Wilfried Wenzel and Rolf Speetzen
Axel Springer Verlag AG
Hamburg, West Germany

# 2.7 DEBUGGING RANDOM ERRORS FROM MEDIA ANALYSIS DATA - A NEW TYPE OF VALIDATION

## INTRODUCTION

It is well known that the results of Through-the-Book, Recent Reading and 'first time read yesterday' are influenced by numerous factors, ie wording, rotation, masthead design, number of publication, filter categories and the like. These factors affect the mental processes of the respondents in areas of title recall, comprehension, working of the memory and the tendency to deliver status-seeking answers.

By far the greatest effect is achieved by the deployment of qualifying traps in the filter system (Tennstädt, 1983). According to the proportion of such qualifying filter traps, the coverage results can swing by more than 100%. However such deviations can only be observed in the case of particular magazines. The smaller the journal and the more infrequent the readership, the greater the difference in coverage.

The theme for this paper is to demonstrate that such dramatic aberrations can be ascribed solely to 'random errors' by the respondents.

In detail, the objective of this paper is:

(1) To demonstrate how the random error hypothesis can be tested.

(2) To develop an initial and simple calculation for the random errors.

(3) To work through an example with realC results of two German MA versions for 91 titles (with different filterings).

(4) To demonstrate thereby how the different results of coverages in both MA versions can be explained solely by the different workings of the same random errors in the two versions.

(5) To calculate the coverage, cleaned of random errors, from the empirical coverage results of the 91 magazine titles in the two MA versions.

(6) To show that debugging random errors is a particular type of validation, which cannot be replaced by any other validation (except an objective measurement as described by Schreiber and Schiller, 1983).

## DATA

The database of this paper is the findings of the MA experiment of 1982 and the actual MA 1982, which were reported in Montreal in 1983 by Eva-Maria Hess and Hans-Erdmann Scheler. Both MA versions show sharp differences in the filter system. The MA 82 had a ratio of qualifying filter traps in the fine time filter of 1:3 (example for weeklies):

*MA 82: I last read or looked through this magazine*

| | |
|---|---|
| 1 | within the last seven days |
| | 8 to 14 days ago |
| 3 | 2 to 3 weeks ago |
| | longer ago |

With the MA experiment 82, on the other hand, the qualifying filter trap ratio was 2:2 (example for weeklies):

*MA experiment 82: I last read or looked through this magazine*

| | |
|---|---|
| | yesterday |
| 2 | within the last 7 days |
| | 1 to 4 weeks ago |
| 2 | longer ago |

The results of both versions are shown in **Table 1** for all 91 magazines. It is apparent, that with a larger proportion of qualifying filter traps:

- very small titles will reflect exaggerated increases in coverage, against very large titles which maintain their coverage levels.

| | *MA-actual* | *MA-experiment* |
|---|---|---|
| Fussballwoche | 1.1 | 3.6 |
| Die Zeit | 1.7 | 3.5 |
| Zuhause | 1.2 | 2.8 |

- infrequently-read titles will reflect major increases in coverage,

| | *MA-actual* | *MA-experiment* |
|---|---|---|
| Quick | 7.5 | 12.6 |
| Für Sie | 6.6 | 11.0 |
| Freundin | 6.3 | 10.1 |

- while titles which are read regularly will again remain with constant coverage levels. (Titles which are read regularly, in particular, are *Funk Uhr, Neue Post, Bild + Funk, ADAC Motorwelt, Hörzu, Fernsehwoche.*)

| | *MA-actual* | *MA-experiment* |
|---|---|---|
| Funk Uhr | 9.2 | 10.5 |
| Neue Post | 7.3 | 8.3 |
| Bild + Funk | 5.5 | 6.4 |
| ADAC Motorwelt | 24.6 | 25.1 |
| Hörzu | 27.7 | 27.2 |
| Fernsehwoche | 11.6 | 9.7 |

**TABLE 1**
**Empirical coverages of 91 magazines in two MA versions and with averages derived through the random error model**

| | *Empirical coverages* | | *Calculated coverages with the simplified random error model* | |
|---|---|---|---|---|
| | *MA 82* | *MA experiment 82* | *MA experiment 82* | *'True coverage'* |
| *Publication* | *qualifying filter traps = 1:3* | *qualifying filter traps = 2:2* | *predicted with random error model from MA 82* | *= MA 82 minus random error* |
| 1 | 7.5 | 9.2 | 8.9 | 8.3 |
| 2 | 1.2 | 2.8 | 2.6 | 0.9 |
| 3 | 1.2 | 2.7 | 2.6 | 0.9 |
| 4 | 0.7 | 2.1 | 2.1 | 0.4 |
| 5 | 0.7 | 2.1 | 2.1 | 0.4 |
| 6 | 0.8 | 2.2 | 2.2 | 0.5 |
| 7 | 1.6 | 3.0 | 3.0 | 1.4 |
| 8 | 3.1 | 4.5 | 4.5 | 3.2 |
| 9 | 8.1 | 9.5 | 9.0 | 9.5 |
| 10 | 1.2 | 2.5 | 2.6 | 0.9 |
| 11 | 1.9 | 3.2 | 3.3 | 1.8 |
| 12 | 2.2 | 3.5 | 3.6 | 2.1 |
| 13 | 9.2 | 10.5 | 10.6 | 10.2 |
| 14 | 2.8 | 4.0 | 4.2 | 2.8 |
| 15 | 6.3 | 7.5 | 7.7 | 6.9 |
| 16 | 0.6 | 1.7 | 2.0 | 0.2 |
| 17 | 1.5 | 2.6 | 2.9 | 1.3 |
| 18 | 1.9 | 3.0 | 3.3 | 1.8 |
| 19 | 2.2 | 3.3 | 3.6 | 2.1 |
| 20 | 2.9 | 4.0 | 4.3 | 2.9 |
| 21 | 0.7 | 1.7 | 2.1 | 0.4 |

continued

**TABLE 1 (continued)**
**Empirical coverages of 91 magazines in two MA versions and with averages derived through the random error model**

| | *Empirical coverages* | | *Calculated coverages with the simplified random error model* | |
|---|---|---|---|---|
| | *MA 82* | *MA experiment 82* | *MA experiment 82* | *'True coverage'* |
| *Publi-cation* | *qualifying filter traps = 1:3* | *qualifying filter traps = 2:2* | *predicted with random error model from MA 82* | *= MA 82 minus random error* |
| 22 | 1.1 | 2.1 | 2.5 | 0.8 |
| 23 | 1.2 | 2.2 | 2.6 | 0.9 |
| 24 | 7.3 | 8.3 | 8.7 | 8.0 |
| 25 | 0.5 | 1.4 | 1.9 | 0.1 |
| 26 | 1.3 | 2.2 | 2.7 | 1.1 |
| 27 | 1.7 | 2.6 | 3.1 | 1.5 |
| 28 | 2.0 | 2.9 | 3.4 | 1.9 |
| 29 | 5.5 | 6.4 | 6.9 | 5.9 |
| 30 | 19.7 | 26.1 | 21.1 | 22.5 |
| 31 | 17.3 | 22.9 | 18.7 | 19.7 |
| 32 | 7.5 | 12.6 | 8.9 | 8.3 |
| 33 | 6.6 | 11.0 | 8.0 | 7.2 |
| 34 | 7.2 | 11.2 | 8.6 | 7.9 |
| 35 | 11.9 | 15.8 | 13.3 | 13.4 |
| 36 | 6.3 | 10.1 | 7.7 | 6.9 |
| 37 | 6.3 | 9.6 | 7.7 | 6.9 |
| 38 | 10.4 | 13.7 | 11.8 | 11.6 |
| 39 | 4.7 | 7.9 | 6.1 | 5.0 |
| 40 | 9.9 | 13.1 | 11.3 | 11.1 |
| 41 | 3.2 | 6.3 | 4.6 | 3.3 |
| 42 | 4.4 | 7.4 | 5.8 | 4.7 |
| 43 | 4.3 | 7.2 | 5.7 | 4.5 |
| 44 | 3.0 | 5.7 | 4.4 | 3.0 |
| 45 | 6.3 | 9.0 | 7.7 | 6.9 |
| 46 | 12.0 | 14.7 | 13.4 | 13.5 |
| 47 | 1.1 | 3.6 | 2.5 | 0.8 |
| 48 | 3.6 | 6.1 | 5.0 | 3.7 |
| 49 | 2.6 | 4.9 | 4.0 | 2.6 |
| 50 | 8.6 | 10.9 | 10.0 | 9.5 |
| 51 | 4.1 | 6.2 | 5.5 | 4.3 |
| 52 | 1.9 | 3.9 | 3.3 | 1.8 |
| 53 | 2.0 | 4.0 | 3.4 | 1.9 |
| 54 | 6.0 | 7.9 | 7.4 | 6.5 |
| 55 | 1.7 | 3.5 | 3.1 | 1.5 |
| 56 | 4.9 | 6.7 | 6.3 | 5.2 |

continued

**TABLE 1 (continued)**
**Empirical coverages of 91 magazines in two MA versions and with averages derived through the random error model**

| | *Empirical coverages* | | *Calculated coverages with the simplified random error model* | |
|---|---|---|---|---|
| | *MA 82* | *MA experiment 82* | *MA experiment 82* | *'True coverage'* |
| *Publi-cation* | *qualifying filter traps = 1:3* | *qualifying filter traps = 2:2* | *predicted with random error model from MA 82* | *= MA 82 minus random error* |
| 57 | 4.4 | 6.1 | 5.8 | 4.7 |
| 58 | 0.9 | 1.7 | 2.3 | 0.6 |
| 59 | 1.5 | 2.3 | 2.9 | 1.3 |
| 60 | 2.4 | 3.2 | 3.8 | 2.3 |
| 61 | 0.7 | 1.4 | 2.1 | 0.4 |
| 62 | 3.8 | 4.5 | 5.2 | 4.0 |
| 63 | 1.0 | 1.6 | 2.4 | 0.7 |
| 64 | 1.1 | 1.7 | 2.5 | 0.8 |
| 65 | 1.4 | 2.0 | 2.8 | 1.2 |
| 66 | 1.4 | 2.0 | 2.8 | 1.2 |
| 67 | 3.2 | 3.8 | 4.6 | 3.3 |
| 68 | 1.2 | 1.7 | 2.6 | 0.9 |
| 69 | 1.4 | 1.9 | 2.8 | 1.2 |
| 70 | 1.8 | 2.3 | 3.2 | 1.6 |
| 71 | 2.0 | 2.5 | 3.4 | 1.9 |
| 72 | 2.0 | 2.5 | 3.4 | 1.9 |
| 73 | 24.6 | 25.1 | 26.0 | 28.2 |
| 74 | 1.3 | 1.7 | 2.7 | 1.1 |
| 75 | 2.4 | 2.8 | 3.6 | 2.3 |
| 76 | 0.3 | 0.6 | 1.7 | 0.0 |
| 77 | 0.9 | 1.2 | 2.3 | 0.6 |
| 78 | 1.3 | 1.6 | 2.7 | 1.1 |
| 79 | 1.3 | 1.6 | 2.7 | 1.1 |
| 80 | 3.9 | 4.1 | 5.3 | 4.1 |
| 81 | 2.2 | 2.2 | 3.6 | 2.1 |
| 82 | 1.6 | 1.6 | 3.0 | 1.4 |
| 83 | 1.4 | 1.4 | 2.8 | 1.2 |
| 84 | 1.1 | 1.1 | 2.5 | 0.8 |
| 85 | 0.5 | 0.5 | 1.9 | 0.1 |
| 86 | 2.2 | 2.0 | 3.6 | 2.1 |
| 87 | 1.0 | 0.8 | 2.4 | 0.7 |
| 88 | 27.7 | 27.2 | 29.1 | 31.8 |
| 89 | 3.7 | 3.0 | 5.1 | 3.9 |
| 90 | 1.9 | 1.1 | 3.3 | 1.8 |
| 91 | 11.6 | 9.7 | 13.0 | 13.0 |

'Rebugging' (MA experiment 82 predicted ↔ 'True coverage')
'Debugging' (MA 82 ↔ 'True coverage')

None of these shifts in the findings should occur, for both MA versions use identical definitions for the readers, and are based on an identical logic in the survey method.

Random errors, however, exert a vastly different effect on the results for individual titles, according to the filter system - despite the same definition for the readers and the same logic applied to the survey system. This paper poses the question whether random errors could be guilty of the differences between the findings of the actual MA 8 and the MA experiment 82.

**HOW CAN THE EMPIRICAL TEST BE APPLIED TO THE RANDOM ERROR HYPOTHESIS?**

*The random error hypothesis states: Random errors impinge on all filter stages of a media-analytic model, and for all surveyed media, with a clearly visible probability above zero.*

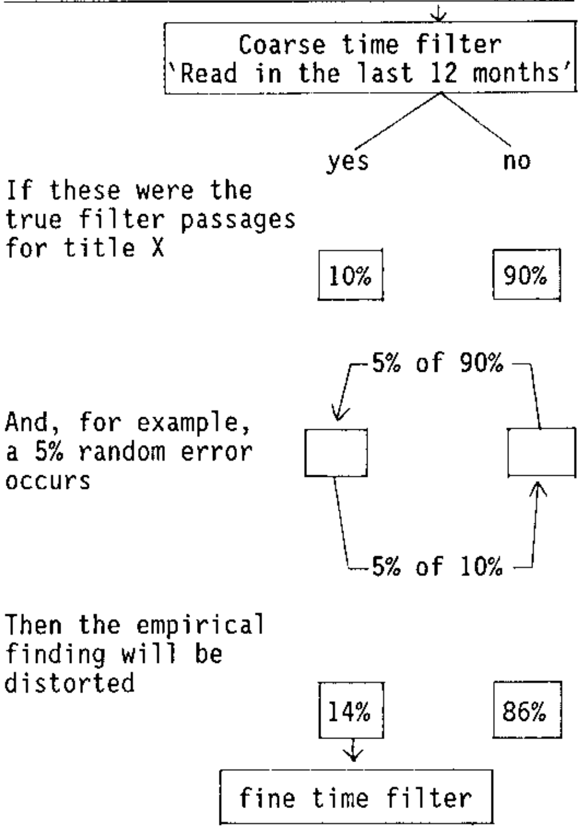
It is important, for the empirical test, to assume that random errors can occur in every filter stage of a survey model. That means, to an equal degree, in the general filter, and the coarse and fine time filters of the Recent Reading model. By random errors, we are not considering systematic errors in responses, which emerge only under special conditions and/or with special titles (eg status-seeking answers). Random errors are much more those errors which occur with equal probability in every choice of response in a multiple-choice situation. It is a further characteristic of a random error that a wrong choice is selected from the available wrong answers, *by chance*.

**Testing the random error hypothesis**

In the Recent Reading models, random errors fill the lightly-populated filter traps from the densely-filled traps.

**Graph 1** gives a hypothetical demonstration on an assumed 5% random error level.

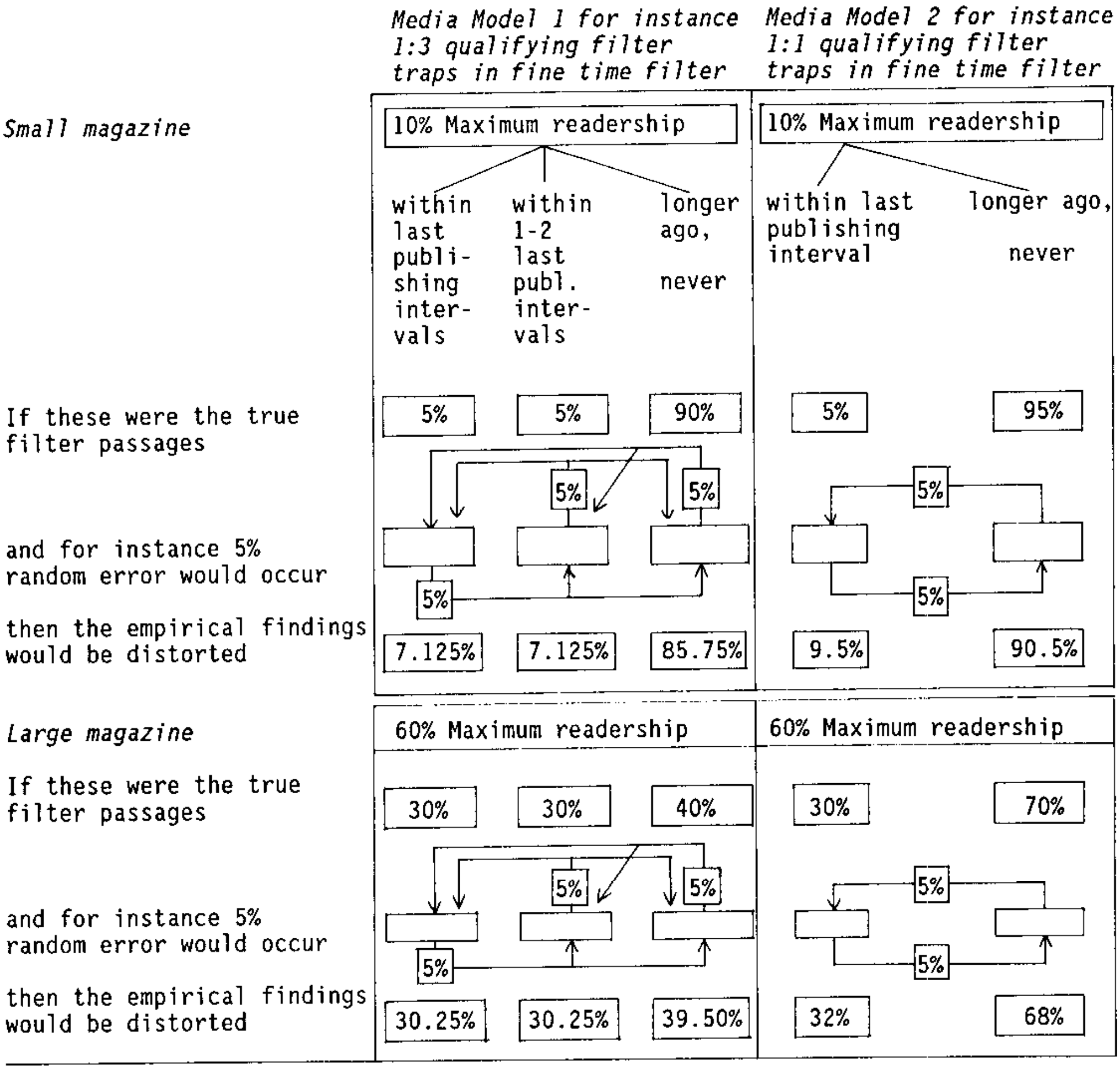
**GRAPH 1**
**Example of the effect of random errors on the catch of the filter traps (Example: coarser time filter with two answer traps)**




Accordingly for each title - by size and share of regular readers - their catch sizes in each filter trap vary, and the correct catchment for each individual title is distorted in different ways by random errors.

To identify the random errors, it is vital that the distortion of each title should be typically different according to the filter system of the media analysis (**Graph 2**). Therefore one needs two media analyses, with different proportions of qualifying traps, to check whether the differences in their findings can be ascribed to random errors. If so, the presence of random error is established.

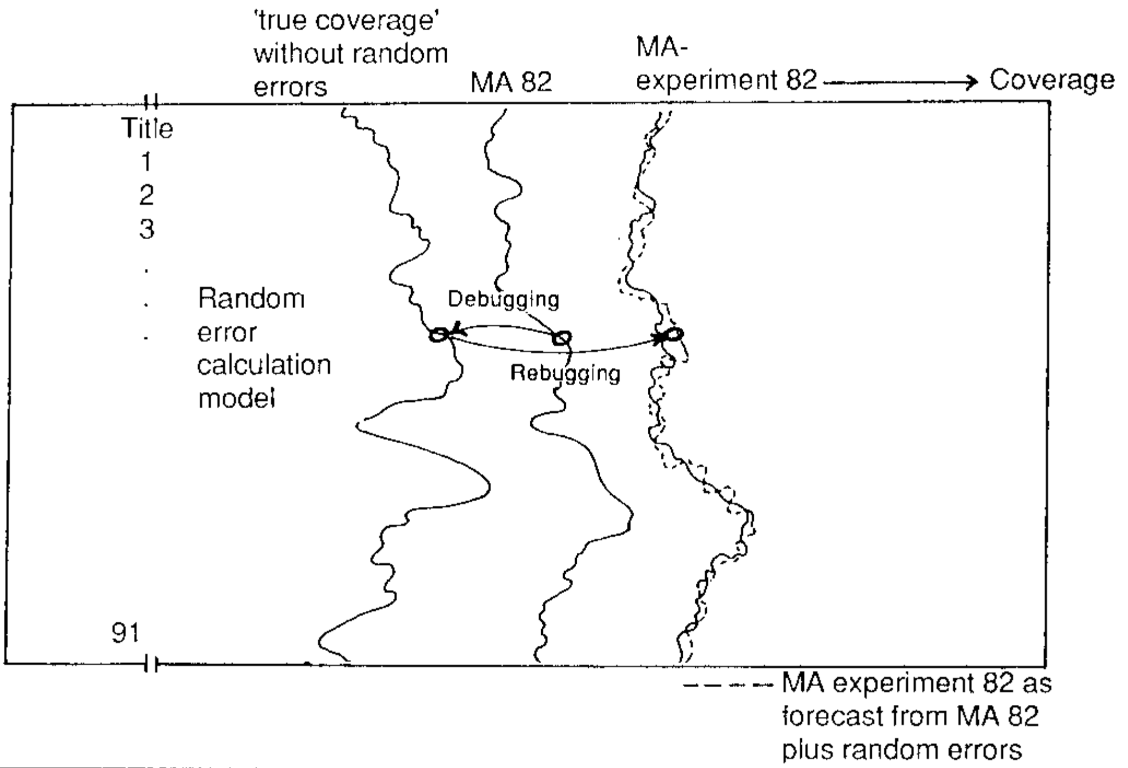
**GRAPH 2**
**The influence of random errors on magazines with different audience sizes in two different media models with different filter systems**




This paper takes the differences in the coverage values of the MA 82 and the MA experiment 82, for all 91 titles, to test the random error hypothesis. The testing is illustrated, in principle, in **Graph 3**. A calculation model simulated random error levels of 1-20%.

This cleaned the 91 surveyed coverage values in MA 82 of the primary hypothetical random error ('Debugging'). The results give the *provisional* 'true coverages' for the 91 titles. Then those random errors which should have been caught in the MA experiment 82 filter system were reinserted into the 'true coverages' ('Rebugging').

*GRAPH 3*
*Principle of the test for the random error hypothesis*




The random error model, therefore, uses the 91 empirical coverages of the MA 82 to predict the coverages of the MA experiment 82. If the prediction succeeds with a given random error level, then the presence and the volume of random errors is established. The most interesting outcome is not the random error level, but the 'true coverage' derived through the application of the random error level. In this paper, the 'true coverage' is taken as the coverage cleaned of random errors.

**HOW OTHER DISCIPLINES USE SIMILAR TESTS**

Similar testing procedures are used in many natural sciences, where 'reality' cannot be measured directly, but through indicators related to such reality. The most obvious examples are in the field of astronomy. Because of the vast distances between the heavenly bodies, coupled with the time taken by the passage of the light, only a theoretical fix can be stated for a planet or star in relation to the Earth. A series of such fixes gives a data sequence for this particular body. Then a calculation model is sought, which will exactly match the theoretical positioning. On completion, this model will deliver the undetectable accurate location from the data sequence.

The similarity to media analysis is obvious. The various survey models, changed experimentally, each deliver 'theoretical' coverage positions. If one could now use the influencing factors, integrated in a calculation model, to transfer the different theoretical coverages, then one could not only evaluate the survey model but one could also calculate the uninfluenced 'true coverage'.

## DEVELOPMENT OF A SIMPLIFIED RANDOM ERROR CALCULATION MODEL

The development of a perfect calculation model to test the random error hypothesis, as the first priority, was too burdensome. Such a perfect calculation model would have to work through all the filter traps, for all 91 titles, in the complete filter systems in both MAs.

This perfect procedure would deliver the random error effect for

- publications of different sizes

- titles with different reading frequencies.

Instead, in this paper, the initial simplified random error model is tested which concentrates on the size effect of the titles, and leaves reading frequencies to one side.

This simplified random error model can easily be stated:

It takes the filter questions of the MA 82 and MA experiment 82 as a 'labyrinth'. The general filter is the entry point, the Average Issue Readership is the exit. It can be demonstrated that a respondent with random answers will find the AIR exit

- in the MA 82 model with 2.7% probability

- in the MA experiment model with 12.5% probability

We are looking for the share of respondents with random answers within all respondents, which will account for the differences between the MA model and the MA experiment model particularly well. In theory this share of respondents with random answers lies between 0% (all answers are correct) and 100% (all answers are random, therefore the survey model acts as an unknown labyrinth for all respondents).

How, for example, random errors between 0-10% work in the MA model, is shown in **Tables 2 and 3** for given 'true

**TABLE 2**
**The effect of 1-10% random errors in the MA experiment 82 model**
**A simulation with the simplified random error model for titles of 1-30% 'true coverage'**

| *'True coverage'* / *random errors %* | *1 %* | *2 %* | *3 %* | *5 %* | *7 %* | *10 %* | *15 %* | *20 %* | *25 %* | *30 %* |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Distorted empirical coverage in the MA experiment 82 model (in %)* | | | | | | | | | |
| 1 | 1.11 | 2.10 | 3.09 | 5.07 | 7.05 | 10.02 | 14.97 | 19.92 | 24.87 | 29.82 |
| 2 | 1.23 | 2.21 | 3.19 | 5.15 | 7.11 | 10.05 | 14.95 | 19.85 | 24.75 | 29.65 |
| 3 | 1.34 | 2.31 | 3.28 | 5.22 | 7.16 | 10.07 | 14.92 | 19.77 | 24.62 | 29.47 |
| 4 | 1.46 | 2.42 | 3.38 | 5.30 | 7.22 | 10.10 | 14.90 | 19.70 | 24.50 | 29.30 |
| 5 | 1.57 | 2.52 | 3.47 | 5.37 | 7.27 | 10.12 | 14.87 | 19.62 | 24.37 | 29.12 |
| 6 | 1.69 | 2.63 | 3.57 | 5.45 | 7.33 | 10.15 | 14.85 | 19.55 | 24.25 | 28.95 |
| 7 | 1.80 | 2.73 | 3.66 | 5.52 | 7.38 | 10.17 | 14.82 | 19.47 | 24.12 | 28.77 |
| 8 | 1.92 | 2.84 | 3.76 | 5.60 | 7.44 | 10.20 | 14.80 | 19.40 | 24.00 | 28.60 |
| 9 | 2.03 | 2.94 | 3.85 | 5.67 | 7.49 | 10.22 | 14.77 | 19.32 | 23.87 | 28.42 |
| 10 | 2.15 | 3.05 | 3.95 | 5.75 | 7.55 | 10.25 | 14.75 | 19.25 | 23.75 | 28.25 |

*Example: At 5% random error, a title in the MA experiment 82 model with 1% 'true coverage' would achieve a distorted empirical coverage of 1.57%*

**TABLE 3**
**The effect of 1-10% random errors in the MA 82 model**
**A simulation with the simplified random error model for titles of 1-30% 'true coverage'**

| *'True coverage* / *random errors %* | *1 %* | *2 %* | *3 %* | *5 %* | *7 %* | *10 %* | *15 %* | *20 %* | *25 %* | *30 %* |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Distorted empirical coverage in the MA 82 model (in %)* | | | | | | | | | |
| 1 | 1.02 | 2.01 | 3.00 | 4.98 | 6.96 | 9.93 | 14.88 | 19.83 | 24.78 | 29.73 |
| 2 | 1.04 | 2.02 | 3.00 | 4.96 | 6.92 | 9.86 | 14.76 | 19.66 | 24.56 | 29.46 |
| 3 | 1.05 | 2.02 | 2.99 | 4.93 | 6.87 | 9.78 | 14.63 | 19.48 | 24.33 | 29.18 |
| 4 | 1.07 | 2.03 | 2.99 | 4.91 | 6.83 | 9.71 | 14.51 | 19.31 | 24.11 | 28.91 |
| 5 | 1.09 | 2.04 | 2.99 | 4.89 | 6.79 | 9.64 | 14.39 | 19.14 | 23.89 | 28.64 |
| 6 | 1.11 | 2.05 | 2.99 | 4.87 | 6.75 | 9.57 | 14.27 | 18.97 | 23.67 | 28.37 |
| 7 | 1.12 | 2.05 | 2.98 | 4.84 | 6.70 | 9.49 | 14.14 | 18.79 | 23.44 | 28.09 |
| 8 | 1.14 | 2.06 | 2.98 | 4.82 | 6.66 | 9.42 | 14.02 | 18.62 | 23.22 | 27.82 |
| 9 | 1.16 | 2.07 | 2.98 | 4.80 | 6.62 | 9.35 | 13.90 | 18.45 | 23.00 | 27.55 |
| 10 | 1.18 | 2.08 | 2.98 | 4.78 | 6.58 | 9.28 | 13.78 | 18.28 | 22.78 | 27.28 |

*Example: At 5% random error, a title in the MA 82 model with 1% 'true coverage' would achieve a distorted empirical coverage of 1.09%*

coverages' taken into account. These two tables will demonstrate the effect random errors exert on the magnitude of the title's coverage. Whether a given random error level is actually able to explain the differences in the results for 91 titles in the two MAs will be shown in the next section.

## CHECKING THE SIMPLIFIED RANDOM ERROR MODEL OF THE MA 82 AND THE MA EXPERIMENT 82

The simplified random error model simulated random error levels of 1, 2, 3, ...20%.

The 14-15% random error level was the optimum to forecast the empirical results of the MA experiment 82 from the empirical results of the MA 82 (Table 4).

How good is the forecast?

The sum total of the squared differences between MA 82 and MA experiment 82 is 358. Of this, the labyrinth calculation model accounts for exactly 50% at an assumed random-error level of 14-15%. With this the sharply differing effect of the same share of random-choice answers in both MA model comes through in a very significant manner.

Positively stated:

*The systematic shifts in the results between large and small titles in both MAs can only be explained by the different effects of the same 14% random error level alone.*

The specific increases in coverages, according to the coverage level in the MA experiment 82, came through - as in this model, answer errors have a stronger effect on AIR.

**Table 1** shows, for each of the 91 titles, the predicated coverages taken from the MA 82 to give forecasts for the MA experiment 82, using a 14% random error level. The quality of the forecast is already very satisfactory. But one has to take into account that it is limited by the simplification of the random error model, which only reacts to effects related to different publication sizes, and not to the effects of different reading frequencies.

The coverage increases related to reading frequency differences in the MA experiment 82 could similarly be tested by the perfect calculation model, noted above.

The 'true coverages' are also shown in **Table 1**, that is the values for all 91 titles, cleaned by the simplified random error model. These 'true coverages' lie close to the MA 82, but are in no way identical.

## RANDOM ERROR DEBUGGING AS A VALIDATION METHOD

This paper records the presence of random errors. It shows clearly that random errors

- have an effect on all conceivable filter systems in media analysis

- have in fact a different effect for the different types of journals.

That means that there are no actual 'correct' and no 'fair' filter systems. It is not enough, even with the most refined validation, to establish which filter system best delivers accurate coverages, because, in every case, some types of journals will be disadvantaged by random error, while others will be enhanced.

The distortion in results can be set aside by a *random error validation* if the experience of this analysis is confirmed in the future.

In practice, a random error validation could be executed thus. Every media analysis uses two different filter systems in equal splits. The random error level is calculated from the differences in the findings. With this random error level, the empirical coverages can be cleaned for all titles. It is immaterial how many qualifying filter traps are deployed in the two media analyses - the coverages, after cleaning, must be about the same for all titles individually, whichever MA version supplies the raw coverage data.

**TABLE 4**
**Which random error level best accounts for the differences between the coverages of MA 82 and MA experiment 82?**

| *Tested random error level* | *Sum of the squared coverage differences in the 91 titles between actual and predicted MA experiment 82 coverage data (forecast from the MA 82 with the simplified random error model)* | |
|---|---|---|
| *%* | *Sum of the squares* | |
| 0 | 358 | |
| 5 | 254 | |
| 10 | 193 | |
| 12 | 182 | |
| 13 | 178 | |
| 14 | 176 | Best explanation! |
| 15 | 176 | |
| 16 | 178 | |
| 17 | 182 | |
| 20 | 203 | |