Too Many Technical Details Hinder Recall of Poll Results

by Wolfgang Wichmann

The findings support the hypothesis that poll reports that include many technical details about how the polling methods were encoded and stored were less successful than information in poll reports that included fewer technical details.

In the last 40 years a variety of studies has fairly clearly answered the question of what journalists do with political poll results.[1] This study was done to answer the question: What do readers do with poll results?

When analyzing the quality of political poll reporting, most researchers have focused on the journalistic output—on articles that include poll results. Referring to results of content analyses, several scholars have described structural and formal aspects of political poll reporting.[2] To analyze the articles' formal quality, the "Standards for Minimal Disclosure" by the American Association of Public Opinion Research (AAPOR) have become widely prominent. These guidelines were issued in 1969 and state that public opinion researchers should "report or make available" at least eight technical details about how a poll was conducted in every poll report.[3] Because of the guidelines' simplicity, researchers in several countries analyzed poll reports in the media, with regard to technical details mentioned and many were attracted by the philosophy of the more, the better. In 1990, Rollberg, Sanders and Buffalo argued for the publication of technical details in poll reports:

> *If newspapers' stories routinely included all eight disclosure standards, along with clearly understandable definitions of those standards, they would be serving two of the purposes of journalism: reporting the news and educating the readers.*[4]

Wichmann is a doctoral candidate in the Department of Communication Science at the University of Hohenheim in Germany.

The researchers Meyer and Jurgensen opposed this practice in 1991:

> *Putting too much in the readers´ hands can create a noisy clutter that reduces the amount that reaches their heads. And that is the basic flaw of the 100 percent or more-is-better model.*[5]

The dispute for and against the total-conformity model has remained unresolved. This study takes a new approach to give more clarity to this discussion. The author believes the quality of political poll reporting should not be measured by the amount of information provided to the readers alone. Instead, an analysis of how the given information is actually processed by the readers is recommended.

The following two steps were done:

- The limited capacity model of mediated message processing by Annie Lang was applied to printed news to analyze how poll reports are cognitively processed by the readers and transferred into their short-term memory.
- Referring to self-evaluative measures on poll reports, the readers' rating of reliability, credibility, difficulty and informativeness of poll reports was investigated.

AAPOR's disclosure standards or similar criteria are still present in the education of journalists. Although it never became mandatory to include all eight criteria, the results of this study can be beneficial to the education of political poll reporters. The relevance of this study becomes even clearer when reminded of an attempt in the U.S. in 1972 to pass a law that would have regulated the amount of technical information in reports about political polls.[6]

Objectives and Method

After AAPOR published its standards of minimal disclosure in 1969, other organizations followed. This trend was accompanied by the publication of Philip Meyer's books on precision journalism[7] and the request of several researchers for higher standards in poll reporting. Given the great number of studies on technical information in poll reports, the following eight items were generally applied while talking about standards for minimal disclosure for poll reports:

- Sample size
- Firm that commissioned the poll
- The exact wording of the voting intention question
- The margin of error
- Definition of population for which the survey represents
- Method used in the survey
- Time of fieldwork
- Name of responsible poll institute

The demur that technical information might distract the readers from the most important aspects was sometimes mentioned, but with the exception of Lordan, no one acted on it.[8] Lordan tested the memorization of technical information in poll reports and found that all technical information was memorized on a very low level. There was no significant difference between readers who had read a story with all eight AAPOR criteria either included in the text or presented in a sidebar. He concluded:

> *The findings of the study do not support the idea that the inclusion of details assists readers in understanding statistically oriented newspaper stories. Respondents recalled an average of two of the eight details, and even when they did recall this information, had trouble understanding what it meant.*[9]

It seems self-explanatory that the quality of poll reporting does not only depend on the amount of presented methodological information. A qualitative aspect about how the results are interpreted, fitted into a story and used to explain a certain situation, is important as well. But it is just as important to know whether the methodological information ever makes it to the readers' heads, even though it was placed in their hands in the form of a news article.

According to the limited capacity model, information processing is seen as a combination of three subprocesses—encoding, storage and retrieval—that are performed at the same time. While encoding, bits of information are transformed from the sensory store(s) into "activated mental representations in working or short-term memory."[10] These representations are then linked with existing information in the associative memory network. This process is called storage and depending on the associations and links, the result can range from "poorly stored" to "thoroughly stored" information. Retrieval is the final subprocess in Lang's model and is described as "the process of searching the associative memory network for a specific piece of information and reactivating it in working memory."[11]

Applying Lang's model of information processing to a relatively new approach. Only Gibbons and his colleagues applied the model to investigate the believability of headlines in tabloids over time.[12] To understand how printed information is processed, Lang's model had to be applied to the specific characteristics of printed information. This study focused on 'encoding' and 'storage.' 'Retrieval' was not tested. Encoding was measured with a recognition-test and storage with a cued recall test.

The independent variable in part A of this study is the amount of technical information in each pre-election poll report, according to the AAPOR disclosure standards. The amount of technical information within the text is divided into three levels: high (eight criteria), medium (four criteria) and low (one criterion).[13] The dependent variables are recognition (encoding) and cued recall (storage).

Hypotheses and Research Question

H1:

Recognition (encoding) should be lowest for paragraphs including all eight AAPOR criteria.

H2:

Recognition (encoding) should be highest for paragraphs including fewer AAPOR criteria.

H3:

Cued recall (storage) should be lowest for paragraphs including all eight AAPOR criteria.

H4:

Cued recall (storage) should be highest for paragraphs including fewer AAPOR criteria.

According to expert literature, several aspects are said to have an impact on how intensively poll reports are read and thus how well the content is memorized: How carefully the stories are read and how hard the readers exert themselves depends on:

- how interesting the subject is
- how relevant the information is
- whether the reader wants to remember it.

Additionally, writing style and the form of presentation also have an impact.[14] Lang refers to earlier studies with TV messages: "They found […] that increasing structural complexity decreased recognition memory for messages globally."[15]

In a study by Thorson and Lang familiar topics were treated as low cognitive load (easy), and unfamiliar topics were treated as high cognitive load (difficult).[16] Recognition and cued recall were better for familiar topics when compared to unfamiliar topics. The amount of resources that is generally allocated depends mainly on the reader's needs and goals. Readers who are interested in politics are expected to allocate more resources to the content.

H5:

Recipients who are more interested and knowledgeable in politics and political polls should do better in encoding (recognition) of information in pre-election poll reports.

H6:

Recipients who are more interested and knowledgeable in politics and political polls should do better in cued recall (storage) of information in pre-election poll reports.

For an analysis of the quality of poll reports, the individual evaluations of reliability, difficulty, credibility and informativeness of pre-election poll reports are investigated.

RQ1:

How are pre-election poll reports rated (seven-point scale) and evaluated (open-ended comments) in terms of reliability, difficulty, credibility and informativeness, if the stories are altered by the number of AAPOR criteria included in the text (low, medium and high)?

One-hundred seventy-nine recipients participated in this study. Ten were non-students, and 169 recipients were recruited from undergraduate classes at a large Midwestern university in the United States. Part A of this study is an experiment based on Lang's limited capacity model, and Part B asked for evaluative self-report measures by the recipients. Each participant had to complete a pre-test questionnaire, read an article[17] and then complete a cued recall, a recognition and a post-test questionnaire.

Experimental Design

The design for the experiment was a 2 x 2 x 3 (Repetition x Low / High AAPOR criteria x Question Order) within-subject factorial design. The experimental design allows a within-subject comparison for high versus low and medium level of AAPOR criteria in the articles, but only a between-subject comparison for medium versus low level of AAPOR criteria in the articles. All data were evaluated by using the software SPSS.

Stimulus Material

The stimulus consisted of two versions of a slightly altered *Los Angeles Times* pre-election report (Articles A and B), originally written by *Times* staff member Ronald Bernstein.[18] Because the original article did not include graphics or sidebars, the stimulus material was only text-based as well. To assure high external validity, the articles were designed and printed in the *Los Angeles Times* corporate layout, cut out and photocopied. Bernstein's article was altered in order to present pre-election poll results of four different heavily contested states in the United States during the election campaign that preceded the Nov. 8, 2006, elections for U.S. Congress and Senate.

Each of the two article versions included four paragraphs with each paragraph presenting the pre-election poll results of one state (Ohio, Missouri, Virginia or New Jersey). The first article (A) presented pre-election poll results by alternating a high amount of AAPOR criteria for two surveys (Virginia and New Jersey), with a low amount of AAPOR criteria for the other two surveys (Ohio and Missouri). The second article (B) presented pre-election poll results by alternating a high amount of AAPOR criteria for two surveys (Virginia and

New Jersey) with a medium amount of AAPOR criteria for two other surveys (Ohio and Missouri).

Each post-test questionnaire included one single paragraph of the article that the recipients had just read. Each paragraph was randomly assigned and asked for four dimensions of credibility and comprehensibility of the given pre-election poll report. All post-test questionnaires asked similar questions related to the four dimensions.[19] All four questions were followed by the open-ended request: "Please comment why?" The personal evaluation was required to determine if the recipients evaluated stories differently, depending on the number of AAPOR criteria included. The evaluation of these aspects was measured on a seven-point scale between reliable/not reliable, difficult/not difficult, credible/not credible and informative/not informative.

For part B of this study, the independent variable was again the level of AAPOR criteria in the pre-election poll report (LC = low, MC = medium and HC = high number of criteria). The dependent variables were the evaluation of reliability, difficulty, credibility and informativeness.

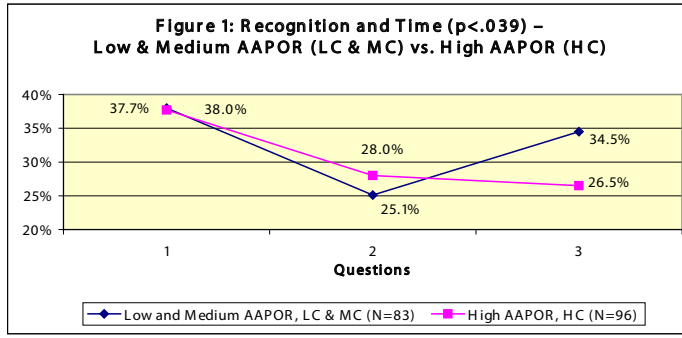
As expected, politically interested readers were significantly more successful in storing information of pre-election poll reports than were less politically interested readers. Additionally, the politically interested recipients seemed more successful in encoding information from the beginning to the end of a poll story.

Findings

Hypothesis 1 predicted that encoding (recognition) would be better for stories with fewer criteria (LC & MC) compared to stories with all eight AAPOR criteria (HC).

The main effect for criteria (with two levels, LC&MC and HC) was significant ($F(1) = 12.722$, $p < .000$, Eta squared = 0.067). As expected, stories with fewer AAPOR criteria were encoded better (LC&MC = 36.6 percent) than stories with all eight AAPOR criteria (High = 31.2 percent). The number of AAPOR criteria explained 6.7 percent of the variance in recognition.

The recognition test included five questions per story. The first three ques-

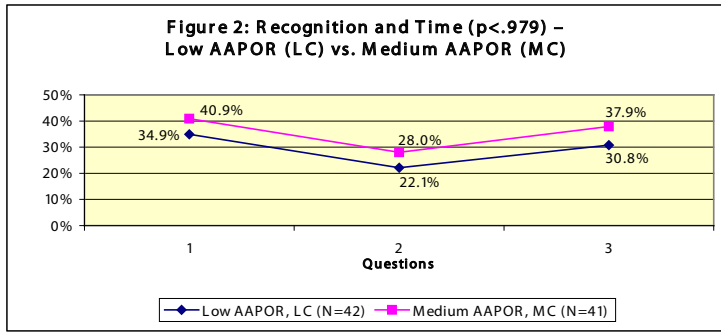


tions were asked to determine whether information was encoded better at the beginning, the middle or at the end of a story. Questions four and five were designed to investigate whether information about AAPOR criteria was encoded by the recipients.

For the first three questions, the interaction of criteria and time was significant ($F(2) = 3.275$, $p < .039$, Eta squared = .018). As can be seen in Figure 1, encoding was similar at the beginning of each story (LC&MC = 38 percent, HC = 37.7 percent). But for stories with all eight AAPOR criteria, encoding diminishes as the story continues. At the end of the text, encoding increases for stories with fewer criteria (LC&MC = 34.5 percent) compared to stories with all AAPOR criteria (HC = 26.5 percent).

The main effect of AAPOR criteria for the encoding of AAPOR criteria information (final two questions only) was also significant ($F(1) = 16.946$, $p < .000$, Eta squared = .087). For stories with fewer AAPOR criteria, encoding of AAPOR criteria information was better (LC&MC = 42.6 percent) than for stories with all eight AAPOR criteria (HC = 32 percent). The difference in AAPOR criteria was responsible for 8.7 percent of the variance.

Hypothesis 2 predicted that encoding (recognition) should be better for stories contain-



ing a low (LC) compared to a medium number of AAPOR criteria (MC).

For the encoding of general information (first three questions only), the main effect for AAPOR criteria was significant (F (1) = 4.313, p < .039, Eta squared = .024). Contrary to the prediction, information was encoded better for stories with a medium number of AAPOR criteria (MC = 35.6 percent) compared to stories with a low number of AAPOR criteria (LC = 29.3 percent).

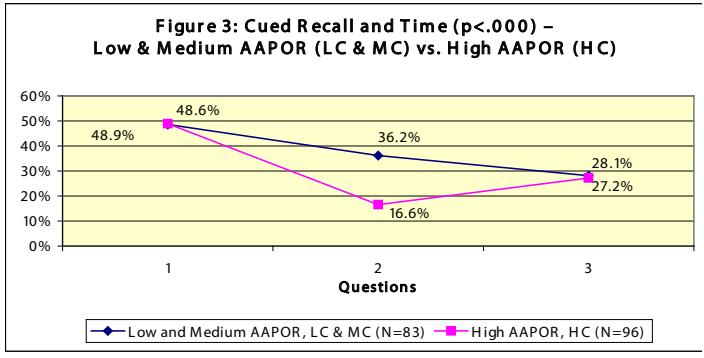
For the first three questions, the interaction of AAPOR criteria and time was not significant (F (2) = .022, p < .979, Eta squared = .000). The differences of means in encoding for questions one to three can be seen in Figure 2. The graph shows that information was generally encoded better for stories including a medium number of AAPOR criteria (MC), compared to stories including a low number of AAPOR criteria (LC).

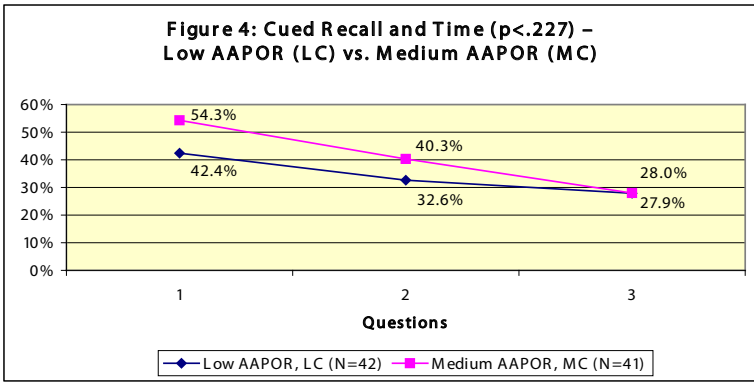
For the encoding of AAPOR criteria related information (final two questions) in stories with a low number of criteria (LC) and a medium number of criteria (MC) the main effect for criteria was not significant (F (1) = .122, p < .727, Eta squared = .001). However, the means for recognition (LC = 43.3 percent and MC = 41.9 percent) are in the predicted direction.

Hypothesis 3 predicted that cued recall would be better for stories containing fewer AAPOR criteria (LC and MC) than for those with all eight criteria (HC).

The main effect for criteria (with two levels, LC/MC and HC) was significant (F (1) = 12.19, p < .001, Eta squared = .064). As expected, stories with fewer AAPOR criteria were remembered better (LC&MC = 37.6 percent) than those with many AAPOR criteria (HC = 30.9 percent). The number of AAPOR criteria explained 6.4 percent of the variance in cued recall.

In addition to the main effect for criteria, the interaction of criteria and time was also significant (F (2) = 11.920, p < .000, Eta squared = .063). The effect size was 6.3 percent. Figure 3 illustrates this interaction. As can be seen, initially and at the end of a story, the number of criteria in the message has no effect on cued recall; in the middle, however, the information input of articles with all eight AAPOR criteria reduces storage capacity more quickly, compared to the





recipients of low / medium AAPOR stories.

Hypothesis 4 predicted that cued recall should be better for stories containing a low compared to a medium number of AAPOR criteria.

The main effect for criteria approached significance ($F(1) = 3.191$, $p < .076$, Eta squared = .018). The means show that contrary to the prediction, stories with a medium number of AAPOR criteria may be remembered slightly better (MC = 40.9 percent) than paragraphs with a low number of AAPOR criteria (LC = 34.3 percent).

The interaction of criteria and time on the cued recall data was not significant ($F(2) = 1.491$, $p < .227$, Eta squared = .008). However, as can be seen in Figure 4, the difference in criteria makes a difference in storage early in the message (LC = 42.4 percent, MC = 54.3 percent). But by the end of the message, all conditions are remembered equally well (LC = 27.9 percent, MC = 28 percent). Thus, information presented early in the message does better with a medium level of criteria.

Hypothesis 5 predicted that encoding (recognition) on average should be performed better by recipients who are generally more interested in and knowledgeable about politics and political polls.

Four relevant variables of the pre-test questionnaire were recoded in a new latent variable called "interest in politics and polls." Cronbach's Alpha[20] as a measure of reliability computed .8415. The new latent variable was split in three similar sized groups of recipients with little interest in politics and polls (N=59), some interest in politics and polls (N=55) and much interest in politics and polls (N=65). By computing an analysis of variance between groups (ANOVA) with the between-subjects variable "interest in politics and polls," the hypothesis was not supported ($F(2) = 0.178$, $p < .837$, Eta squared = .002). Recipients with much interest / knowledge in polls encoded more information (32.4 percent) than recipients with some (30.9 percent) or few (31.5 percent) interest / knowledge in

politics and polls. However, the differences were not significant.

Hypothesis 6 predicted that storage of information (cued recall) on average should be performed better by recipients who are generally more interested and knowledgeable about politics and political polls.

By using the latent variable "interest in politics and polls," another analysis of variance between groups (ANOVA) was computed. The hypothesis was supported. For the correlation of the cued recall results and the poll/political interest of the recipients, the effect of interest/knowledge in politics and polls was significant ($F(2) = 7.011$, $p < .001$, Eta squared = .074). As predicted, information was remembered better by recipients with much interest/knowledge in politics and polls (41.2 percent) compared to recipients with some (32.3 percent) or little (28.5 percent) interest/knowledge in politics and polls.

Just like the "average" reader in this study, politically interested readers had problems encoding and storing information in articles with all eight AAPOR criteria. Additionally, politically uninterested readers had massive problems encoding information late in high-level criteria stories (HC). Contrary, politically interested readers were able to keep encoding results stable after the positive effects of a new story (primacy effect) had decreased.

Research question one asked how the average scores differed in the four dimensions of evaluative self-report measures (reliability, difficulty, credibility and informativeness), if the amount of technical information (AAPOR criteria) was changed in a pre-election poll report.

The answers to research question number one are given in two separate sections—by analyzing the ratings on the seven point scales and by content analyzing the written comments for each dimension.

For further analyses, the results of the evaluative self-report measures were regrouped according to the criteria level of the poll-report that had been evaluated on the post-test questionnaire (LC, MC or HC criteria level).

By computing analyses of variance (ANOVA) for all four dimensions and the three levels of AAPOR criteria, the difference in means for reliability of low level (LC) and high level (HC) criteria articles was significant ($F = 4.049$, $p < .046$). The difference in AAPOR criteria (independent variable) was responsible for 2.9 percent of the variance in evaluation of reliability for the two groups of articles. Pre-election poll reports with only one AAPOR criteria were rated significantly more reliable than paragraphs with all eight AAPOR criteria. The recipients of this study rated the "reliability" of the given pre-election poll reports with a mean of 4.793 on a seven-point scale, with one being "does not seem reliable" and seven being "seems very reliable."

The differences in means for all other dimensions and levels of AAPOR criteria were not statistically significant.[21] The analysis of ratings on the seven-point scale should not be overrated. A first analysis of personal comments revealed that the recipients often interpreted the given scale differently. Together with a

thorough content analysis of the recipients' personal comments, the scale ratings offer more reliable information

Overall, the pre-election poll reports of this study were rated rather reliable than unreliable (4.8 on a seven-point scale). Surprisingly, pre-election poll reports with all eight AAPOR criteria were rated significantly less reliable than articles with only one AAPOR criterion. A content analysis of comments on the recipients' evaluation revealed possible reasons:

The comments showed that statistics in general are perceived favorably by recipients ("It's backed up with statistics and historical facts") and thus might generally lead to a perception of more reliability ("Because it gives the results of votes and margin").

However, if specific technical information about polls was provided, many recipients used the given information to rate the poll results as not as reliable ("the survey includes people with telephones, not everyone" or "small sample number, large margin of error"). Overall, the method of how the survey was done (telephone survey), the margin of error (4 or 5.5 percentage points) and the sample size (385 or 593 likely voters) led to a negative interpretation of the survey results. In contrast, the names of the responsible polling organization (the fictitious "Triple A Research Center"), the poll sponsor (Bloomberg and the *Los Angeles Times*) and the newspaper in which the story was published (*Los Angeles Times*) led to an interpretation of greater reliability.

Further, one has to keep in mind that the given survey results might indeed be criticized for a relatively small sample size and a relatively high margin of error. However, further research should investigate whether the interpretation of reliability by recipients is directly related to general skepticism toward technical details of polls or to the factual qualities of the pre-election poll.

The pre-election poll reports in this study were rated more credible than not credible by the recipients (4.8 on a seven point scale). A content analysis of the comments provided by the recipients suggests that the numbers and statistics in the reports were not the most important information to evaluate credibility. Readers instead complimented the "detailed explanation," "a lot of supporting material" or the fact that "they cited sources well."

Not surprisingly, the name of the polling institute (the fictitious Triple A Research Center) and the names of the responsible media groups (Bloomberg and *Los Angeles Times*) were often used to rate the poll reports as more credible. In contrast, if the names of the institutions were not given or found by the readers, the poll reports were generally rated as less credible.

Conclusion

The findings of this study provide a new perspective on how recipients perceive and store information in pre-election poll reports. But there is no doubt that further research is needed. The reader should keep in mind that, although the number of recipients is acceptable, the majority of participants in

this study were college students. Future research should replicate this study with a more representative sample and find out whether the findings are stable for newspaper reports on other topics than political surveys and whether the major patterns are similar for different types of media.

Nevertheless, the findings of this study might be taken as additional input for new guidelines on how political poll results should be published in newspapers. The following four suggestions might be used to supplement the standards for minimal disclosure by renowned organizations such as AAPOR or WAPOR and to complement the education of professional journalists.

Pre-election poll reports with all eight AAPOR criteria were encoded and stored less successfully than were reports with fewer criteria. Additionally, encoding of technical information was diminished in articles including all eight criteria. Thus, technical details of political polls should be reduced to relevant key facts in the written text of a poll report.

Contrary to the prediction, poll stories with only one AAPOR criterion were encoded and stored less successfully than reports with a medium number of AAPOR criteria. Moreover, as written comments of recipients suggested, technical details are relevant for the reliability rating of poll reports. This leads to the suggestion that relevant key facts should include the name of the polling institute (source of information) and further AAPOR criteria, if they are newsworthy to the readers because of exceptionality or oddness.

Information early in the article was generally encoded and stored better by the recipients. Additionally, texts with only one AAPOR criterion reduced storage quality from the beginning to the end of a story, possibly because of being less interesting. Thus, if AAPOR criteria are exceptional or odd, they should be explained for statistically inexperienced readers, if possible, early in a poll report.

As expected, politically interested readers were significantly more successful in storing information of pre-election poll reports than were less politically interested readers. Additionally, the politically interested recipients seemed more successful in encoding information from the beginning to the end of a poll story. Yet, further research is needed to verify this trend. However, information processing of articles with eight AAPOR criteria suffered, compared to articles with fewer criteria. Thus newspapers with an assumed politically interested readership should not overload their audience with too many technical details in the text.

Although these recommendations might need further discussion, as a journalist by trade, the author feels committed to support colleagues Meyer and Jurgensen in their denial of the "100 percent or more-is-better model."[22] The results of this study showed that heavy use of technical details in the text of a poll story hinders encoding and storage of important information.

But for ethical and professional reasons and as a service, technical details about how a poll was conducted must be available to the interested audience. As Wilhoit and Weaver stated in their "Newsroom Guide to Polls and Surveys:"

By reporting survey results as clearly and accurately as possible, a journalist can help to insure that survey research standards remain high.[23]

A survey masthead, as suggested in Germany by Walter Rudolf in 1983, is one possibility.[24] A likely masthead for example could contain the AAPOR criteria and display them beside of a written text. Or the criteria could also be added as explanatory footnotes to graphics in poll reports. Also, in times of the World Wide Web, the technical details of how a poll was conducted should always be made available online.

Notes

1. David L. Paletz, Jonathan Y. Short, Helen Baker, Campell B. Cookman, Richard J. Cooper and Rochelle M. Oeslander, "Polling in the media: Content, credibility, and consequences," *Public Opinion Quarterly* 44, no. 4 (winter 1980): 495-513; M. Mark Miller and Robert Hurd, "Conformity to AAPOR Standards in Newspaper Reporting of Public Opinion Polls," *Public Opinion Quarterly* 46, no. 2 (summer 1982): 243-249; Michael B. Salwen, "The reporting of public opinion polls during presidential years. 1968-1984," *Journalism Quarterly* 62, (summer 1985): 272-277, Frank Brettschneider, "Wahlumfragen und Medien. Eine empirische Untersuchung der Presseberichterstattung über Meinungsumfragen vor den Bundestagswahlen 1980 bis 1994 (Election Polls and the Media. Empirical Research on the Reporting of Opinion Polls before the Federal Elections 1980 through 1994), *Politische Vierteljahresschrift* 37, no.4 (December 1996): 475-493; Robert Andersen, "Reporting Public Opinion Polls: The Media and the 1997 Canadian Election," *International Journal of Public Opinion Research* 12, no. 3 (autumn 2000): 285-298.

2. Paletz et al., "Polling in the media: Content, credibility, and consequences," Miller and Hurd, "Conformity to AAPOR Standards in Newspaper Reporting of Public Opinion Polls," Andersen, "Reporting Public Opinion Polls: The Media and the 1997 Canadian Election."

3. AAPOR, "Standards for Minimal Disclosure," aapor.org, as revised in 2005, <http://www.aapor.org/Disclosure_Standards/1570.htm> (Dec. 2, 2009)

4. Jeanne N. Rollberg, Luther W. Sanders and M. D. Buffalo, "Down to the Wire: How Six Newspapers Reported Public Opinion Polls During the 1988 Presidential Campaign," *Newspaper Research Journal* 11, no. 4 (fall 1990): 91.

5. Philip Meyer and Karen Jurgensen, "Beating disclosure to death: A rejoinder to Rollberg, Sanders and Buffalo," *Newspaper Research Journal* 12, no. 3 (summer 1991): 5.

6. H.R. 5003 *Public Opinion Polls. Hearings before the Subcommittee on Library and Memorials of the Committee on House Administration. House of Representatives. Ninety-Third Congress. First Session on H.R. 5003 to provide for the Disclosure of certain Information related to certain Public Opinion Polls*, Sept. 19, 20, 21; Oct. 5, 1972; (Washington: U.S. Government Printing Office, 1973).

7. Philip Meyer, *Precision Journalism: A Reporter's Introduction to Social Science Methods* (Bloomington, London: Indiana University Press, 1973); Philip Meyer, *The New Precision Journalism* (Bloomington, Indianapolis, IN.: Indiana University Press, 1991).

8. Edward J. Lordan, "Do methodological details help readers evaluate statistic-based stories?," *Newspaper Research Journal* 14, no. 3/4 (summer/fall 1993): 13-19.

9. Lordan, "Do methodological details help readers evaluate statistic-based stories?," 18.

10. Annie Lang, "The limited capacity model of mediated message processing," *Journal of Communication* 50, no. 1 (winter 2000): 48.

11. Lang, "The limited capacity model of mediated message processing," 50.

12. Jeffrey A. Gibbons, Angela F. Lukowski and W. Richard Walker "Exposure increases the believability of unbelievable news headlines via elaborate cognitive processing," in: *Media Psychology* 7, no. 3 (2005): 273-300.

13. The average number of AAPOR criteria in pre-election poll reports ranges between three and four, according to earlier studies by Salwen (1985) and Miller & Hurd (1982). Following a study by Marton & Stephens the following four criteria were included in all pre-election poll reports with a medium amount of AAPOR criteria (MC): Sponsor, Poll Institute, Sample Size and Population Studied. In all pre-election poll reports with a low amount of AAPOR criteria (LC), only the Sponsor was mentioned, Krisztina Marton and Lowndes F. Stephens, "*The New York Times*' conformity to AAPOR standards of disclosure for the reporting of public opinion polls," *Journalism and Mass Communication Quarterly* 78, no. 3 (autumn 2001): 484-502.

14. Annie Lang, Jennifer Borse, Kevin Wise and Prabu David, "Captured by the World Wide Web: Orienting to structural and content features of computer presented information," *Communication Research* 29, no. 3 (June 2002): 215-245.

15. Lang, "The limited capacity model of mediated message processing," 59.

16. Esther Thorson and Annie Lang, "The effects of television videographics and lecture familiarity on adult cardiac orienting responses and memory," *Communication Research* 19, no. 3 (June 1992): 346-369.

17. Since the theory for the allocation of information processing is based on the characteristics of an ongoing flow of information, time and reading pace needed to be controlled, at least to a feasible degree. Thus, the participants were asked not to "jump" within the text. Instead they were asked to read the text from the beginning to the end, as if they were reading it at home.

18. Ronald Bernstein, "Election 2006, Senate rule for Dems? According to polls, Democrats could take over majority position if they can win at least 2 races in 3 red states," *Los Angeles Times*, Oct. 25, 2006.

19. "How would you rate the reliability of the information presented for ….?" "How would you rate the difficulty of the text in terms of reading?" "How would you rate the credibility of the text?", "How would you rate the informativeness of the story?"

20. Cronbach's Alpha is a measure of reliability of a psychometric instrument. It helps to determine whether several variables can be regrouped to a "latent" variable (LV) or if several variables apply to the same group of respondents. Therefore the selected variables must be additive.

21. The means for the other given scales were: reliability 4.793 (on a seven-point scale, with 1 being "does not seem reliable" and 7 being "seems very reliable"), difficulty 4.022 (on a seven-point scale rating from 1: "Easy. No problem" to 7 "Confusing. Needs revision."), credibility 4.809 (on a seven-point scale with 1: "Does not seem credible" and 7: "Seems very credible") and informativeness 4.847 (on a seven-point scale with one "Not very informative" and seven "Very informative").

22. Meyer and Jurgensen, "Beating disclosure to death: A rejoinder to Rollberg, Sanders and Buffalo," 5.

23. G. Cleveland Wilhoit and David H. Weaver, *Newsroom Guide to Polls and Surveys*, Reprint (Bloomington and Indianapolis, IN.: Indiana University Press, 1990), 70.

24. Walter Rudolf, "Wähler-Umfragen, Wähler-Nachfragen. Demoskopie bei Wahlen" (Electorate Polls, Electorate Inquiries. Opinion Polling at Elections), *Frankfurter Allgemeine Zeitung*, March 5, 1983, p. 10.