yield a value
'coverage x reading days'.
Correspondingly the 'reader during the
last hour' would lead to a value
'coverage x reading hours'. The

information of 'how many copies were
read for the first time', included into
any RR-question, would solve the
problem of parallel and replicate
reading.

- The person's answer about his
'reading at this moment' may be correct
in nearly all cases - his statement
about other short periods of time such
as for instance 'reading during the
last hour' will become less certain for
the starting point 'one hour before'
may not be remembered accurately in
many cases.

- So it seems to be reasonable to take
as an intermediate between 'at this
moment' and 'yesterday' the period of
time 'last eight hours'.

This divides the day into three equal
parts which may be arranged such that
they correspond to the course of the
day: eg 6 am to 2 pm, 2 pm to 10 pm and
finally 10 pm to 6 am.

In order to make full use of the
expected precision of recall the
interviews would have to take place
exactly at these points of time. This
would be possible at most by telephone
and only with a low sample response
percentage.

Considering these possibilities one
realises at once our 'third
disadvantage': in order to measure
coverages by asking for the behaviour
during these short periods of time it
would be necessary to arrange the
interviews equally distributed around
the whole day - impossible (even by
phone) as one knows from the experience
of arranging interviews equally
distributed between the days of the
week.

Not yet having looked at the
disadvantage of statistical
uncertainty, the absence of any chance
of practicability leads us to the
conclusion that it does not make sense
to take into consideration periods of
recall shorter than one day.

We point out that this statement refers to using the short periods of recall as a means to measure coverage values. Things may look different if one wants to use short periods of recall for other purposes.

HYPOTHETICAL CURVES: ADVANTAGES AND DISADVANTAGES IN DEPENDENCE ON THE LENGTH OF THE PERIOD OF RECALL

We now concentrate our attention on the method 'yesterday recall'. We assume that the condition of equally distributing the interviews to the days of the week has been met, so that errors from this source are excluded.

The aspects 'better ability of recall', 'possibility of validation' and 'necessity of asking for first-time reading' (our first and second advantage and our first disadvantage) have to be regarded as a whole.

In order to estimate the degree of improvement which may be expected from shortening the period of recall to one day, one has to quantify the effects of advantages and disadvantages and to weigh them against each other.

So it would be desirable to have three curves drawn into a diagram, which shows the relationships between:

The length of the period of recall and

(a) the percentage of errors produced by the interview questions *including* first-time reading

(b) the percentage of errors produced by the interview questions *without including* first-time reading

(c) the sampling error which results from calculating the coverages from values measured by the short period of recall.

Chart 6 shows the *hypothetical* shape of these curves:

Which parts of these curves are

available today, and what research has been done?

Defining 'validation' as the possibility of comparing the respondent's answer with what he has really done then it may be stated:

A full validation of the RR-method for time periods longer than one day is not really possible* (Opfer and Müller-Veeh, 1983). This has also been pointed out by Douglas and Lysaker (1983). This means, the curves (a) and (b) do not exist up to now - the best one can do is to fill in some point of these curves.

A lot of work has been done on the validation of the 'yesterday recall'. We shall refer to this in the next section. It is also the topic of two other papers of this Symposium.

The curve (c) can be calculated - we shall do this in Section 6, such that the diagram finally will show one point of curve (b) and the shape of curve (c) for one example.

VALIDATION EXPERIMENTS FOR THE METHOD 'YESTERDAY RECALL'

As already stated different experiments exist concerned with validation of the method 'yesterday recall'. Attempts have been made to measure the error of underclaiming and, as far as possible, also the error of overclaiming. The experiments consisted in observing people and asking them later.

The ARF-Certitude Test (1978 and 1979) used a version of TTB, the experiments of Eadie/Lysaker (1983) and Douglas/Lysaker (1983) a version of RR.

---

* Reason: RR refers to a longer *period* of time. The situation is for the method TTB; but in total a validation of TTB is possible to a higher degree than of the method RR.
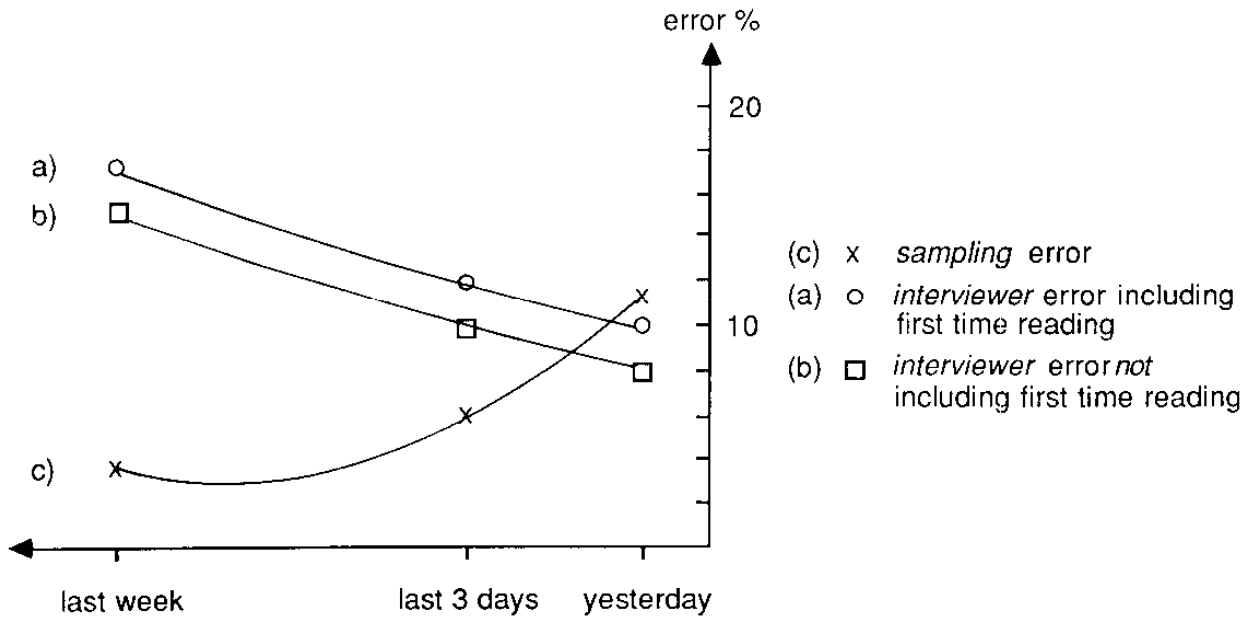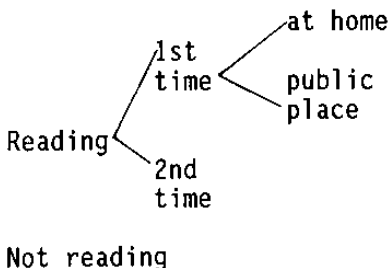
CHART 6

error %

a)
b)

20

10

(c) x   *sampling* error
(a) o   *interviewer* error including
         first time reading
(b) □   *interviewer* error *not*
         including first time reading

c)

last week   last 3 days   yesterday

CHART 7

Interview - Did you read any issue of this title yesterday?

Eadie/Lysaker (1983) P.147

	Yes		No
	Was it for the first time		Not reading
	Yes	No	
O B S E R V A T I O N — Reading — 1st time — at home / public place	correct 88 %	underclaim 10 %	underclaim 2 %
— 2nd time	92 % overclaim*	2 % correct*	6 % incorrect* but OK
Not reading	overclaim*	incorrect* but OK	correct*

* Not measured in this experiment - planned to measure

We have explained why in this paper we restrict ourself to the latter method. The formulation of the questions used by Eadie/Lysaker/Douglas was also discussed; details of the experiment's concept should be read in the original papers.

Roughly, the design was such that at two types of places (at home and in a waiting-room) persons were observed as they read a new issue of the title *Newsweek* (first-time readers).

These people were asked during the next day if they read yesterday in any issue of this title and whether it was for the first time.

This design leads to nine different combinations of behaviour and answer in the interview; three of these cells belong to correct answers, two of them correspond to incorrect answers which do not however disturb the result. Two cells contain answers which are incorrect by underclaiming, and two of them show the error of overclaiming. Eadie/Lysaker (1983) only measured the error of underclaiming. Their results are shown in **Chart 7** as well as the gaps which they planned to fill by further research.

As a summary we state:

- The observed first-time readers answered correctly in about 90% of the cases. This seems to be a satisfactory result; it should be verified with higher sample sizes and other titles.

- For the error of overclaiming, which had not been measured, an investigation was announced by the authors:

Overclaiming of type (1) can be measured by a similar design: second-time readers will be 'produced' and asked later.

Overclaiming of type (2) principally resists being measured: not-reading is hardly to be observed over a whole day. The authors propose to try a measurement by using a Belson-type

interview (which is in our opinion a much weaker criterion than observation).

So we finally come out with one single point of curve (b) for our diagram, and even this does not yet contain all aspects: correctness is only related to the possibility of underclaiming.

RELATIONSHIP BETWEEN THE LENGTH OF THE PERIOD RECALL, LENGTH OF THE PUBLICATION PERIOD AND THE SAMPLING ERROR

We now want to quantify the sampling error which occurs by using periods of recall which are shorter than the interval of publication; ie we want to construct curve (c) for the diagram as formulated above.

Here we do not depend on the results of validation experiments and therefore our knowledge is much better. In this connection we refer to the paper of Tom Corlett (1964) who treated the example 'reader during the last three days' for a monthly occurring title.

The 'factor of shortening' which has to be applied when coverage values are calculated from the values 'reader during the shorter period' depends on the relationship 'length of interval referred to in the questionnaire to length of the interval of publication'.

Some examples:

- 'Reader yesterday'. The factor of shortening for weekly occurring titles equals 7; for monthlies it equals 30 (more accurately: 30.4).

- 'Reader during the last seven days': For the weeklies there is no shortening factor (value 1.0); for the monthlies it equals 4.3 (30.4 : 7).

By the same factor we have to multiply the sampling error which is connected with the measurement of the 'readership in the short period'.

Example:

Assume the value 'reader yesterday' of 5% was measured for a weekly occurring title. The sampling error equals 1.6% (90%-level, sample size 1000)*. The estimated coverage of this title equals 5% x 7 = 35%, with a sampling error of 1.6% x 7 = 11.2%. If we had measured the coverage by using the period of recall 'last week' the sampling error would only equal 3.5%

This disadvantageous circumstance results from the U-shape of the relationship 'tolerance range depending on the measured coverage value' (see the formula* below) as shown in Chart 8.

In order to complete the desired diagram by curve (c) we finish the calculation above by using different periods of recall (Table 1).

SHORTENING THE PERIOD OF RECALL: DIAGRAM WITH THE ADVANTAGES AND DISADVANTAGES

We have said it would be desirable to construct three curves in order to decide whether it is worthwhile to measure coverages by using shortened periods of recall in the interview.

Drawing the diagram with the results available up to now is shown in Chart 9. One point of curve (a) and the total curve (c) may be filled in.

SUMMARY AND CONCLUSIONS

The purpose of this paper was to trace the idea of shortening the period of recall referred to in the interview in order to get more valid coverage values.

* Formula for the 90%-level:

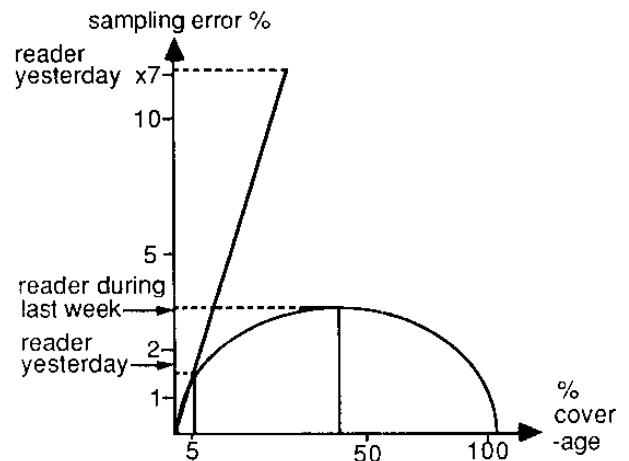$$S = 1.64 \cdot \sqrt{\frac{2 \cdot p \cdot (100-p)}{n}}$$

S = sampling error
n = sample size
p = measured coverage
(the '2' is a factor of clustering)

CHART 8

Sampling error depending on coverage

We found that this idea leads to two advantages opposed by three disadvantages. The effects of these components have to be balanced if one wants to appraise the contribution of shortening the period of recall to the measurement of coverage values.

Our work on this topic may be summarised by the following statements:

- The impracticability regarding the demand of equally and representatively distributing the interviews to the shortened periods of time leads to the consequence that periods shorter than one day can hardly be used to measure coverages.

- For the purpose of measuring coverage values by use of short periods of recall a questioning method based on RR is more appropriate than a method based on TTB - this is a matter of practicality, too.

- So we focussed our attention on a version of the so-called 'yesterday recall' which follows the method RR; it has been formulated and tried out by Eadie/Lysaker (1983) and Lysaker/Douglas (1983).

TABLE 1
Sampling error

Weekly occurring title, p = 35%
N = 1,000, level 90% (t = 1.64)

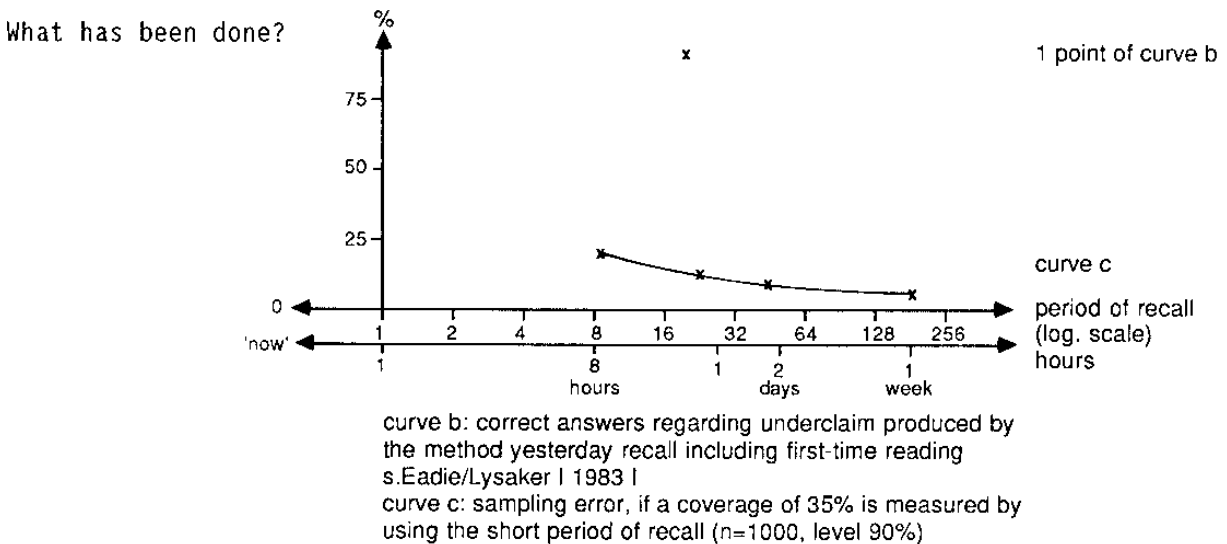| coverage %	during	sampling error %	factor	multiplied sampling error %
1.6	last 8 hours	0.92	21	19.3
5.0	yesterday	1.6	7	11.2
10.0	last 2 days	2.2	3.5	7.7
35.0	last week	3.5	1	3.5

- This question included an appendix:
'How many issues did you read yesterday' for the first time. In order to calculate coverage values, this information is the more important the shorter the period of recall is.

- Assuming that it is really unnecessary to ask for first-time reading when periods of one week or longer are used and let us consider as an example the weekly occurring titles; then the problem condenses to the question:

'How high is the level of correct answers resulting from the use of the 'yesterday recall' with inclusion of the first-time reading compared with the level of correct answers produced by the question 'read during the last week' without this inclusion?'

- This increase of the level of correct answers has to be contrasted with the increase of the sampling error which is connected with calculating the coverage from the value 'reader yesterday'.

CHART 9

curve b: correct answers regarding underclaim produced by the method yesterday recall including first-time reading s.Eadie/Lysaker I 1983 I
curve c: sampling error, if a coverage of 35% is measured by using the short period of recall (n=1000, level 90%)

- The increase of the sampling error can be calculated. We demonstrated this for the example of weekly occurring titles. Here, as an example, the sampling error multiplies by a factor of three when one measures the 'reader yesterday' in order to calculate the coverage-compared with the measurement of the 'reader during the last week' (equal sample sizes assumed).

- Whereas the increase of the sampling error is under control, one cannot say the same for the increase of the level of correct answers. This increase had to be measured by some kind of validation experiment and for this the possibilities are very limited.

- The 'reader yesterday' may be partially validated (one of the advantages of the method), and this has been done by the authors quoted above in the summary. They stated that 10% of observed first-time readers gave incorrect answers due to underclaim. Overclaiming had not been measured, but the authors planned to try it. But overclaiming is hard to measure for any version of RR - even for the period of one day (one cannot observe non-readers during a whole day).

- Even more difficult and perhaps impossible is the validation of the RR method for longer periods of recall (Opfer and Muller-Veeh, 1983).

- So finally the degree of improvement which may be expected by shortening the period of recall is hardly measurable with the tools available today (we look forward to electronic scanning methods).

- One has to be content with the fact that the method 'yesterday recall' has

a chance of being validated to a high degree. And this may be used indirectly for validating other methods - a very important fact.

REFERENCES

ARF Certitude Tests

(1978) No.1 'Measures of magazine exposure: laboratory waiting-room' New York: Advertising Research Foundation

(1978) No.2 'Measures of magazine exposure: Yesterday recall of subscriber at home reading' New York: Advertising Research Foundation

(1979) No.3 'Measures of magazine exposure: Through-the-Book recall of natural waiting-room reading' New York: Advertising Research Foundation

Corlett, Tom (1964) 'The IPA National Readership Survey: Some problems and possible solutions' *Journal of the Market Research Society* p 13-24

Douglas, S and Lysaker, R L (1983) 'The audience levels produced by the 'claimed first-time reading' method' (Montreal Proceedings)

Eadie, W P and Lysaker R L (1983) 'Developing a magazine readership validating technique' (Montreal Proceedings)

LA-MED (1983) Leseranalyse medizinischer Fachzeitschriften' Bd.2

Opfer, Gunda und Müller-Veeh, Dieter (1983) 'Validierungen im Mediabereich' Zeitschriften Arbeitsgemeinschaft Media-Analyse e.V. Schriften Band 9.