

This article was downloaded by:[New York University]
On: 30 June 2008
Access Details: [subscription number 762310476]
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journalism Studies

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713393939>

TRACKING THE ONLINE AUDIENCE

Phil MacGregor

Online Publication Date: 01 April 2007

To cite this Article: MacGregor, Phil (2007) 'TRACKING THE ONLINE AUDIENCE',
Journalism Studies, 8:2, 280 — 298

To link to this article: DOI: 10.1080/14616700601148879
URL: <http://dx.doi.org/10.1080/14616700601148879>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

TRACKING THE ONLINE AUDIENCE

Metric data start a subtle revolution

Phil MacGregor

This research examines the way online journalists react to new ways of knowing about their audiences. Interviewing 19 online journalists in print, broadcast and net-native media, the research analyses emergent narratives on journalists' use of tracking data from website servers. The study concludes that competing attitudes are at work in the way journalists react to server data about users and audiences. News and brand values influence journalists towards traditional behaviours, which are shown to form a strong counterweight to market pressure for growth in audience numbers. In addition, a vigorous narrative of change is evident. Online journalists are seen to claim a more reasoned and evidence-based method in choosing what to publish, which is less subject to "instinct" and whim. Tracking data are therefore directly revising the way "news values" are implemented in the respondent sample. Overall, social and organisational context rather than technology alone shape the way these online professionals react to their new tool.

KEYWORDS audiences; interactive; journalist; multi-media; news values; online

Introduction

This paper empirically investigates how online journalists say they react to direct knowledge of their audiences. It covers the use of a recent technology in journalism—the computer server—which potentially revolutionises the relationship with those for whom journalists publish. This device for the first time in journalism history allows journalists to observe, almost directly, the audience, as they access website content. Through electronic measurements, the journalists can see geographical regions from where editorial content is viewed, when it is viewed, and in what numbers (hits). Before the Internet era, journalists encountered a gap in this type of knowledge of their audience. The extent to which this gap is filled online, and how the process is understood by journalists, is the subject of this study.

Tracking the Audience: Literature

This research intersects three main themes of published literature. First are the studies of journalists' attitudes to their audiences. Second, some elements of market pressure on journalism in the Web era are briefly outlined, and the argument that tracking data might act to stimulate desire to maximise media growth; lastly, news values operated by working journalists are debated—those gatekeeping practices by which journalists choose content to publish. Other prolifically researched online journalism topics include interactivity (Massey and Levy, 1999; Schultz, 1999; Sundar, 2000), the extent of the novelty of the new media (Pogash, 1996), audience empowerment and journalists adaptations (Halls, 2001), problems of definitions (Deuze, 2004), self-perception of journalists and

Journalism Studies, Vol. 8, No 2, 2007

professionalism (Deuze, 2004; Quinn and Trench, 2002). Tools and technologies are integral to many of these.

How Journalists Model the Audience

The literature covering the combination of journalist, audience and technology is modest.

That technology links audience and journalist into a dynamic harmony is absent in early writings. Some scholars propose that technology's effect on journalism quality is regressive not progressive (Katz, 1992). Rarely are writers and academic studies of the pre-Internet era accessed in detail or quoted in new media research. Typically, journalists are labelled as displaying "traditional" attitudes (Quinn and Trench, 2002) without further elaboration.

Two studies from the pre-Internet era focus on three things: the vague notion of the audience that journalists have; their desire for autonomy and independence from the audience, and the lack of useful technology to test their hypotheses about the audience. The last of these is crucial to this study. Schlesinger (1987) and Gans (1980) analysed journalists in case studies of newsroom practices. Journalists in both accounts are convinced the audience cannot know what it wants. Schlesinger, like Gans, remarks on a lack of technology—and he noted that audience research for the BBC at the time of his study was both "sporadic" and "ambiguous". His overall judgement is, "There is no satisfactory method of feedback which enables them [journalists] to become more aware" (1987, p. 133).

The relationship between editorial and its audience is characterised by Schlesinger, in the absence of audience contact, as "the missing link". No direct connection is sought or established with audiences; instead "news values" and professional practices orientate journalists towards an idea of their public. Tunstall barely recognised the audience as a factor in journalists' minds, partly because he says they are so "source-oriented" (1971, p. 41). For Splichal and Sparks, analysing student journalists worldwide, the characteristics of the emerging professional are that: "To the extent that they do exhibit uniform traits these tend to stress a desire for the independence and autonomy of journalism" (1994, p. 179).

In pre-Internet accounts, journalists are thus seen to prefer to remain aloof from the audience and from most market research about it. The observed tendency of journalists and editors instead to turn to each other as substitute audiences in the absence of a technique to communicate with the real one is reinforced by Sumpter (2000).

Two contrasting threads emerge from scholarship about the online audience relationship with journalists. The conservative themes dominate, projecting practices that are consistent with the findings of pre-Internet studies. But a more progressive strand is also apparent—emphasising parity with audiences—though often expressed more as aspiration than fact.

Stressing that institutional circumstances shape journalism, Dahlgren notes that things should change in cyberspace: "The advent of cyberspace will inevitably impact on the factors which shape how journalism gets done—and may well colour how we define what journalism is" (1996, p. 187). His prediction seems validated in the conclusion of Boczkowski (2004a) who decided moving print to non-print delivery was more than mere technical change, but involved "fundamental cultural transformation" to the staff involved.

Boczowski (2004b, p. 205) examined the news producers' representations of the audience and found the vision of journalists about their users variously shapes the way they adopt technology. The flexible relationship found between journalism and technology echoes the current research.

Market Pressure on Journalism

Website user numbers—which are revealed by tracking data—to an extent mirror the job that is performed in traditional media by circulation figures, and viewer and listener numbers. These have always been of strategic interest to publishers. Direct connections were made early in the development of online journalism to the potential of server data to supplement circulation figures.

Audit Bureau of Circulation figures perform a structural role in helping to find out if the Press's institutional goals are being met. It might be predicted that server tracking data on numbers (how many hits), chronology, (when hits happen), and geographical information (location of origin of hits), will also be understood in terms of market goals—to “serve” customers and expand markets. This kind of expectation was investigated by Cohen (2002), who predicted the tension between market goals and new technology to become more apparent online. McManus (1994), whose work is a springboard for Cohen (2002), also credits journalists with robust attitudes to ward off compromising market pressures since they prefer “not to negotiate the terms of the news environment” (p. 4). These “terms” could be construed in large part as news values, and McManus's comment is strongly supported in this paper.

The desire for journalists to connect to audiences more directly through tracking data are rendered likely, given the news workers' statements already noted by scholars researching journalism, and the potentials of tracking data. Deuze concluded from a study that “69% of these new media professionals agreed to the proposition that a strong interactive relationship with the audience is an essential building block for any news site” (2001b, p. 362). Johnson and Kelly (2003) likewise found that a majority of online respondents viewed understanding the audience as very important, alongside market considerations and providing an alternative to other media.

Also, much practical material has been written, though not necessarily implemented, about “target markets”, and the aspiration towards closer audience connections online. A BBC trainer Jonathon Halls echoed a common appeal when he exhorted journalists this way:

In the good “ol” days you might have been able to get away without being conscious of your audience but in the new media world where we talk about interactivity, personalisation and niche audiences, knowing your audience is now more important than ever. (Halls, 2001)

However, these affirmations contrast strongly with the top-down, unequal, flow from journalist to audience many writers say still typifies online journalism as a communication practice (Heinonen, 1999), a finding of continuity with the past reinforced by O'Sullivan (2005), and which is encapsulated by Quinn and Trench: “Mass media have traditionally relied on their own judgment of what stories are worth telling, on a very largely one-way mode of communication and on an *internalised image* of their publics”

(2002, p. 5). They add the traditional “mental image” audience model, in contrast to an evidence-based model, persists online.

Quinn and Trench (2002) who conducted a pan-European study of online newsroom relationships with the audience, echo Gans and Schlesinger, when they conclude dealings with users are still “hierarchical”. They suggest continuities far outweigh changes in editorial attitudes. Another conclusion significant to this study is that the journalists, they say, show “*markedly weak interest*” in tracking usage of their sites, or employing the results to remodel their practices, or site design.

Quinn and Trench add: “One of the general features emerging from the case study interviews was the lack of detailed information held by the respondents on user demand and user profile” (2002, p. 35).

Deuze (2001b) describes typologies of online journalism, looking for “added value” in the media. However, his formalities sometimes miss the substance of differences. He notably fails, for example, to mention speed of delivery as a prime asset of journalism online which Paulussen (2004), by contrast, shows is rated by journalists as the dominant virtue of the Web.

News Values, Vague Values

News values, a phrase with multiple meanings, have been variously described by academics and professional journalists (Fowler, 1991; Hall, 1978; Westerstahl and Johansson, 1994). News values overlap gatekeeping concepts. It is not possible to rehearse the cultural and organisational factors around this vast debate, but news values’ vagueness and opacity is a common theme of several writers. Tunstall is typical:

News organisation executives and specialist correspondents all operate within what they regard as some set of news values which are in general fairly vague, but which in the individual journalist’s specific work situation are more clear-cut . . . These news values constitute only general guidelines; there is always some discretion. Just how much discretion (autonomy) exists in any particular case is unclear because what constitutes the guidelines is also unclear. Uncertainty is inherent in both the news values and the degree of discretion in relation to news values. (1971, p. 263)

Online, Johnson and Kelly (2003) claimed that news workers held “news values” in general “less dearly” than their offline counterparts, though the writers’ definitions of news values are imposed by the research, and very wide. Morris (2003) found one of four types of journalist interviewed, objective egalitarians, wanted audience feedback in order to determine the nature of news. Purdey (2000) problematised news values further, claiming the consensus on what they are is breaking down.

On the other hand, Outing’s journalistic survey (2005) of editorial processes online and server tracking data, mostly from examples in the United States, found that traditionally understood news values were controlling interpretation of tracking data.

After conducting interviews for this article, I am somewhat assured that editors’ judgement remains tantamount—at least for now—and that most online news editors are treating site statistics with caution (Outing, 2005).

Tracking the Audience: Research Method

The study analyses several overlapping themes for journalists across the sample of informative and news media. The topics are:

1. The extent of editorial take-up among British-based online journalists of data from server technology.
2. The perceived utility and limits of the server data to these journalists.
3. Perceived relationships between server data and pre-existing social and organisational practices of news production.
4. Cause and effect relationships, in either direction, between the server tracking data, and perceived news production practices and ideals.
5. Apparent dilemmas posed by the data and how online journalists sampled say they address them.
6. If and how server data affects the purposes of online journalism; and writers' reaction to perceived market pressures.

Tunstall (1971) proposes that journalism is split into sectors, summarised as print, broadcast, magazine, with each being subdivided. His version is not universally accepted and greater complexity was introduced by Splichal and Sparks (1994) and for Online, by Deuze (2004). As a general guide, but recognising that net native online websites are an additional category, Tunstall's divisions have been followed here.

The starting point is the online journalist, who is either part of, or independent of, any parent medium. Many published studies cover one medium only, mostly newspapers, or combine news media with other types of information source (Boczkowski, 2004a; Light, 2003; Massey and Levy, 1999; Schultz, 1999) rather than take the online journalist as a generic type. This research specifically crosses the print–broadcast divide, and includes net-native and magazine journalists.

Research into online journalism has predominantly used surveys and questionnaires for gathering data. More often than not, detailed reasoning for journalists' actions or opinions are not available in previous research, while questionnaires tend to frame concepts on behalf of respondents (Quinn and Trench, 2002; Singer, 1997). In these methods, complexities of response are potentially reduced. Also, those who answer online surveys are often self-selecting.

The research uses qualitative one-to-one interviews. Nineteen journalists in senior and mid levels of editorial hierarchies participated in semi-structured one-hour interviews conducted over 18 months face to face, or by telephone. Participants qualified only if they were working now or recently in online journalism. No adjustments were made for the time elapse. About 160,000 words were gathered, transcribed verbatim, and put into searchable format.

Sampling was aimed to gather information-rich participants. Choice of medium and personnel aimed to capture "progressive" practice.¹ Maximum variation sampling was used for choice of medium, the variations being in ideological standpoint, type of medium, and size of staff and audience. Given the several methods of defining journalism online by formal characteristics (Deuze, 2004; Deuze and Dimoudi, 2002; Freedman, 2003), selection decisions are bound to seem arbitrary. By rule of thumb, an attempt was made to collect mainstream journalistic opinion from the larger commercial or public service English-language sites, balanced by smaller, often "alternative", "activist" or

“agenda-driven” sites. As stated, the sample was intended to be media-inclusive (print, broadcast) and net-native (Shianews, OneWorld, Blink, AOL). A cross-section of funding types in the sample prevents a picture forming that might be skewed by a particular financing system. A degree of expediency occurred. Some journalists previously known to the researcher helped recommend participants, or became such themselves. Some participants recommended others (snowballing; Lindlof, 1995).

The sample is not representative of all online journalism in any formal sense. Generalisation applies only to the group sampled; with the caveat that many sites chosen are large (circulation, resources, staff numbers) and reputable. Questions to participants covered job description, demographic knowledge of the audience and how it was obtained. Journalists were asked if they knew the age, location, gender, economic group, or any other characteristic of individuals in the audience. They were asked about server data, time and frequency of accessing it, how it is provided and by whom, and their perception of its use, either at the time, or in framing later editorial policy. They were asked if there was more data they wanted. Upcoming changes in technology formed the final section. The focus of the investigation was not declared to participants.

Journalists interviewed were mostly at very senior levels, such as editors in charge of teams. Senior journalists were preferred, since they are more likely than subordinates to be making gatekeeping decisions, more likely to be aware of the range of activities at their site, and more likely to know of general analysis and strategy. In more than half the cases the interviewee was the most senior in the online section. Each participant is evaluated independently. More than one journalist was interviewed from the larger organisations (*Financial Times*, BBC, CNN, AOL, EMAP). The individuals are not identified.

Tracking software became an object of special scrutiny in the analysis. The research looked for common and divergent themes among participants, in questions of perceived fact or opinion. Themes that evolved during the analysis were organised by research questions. Extent, speed of response and reasons for response to tracking data became primary objects of focus, coupled to how their data were interpreted. Software search terms helped aggregate responses on specific points. Where possible, responses and themes were classified using terms offered by participants. The term “interactivity”, for example, which frames some new media research and is said to be a defining characteristic of online journalism (Deuze, 1999; Newhagen and Rafaeli, 1996; Singer, 1997) is not used in analysis partly because the term was only offered six times across all interviews.

Weight is not exactly proportionate to numbers of journalists giving a view, as significance is also given to responses that succinctly frame dilemmas. Past practices are accessed for comparison from descriptive literature, but mainly from the interviewees’ self-perceptions, and also by referring to the author’s experience of 15 years in UK print editorial departments.

The questions allowed specific assessment of responses to server tracking data. These were graded by degree, by time of response, and type of response. At one extreme was stimulus/response—where the editorial action was direct and total in response to tracking data—such as instant removal of a story from the website. On the other extreme was nil response, at the time or later. Between the extremes lies a continuum, but with possibilities of either (1) partial, immediate response to data; or (2) partial, delayed response.

The causal chains behind these response categories were observed where participants indicate them—falling crudely between the two alternatives of either

technological cause, or social and organisational causes, such as reliance on “news values” or “brand values”. “Brand values” are defined equivalently to news values as the mental routines affecting the inclusion and exclusion of content. They are the gatekeeping processes, applied to magazines. There is no term coined for them in literature on magazines in the way “news values” cover news. So brand values here are used as equivalent to news values, but as applied to magazine writers on and offline.

The terminologies—hits, page impressions, unique users, page views—differ slightly from media to media, as does what they signify. It has been accepted nevertheless that there is a rough identity to the names and corresponding realities between them. Similarly, the terms “tracking data”, “tracking software”, “server data”, “log statistics”, “log data” or “site metrics” are often interchangeable and are taken to have identical meanings in this discussion.

Tracking the Audience—Findings

How Server Data Are Used

Server logfile data is only one of the communication devices available to online journalists—but it is one of the few non-human, semi-automated ones. It does not require audience intervention because whether the audience realises it or not, their activity is logged.

The journalist participants reveal that prolific use of server logfile data (tracking data) has arisen and developed in almost all publishing houses and editorial organisations consulted. A variety of uses, and degrees of use, are revealed, and journalists demonstrate an intelligent, sometimes critical interest, in the information provided. They make extensive interpretations.

In large media organisations, the journalists obtain the tracking data through a third party, from hired companies such as Omniture, which analyse data from server log files and make results available to the editorial department. It is often then accessed directly on newsroom computers. In some cases journalists report the information has passed through the marketing departments already. Editorial offices, and news managers in larger newsrooms, consult the data to see the numbers of unique users on the site at any one time, or the numbers of hits at one time, or the total number of page impressions, or all of these. Journalists interviewed see items—like individual stories—listed in order of popularity, and with actual numbers as well as percentages of users.

Journalists regard this data as one strong indication of how the item—be it a single story, or Web page, or a site section—has performed or is performing with the audience. In an even more detailed way it may tell the journalists which part of a story is most accessed. Statistics can reveal to all the journalists distilled summaries, such as the total numbers of hits on the whole site as a snapshot at one time, or in sequence over chosen periods—a day, a month, a year.

The rates of computation and availability of results varies from organisation to organisation. In addition, user access is usually broken down into statistics referring to each country or region of access. Tracking data can reveal the geographical origin of hits, and also the exact time that the hits occurred, and the length of time a user stayed or stays on a story. This is usually learned after it happens, not as it happens.

Editors and journalists can learn from this exactly when surges occurred in site traffic, and how surges are or were distributed across the world. As fodder for the editorial

calculus, whether to publish and how much, such knowledge has implications. At *The Guardian* for example, about 40 per cent of the online readers are seen from logfile data to come from the United States.

The precise timing of hits helps editorial staff assess users' habits. At CNN and *The Guardian* newspaper, journalists commented on the "wake-up" surge of traffic from America. At the *Financial Times* the sole contributors from the Antarctic and Vatican City who voted in a popularity poll could be picked out individually. When spikes in user numbers occur at local lunchtimes around the globe, inferences are made that office workers are the primary site users. Inferences about lifestyles vary from site to site but are made with some confidence—a teen magazine editor reveals the rival attractions between which her site is accessed:

Our first peak during the day is about 4.30-ish. So they've obviously just got home from school and they log on to the computer. Then after 7 o'clock—after they've had their dinner, and they've watched the soaps—then they log on again. So that the evening time, before they go to bed, that's busy. (EMAP)

At a regional BBC site the midday peak is seen to signify office workers, whereas "in the evening peak it is the silver surfers, or the people who are going on to flights or whatever, and they have a computer at home".

Tracking data also illustrate editorial limitations in relation to geography—such as low penetration of a site in a country or region. OneWorld, with its mission for cross-communication between developing countries and non-governmental organisations worldwide, learns from server tracking data that only developed nations create the traffic. It is not the message they would prefer, but they have to believe it.

Significantly, almost all site journalists experienced some delay in obtaining numbers of hits, page impressions and any breakdown of tracking data. The typical time lapse was 24 hours—a day or so after a piece of content was first published. At that time its popularity can be seen but there was only a mixed desire to get access to the statistics any sooner. A BBC respondent said: "One day [delay on story data] is perfect because the story is still fresh and it might still be up" (BBC). However, at *The Guardian* a new tracking system (implemented summer 2005) that gives almost simultaneous vision of tracking data is clearly desired and welcomed.

Decisions and Revisions from Site Metrics

Decisions to alter practice directly, totally and immediately under the stimulation of tracking data are exceptional. Normally, an individual content item will not be re-edited, re-published, moved, archived or retained any differently. This is sometimes because 24 hours is longer than the normal publishing lifespan of the material, so by the time statistics are seen, the moment for intervention has passed. Some evidence suggests that if the data are obtained in real time, journalists react more directly to its stimulus:

We do have a real-time statistics page that gives me a percentage—you can see where people are clicking, with a ten-minute delay. That's quite useful, because when I come in on shift I would be tempted to take down the top one [story], but if it is rating very well, I will keep it. (CNN)

However, as yet few sites have the ability to get this information and, as shown above, they are not pressing for simultaneous knowledge of user activity via server data.

But CNN and AOL do have in-house software that measures the percentage of clicks on any one part of the site in real time, and this reveals “whether a story is going hot or cold”.

If few examples were admitted of immediate reaction to server data to change what happens to a content offering, there is much stronger evidence of partial, incomplete, but immediate response to data about a story. Decisions are often taken to deepen and extend a story or to abandon further development (*Financial Times*, AOL). Overall, one editor said Web working had led news to be shaped “in a completely different way” (*Financial Times*). And she explained:

I mean you have to make a judgement on the fly, and if we see that we did have the right judgement then we know we’re on the right lines for expanding that coverage or analysis the next day, that sort of thing. (*Financial Times*)

The third option is that response is partial and delayed. In fact this type of behaviour is probably the most common and deepest in its impact on story choices. In this case, stories are often grouped together as types and watched over time. Instances are assembled into patterns, not to change what happens to that one story, but to change the fate of a type of story in the future publishing decisions.

Overall the responses range from stimulus/response behaviour—direct, immediate, action taken in reaction to tracking data—to deliberated, mediated responses in which the statistical picture is weighed against many other factors. We shall see some of these mentioned in the reaction to trends, below.

Usefulness of Tracking Data—Perceptions

So far some practices have been outlined in terms of the speed and extent of reaction. How are these practices perceived? These will be divided into positive assessments, and negative or neutral assessments.

Positives of efficiency. Positive gains perceived to editorial efficiency feature highly in these journalists’ responses to tracking data. These were: (1) That they can see most popularly accessed stories; (2) that they can assess trends over time across the site; and (3) they regard tracking data as supplying “objectivity” compared to other interactive human feedback.

There is a strong theme of exploitation of information about the popularity of stories. In several departments, the numbers from server log files are broken down into a “top 10” or “top 100”. In the two UK national newspaper offices consulted, and one regional BBC office, such lists are avidly observed. The *Financial Times* uses it to assess and weight the relative strengths of its current or immediately past website content.

At the bottom we have got the top 10 most read articles (on FT.com), which we tell all the editors [offline] about—because obviously they’re interested as far as their page is just to see which ones [stories] really flew. And sometimes it’s quite surprising . . . So every day the editor (of the *Financial Times*) sees that top 10. (*Financial Times*)

This passage indicates shows how highly the material is valued in terms of traditional editorial aims and procedures. Similarly, at AOL, two editors agree:

[Hit rates are accessed] constantly. I'd say we do it the day after a big story and then, well probably every day, and possibly during a story as the story's breaking and we'll check how well something's doing. (AOL)

And, again, if something does really, really well then we'll hammer it and keep using it. (AOL)

The same point is emphasised at *The Guardian*:

Guardian Unlimited are much more aware of what goes down well and what doesn't because of that [server tracking data]—because every site editor watches the statistics *pretty obsessively*. It's the one piece of definite measurable feedback they get in their work. So they obviously keep a hawk eye on what's coming back.

And the quality checks are spread further down the editorial hierarchy to create a "staggeringly different" environment:

... because previously, well newspaper demographic information would go as far as the editor and sales information ... Online you're getting a much more direct verdict on the quality of a given story. (*The Guardian*)

In short, the statistics were described as "fantastically useful" (BBC). These extensive quotes display widespread high value placed on the statistics and extensive use of them at the core of the editing process. The site data were often rated as the best way journalists could know their audience—superior to email or discussion boards or other human interactive systems. Stories can be tracked instantly in terms of how many readers there are, or whether the hits were continuing to come or "drying up" (*Financial Times*).

The above reactions imply some fairly immediate reactions and assessments of site content. But very often there is a delayed and not an immediate adaptation of practices—if any—in response to data. "Obsessive" observation with "a hawk eye" does not always lead to instant real-time reaction. Comments from several participants indicate that, rather than provoking instant reaction, statistics on the Web are best used as indicators of trends. The tendencies are watched, the patterns assembled and assessed. A magazine's electronic media editor said:

If we get something a bit spectacular come up, we then start looking. Because it's much easier to try and pick up a trend. I think Web work is about trends not about individual spikes. So I think if we could start to pick up a trend that, whenever we press a particular button this happens, then we'd definitely be circulating that as good practice. (IPC Media)

Trends have created some delayed adaptations to practices in the way certain types of story evaluated. A regional BBC site editor observed that since regular crime stories fare less well than expected online as shown by statistics, their importance has been downgraded:

I think we are beginning to learn what does well and it is quite surprising . . . but I think we may be starting to target stories that we didn't really consider before and we don't emphasise other stories, such as local crime. (BBC)

Other examples of delayed changes to content categories after observing trends include: emphasising content about technology, upgrading picture galleries in magazine sites, and prioritising customer product reviews, recipes and dating services. Priorities have been altered to allow these category offerings more length and prominence.

Even positive editorial lessons from data are, therefore, drawn cautiously. The implications of a particular peak or trough in statistics are sometimes seen as unreliable. In one case (IPC Media), a statistical spike was identified as a rogue result of changes in technology. In other examples, traffic surges occur because of the accidental factor of links to an item of content from other websites. A further reason respondents gave to mistrust data is that the number of clicks on a page may actually reflect its position in the design of the site, rather than its readership appeal:

No it wasn't [useful to see click rates]. I think it's of very limited use because I think people just sort of click on a headline and a lot of the stories were just ones that were on the home page, and it didn't necessarily . . . if you put things prominently, people click on it, but that doesn't mean they read it. (*Financial Times*)

Server data are, of course, only one channel of contact between journalist and audience. A wider picture is given by email, bulletin boards, forums, chat forums and other "interactive" aspects. When the results of the Michael Jackson trial were announced at midday (summer 2005), BBC interactive did not have to wait till the following morning for server statistics to appreciate the massive popularity of the story. Emails, postings and professional instinct were enough.

One final positive side to Web statistics was seen as its objective character compared with emails, chats and bulletin boards. The benefit of statistics is that they can be more representative and tell a more accurate story than live human feedback, which forms most other interactive feedback. Thus extreme or vociferous opinion that occurs in forums is not to be taken as typical of the whole user base of a website; objective Web trends can help reinforce this editorial balance. At AOL much feedback at one point during the General Election of 2005 suggested right-wing conservatives dominated the website—but it did not even take server data to appreciate the non-representative nature of the feedback.

Negatives—narratives of rejection. Editorial reactions to metric data so far outlined work broadly, even extensively, to augment operationally useful knowledge. However, there is another narrative within the journalists' responses. This goes beyond mistrust. It amounts at times to outright dismissal of the messages, when journalists actively resist giving tracking data too much importance.

Reservations are formulated in several ways. The main one is the need to adhere to brand values or news values. These can conflict or cut across messages received from site data. They form the intellectual matrix within which the metric data is evaluated and sometimes found wanting, especially when the data indicates popularity of a content offering. A second reservation stems from shortcomings of data, in that it comes as "cold statistics", rather than as living thought and feeling. The third is that reactions to the data are just that—reactive, and give only indirect messages on how to attract traffic to new

content yet to be published. Statistics do not indicate what to do next. This omitted dimension is sometimes termed “leadership” by the journalists.

And fourthly, it is too laborious retrieving it.

When negative reactions do occur, they are due to a conflict with existing values that determine which content ought to be chosen to publish. Examples are given of a type of response shown by several of the journalists:

You could easily tone down your offering to the most popular, but that is not the point . . . You might find out there are certain things the audience responds to—you get lots of hits on a story—but that is not what the site is about. (BBC)

And again:

Clearly we want to still have editorial control, because we have to stick within the BBC’s producer guidelines and all that. So if everybody was saying “do X” and we felt it wasn’t quite the right thing to do, we’d have to think very carefully about it.

In the case below, the reservations to direct response stem from the right of the journalist to choose which story to publish:

You have to be responsive to what people want to read but . . . the whole essence of news values is that you make choices based on what you think is the right story [for the *Financial Times*]. (*Financial Times*)

The same sentiment is voiced emphatically at CNN:

Obviously we are not going to do a story if nobody is going to read it but we are not led purely by that—to be perfectly frank if I just wanted to chase what people on the Internet wanted to click on, I would do stories about soft porn and football and nothing else. We are a news site so we have to be treated as news and we have to cover stories which do not always have mass appeal. (CNN)

And at a net-native site:

We are a bit driven by news values and a bit driven by things we just think are important, and just need to get out there. (Black Information Link)

The strength and ubiquity of this response demonstrates its primary position in balancing news and brand values against direct reaction to tracking data. Among the responses is a repeatedly stated wariness of chasing popularity.

We certainly don’t want to have an overall reputation of being populist because . . . the image that we want to cultivate is of not being populist, but actually being positive about black people in the UK, and sometimes taking stances that are different from the mainstream. (Black Information Link)

Sheer adherence to chasing user numbers or numbers of clicks is criticised also for loss of editorial independence, or “slavery” to the audience. The notion of following clicks is so crude as to be “dangerous” (*Financial Times*).

As to the second reservation—shortcoming of data—negative or neutral feelings were registered in several ways—the time it takes to obtain data (though this is sometimes acceptable), but especially its limited value in giving a living profile of the user. Site statistics are described as cold, hard, “quantitative rather than qualitative”, and therefore inadequate as a means to inform journalists about the user. It is “frustrating that all you have is numbers” (BBC).

And again:

Well it [data] doesn't give you a profile of the individual. It doesn't tell you anything about the individuals . . . We are more interested in how we're changing opinions. So you'll never get that out of a computer. (OneWorld)

In this respect server data are derisively classed with information obtained by company marketing departments, seen as brutally geared towards selling:

I mean I think there's a bit of resistance in most editorial departments—I guess in most media organisations—to wanting to join in the demographic analysis that the marketing side go through (and the advertising side), because in a way it seems kind of constraining. (*Financial Times*)

A third reason for turning away from messages in tracking data stems from the perceived need to show leadership, and anticipate user wishes, not simply to react. News values are not a simple formula. Respondents point out that particular decisions still need to be made fast, “on the fly” and with many goals in mind.

Although Web statistics help identify trends, and let editors tailor news to the audience, or identify the stories that need development, they do not tell anybody what should be published next. “I wouldn't just slavishly go for what people are interested in, because *people don't always know what they are interested in* until they are presented with the information” (*Financial Times*).

News is a question of “anticipation”. One respondent specifically identified the danger that market information about the user is confused with a knowledge of what they want.

Fourthly, there is the time factor, the sheer burden of editorial tasks limiting its use, attested to by almost all respondents.

Tensions and Shifting Values

However, there are responses that are neither negative nor positive, but whose importance lies in that they clearly imply evolution in practices, new tensions and dilemmas resulting from the messages given by tracking data. Of course, revolutions in news practice happen slowly. We have seen some positive uses provided by tracking data, and some negative or neutral responses, resulting in inaction, or deliberated, cautious, integration of its apparent messages into editorial practices. In the sense of covering familiar ground, these responses are mostly conscious and unproblematic. The new ones, on the other hand, sometimes reveal unresolved inconsistencies.

For example, tracking data sometimes create dilemmas not simply answered by appeal to “news” or “brand” values. At Britain's most accessed news website, *The Guardian*, conflicting pressures arise between the values attached to their niche audience

of the newspaper, and a new mass market that has developed online. Forty per cent of the website is US traffic. Its editors are uncertain whether to follow the numbers—a mass market online—or whether to retain the distinctive minority attitudes of the print newspaper. The reaction is gauged thus:

The information's [tracking data] there if we ask for it. The questions lie around should we ask for it? And if we do, how should we act on it? To what extent should we allow our decisions to be influenced by that information? Because it is also the question about to what extent are we a product looking for as big as possible a market versus having a set of values which we accept appeal to only a small, or smaller, part of the population. (*The Guardian*)

Niche audience is not the only inhibitor on chasing popularity. Another quality is creeping into editors' thinking. The following quote shows that it is the quality of interaction—not the quantity of clicks—that is helping decide what content is carried. Content quality is being challenged by desire for quality of audience response:

Online news is changing. I think it's not just about how well a story will do . . . there are lots of different measurements of it and it keeps changing. Sometimes it's how many clicks it does. Sometimes it's how many people left messages, and sometimes it's how many people left a vote in a poll. (AOL)

Editorial objectives are therefore more diverse than just those of obtaining volume. What type of medium online the particular item was accessed through—radio, video, or words or graphics—was also valued. It was important:

That we *engaged* the member rather than just informed them. It's all about interactivity. Or, increasingly, how many people watched the video clip attached to that story. Because now, in the back of their minds, they'll know that, when they come to our service next time, we'll have a piece of a video there that they can actually watch it. (AOL)

To engage is a new value, but not really a news value. This focus on the quality of interaction heralds change.

Even more emphatically, there are signs in several participants of an erosion of the superior view journalists traditionally have of their independent ability to spot news and the value of news. Traditional journalists affirm that news choice and "instinct" is their preserve. Participants cite the approach in some cases with little qualification. "I have a pretty good idea of what kind of a story is going to be popular on the net in any one place and you just get a feel for it, and an instinct" (CNN).

But there is a new tendency. Increasingly the independence of "instinct" is being undermined with a cross-current resulting from the use of data. This change amounts to a significant evolution in an aspect of news professionalism.

In the old days we'd say: "That's a good story" . . . and I would say that's the primary means still of picking a story [online]. So we usually follow our instincts to a great extent. *But I think we'll see in the near future more reliance on knowledge based on what metrics are throwing up.* (*The Guardian*)

Humility and a recoil from too much self-certainty is evident here: "There is a slight sense I get from on-line journalists that they're moving less, you know, they're less reliant on instincts and less willing to trust them" (OneWorld).

Much the same balance is put like this:

So as much as our instincts as news editors can tell us what relevance or importance, or how good a story will be for our readers (those are just our natural journalistic instincts), *we actually look at the figures [tracking data] as well to see whether or not that actually bears out in reality.* I mean most often I'm glad to say it does. (*Financial Times*)

Results indicate a new self-consciousness in this area of news practice—though not yet a displacement of traditional ways.

Tracking the Audience: Conclusions

In respect of the first research consideration, the extent of its adoption, the server tracking data is widely consulted by this set of participants. It is sometimes viewed with great intensity—"obsessively"—to deliver a variety of editorially significant messages. This results in delayed, but also, in instant adaptations to the content offered online. In so doing editorial priorities are re-weighted and in consequence a basic practice of news judgement is being affected in the online newsroom.

As to the second question, tracking data is perceived as useful in all sectors of media consulted in this study—print, broadcast and net-native. Its wide adoption comes in tandem with other interactive systems such as email or discussion boards, which it sometimes complements, but also crucially qualifies by presenting an accurate statistical balance. It therefore establishes a valued independent perspective on human interactive feedback received by these journalists. It validates and informs editorial practice as has never been possible before, and fills in Schlesinger's "missing link"—helping to shape decisions of both editorial detail and strategy online. Sometimes its lessons affect offline news judgements. These observations cast considerable doubt on the finding of Quinn and Trench (2002), who noted "remarkably weak interest" in editorial use of tracking data.

Limits are placed on use of server technology firstly by the existing mental and professional patterns of journalists (their news and brand values), secondly, by time shortages, and thirdly, by what is regarded as the incomplete "efficiency" of the system producing the data, for example its delayed availability, or incompleteness. The data appear not to present a picture that radically challenges past experience.

In the first two research questions above, therefore, social shaping of the use of tracking data pervades both adoption and limitations, echoing Boczkowski (2004b). A variety of factors are entailed in the shaping processes. Technology appears from the evidence to have had little direct agency towards change. Some responses show the opposite, in that it elicits vigorous restatements of traditional attitudes, particularly, for example, that any slide towards populism is dangerous. "Slavery" to audience whim is sharply defined as perilous, from exactly the same principles outlined in the pre-Internet studies of Gans and Schlesinger.

However, this narrative of continuity—that adoption is shaped by factors with recognisable antecedents in editorial practice—ignores important tendencies for change. There are significant developments evident in the answers to remaining research questions. In respect of the third: the most prominent change is to the news "instinct" which evidence shows is increasingly scrutinised by some British online journalists. Some admit now that they double-check their instinctive guesses with tracking data. They no longer implicitly trust themselves.

As a result, an apparently more empirical, and less intuitive framing of the method of how to make content choices is emerging. Journalists check their hunches against the statistical data, and sometimes adapt their actions on the basis of “hard” information. There is thus, in relation to the fourth research question, an emergent though incomplete cause and effect relationship between tracking data and some content publishing decisions. This move towards evidence-based method, rather than “gut reaction”, is partial, but is nevertheless overt and significantly enabled by the new technology and tracking mechanisms. Where Tunstall states: “Uncertainty is inherent in both the news values and the degree of discretion in relation to news values” (1971, p. 263), it can now be said that the uncertainty is perceived in this sample to be diminished by server-provided knowledge.

Further novel tensions (the fifth question) presented by tracking data are emerging between the audience demographics, on the one hand, and existing news or brand values, on the other. Both news and magazine websites’ actual audience sometimes conflicts with the type of audience that has matured offline. Conflict between specific publishing wings of a media brand—between website content and printed content—appears to present unresolved dilemmas for editorial departments—graphically presented by tracking data.

As to the sixth question of the wider aims of journalism: these are evolving online. Interactivity has entered the list of some online journalists’ desirable priorities. Several participants comment on the importance of interactivity (engagement) itself as a goal of publishing, rather than other ideals such as the desire to inform, or the need for a type of audience. Some journalists work to gain a response from the audience, and mere response—as well as its content—is a measure of success. Audience participation competes with the desire to inform, which fits the picture obtained by Deuze (2001b). For this ambition, tracking data are seen to serve a significant purpose in telling journalists which offerings attract traffic.

Although server technology measures user activity, it does not appear to mould journalists’ motives. Server technology is not provoking the wish for interactivity so much as being the measure of it. Although it is partly implicated in changing editorial ends, as well as means, server technology did not produce a stated wish to create a scenario not yet in place. Journalists’ use of this particular type of data and technology is mostly reactive, even though they welcomed, but did not initiate, improvements in it from time to time. Their motives generally can be seen as being affected by inherited social and organisational, rather than technological forces.

In sum, therefore, server data are helping:

1. To refine existing professional practices of some British-based online journalism.
2. To refine and guide emergent social practices and aims of some online newsrooms.
3. To define emerging tensions in journalism—such as those developing between online and offline audiences.

They are not:

4. Independently producing, enforcing or determining new professional procedures or beliefs.

Possibly the most significant ideological conclusion stems from comparisons of tracking data with measuring systems from the past. The hunch that journalists online might be driven by technology towards populism and the hunt for audience numbers is an ideologically riven field. The editors and journalists agree the need for numbers. They

realise that volume is essential to survival. But this sensitivity seems to enhance in an equal degree the counter response that site values must be preserved in well-articulated, often non-populist, editorial news or brand values. It is far too simple to suggest that journalism means an uncritical hunt for markets. Numbers need to attach to specific audience qualities, and this link between numbers and values is still seen as the true grail, the truly virtuous circle.

NOTE

1. The organisations consulted were: Sky Sports Online (one journalist), CNN London, and CNN Hong Kong office (one journalist in each), AOL London (two journalists), *Guardian Unlimited* (two), *Financial Times* (three), IPC Media (one), BBC Online London (two), BBC Online Southampton (one), Shianews.org (one), Blink.org (one), OneWorld (one) and EMAP (two). Participants were promised anonymity. Therefore titles of journalists consulted in each organisation cannot be given. Titles without organisation include: Editor Online, News Editor, Online Editor, Electronic Publishing Editor, Online Journalist, Supervising Editor, Channel Editor, Assistant Online Editor, Launch and Managing Editor, Travel Editor. Of these media chosen, five organisations are not primarily funded by advertising revenue (BBC; OneWorld; Blink—Black Information Link; Shianews; AOL). The *Financial Times* Online is part paid by subscription and part by advertising.

REFERENCES

- BOCZKOWSKI, PABLO (2004a) *Digitizing the News: innovation in online newspapers*, Cambridge, MA: The MIT Press.
- BOCZKOWSKI, PABLO (2004b) "The Processes of Adopting Multimedia and Interactivity in Three Online Newsrooms", *Journal of Communication* 54, pp. 197–213.
- COHEN, ELISIA (2002) "Online Journalism as Market-driven Journalism", *Journal of Broadcasting & Electronic Media* 46(4), pp. 532–48.
- DAHLGREN, PETER (1996) "Media Logic in Cyberspace: repositioning journalism and its publics", *Javnost/The Public* 3(3), pp. 59–7.
- DEUZE, MARK (1999) "Journalism and the Web: An analysis of Skills and Standards in an Online Environment", *International Communication Gazette* 61(5), pp. 373–90.
- DEUZE, MARK (2001a) "Understanding the Impact of the Internet: on new media professionalism, mindsets, and buzzwords", *EJournalist* 1(1), <http://www.ejournalism.au.com/ejournalist/deuze.pdf>
- DEUZE, MARK (2001b) "Online Journalism: modelling the first generation of news media on the World Wide Web", *First Monday* 6(10), http://firstmonday.org/issues/issue6_10/deuze/index.html
- DEUZE, MARK (2004) "What is Multimedia Journalism?", *Journalism Studies* 5(2), pp. 139–52.
- DEUZE, MARK and DIMOUDI, CHRISTINA (2002) Online Journalism in the Netherlands: towards a profile of a new profession", *Journalism* 3(1), pp. 103–8.
- FOWLER, ROGER (1991) *Language in the News: discourse and ideology in the press*, London: Routledge.
- FREEDMAN, DES (2003) "New Media: technology and cultural form", in: Ramon Salaverria and Charo Sadaba (Eds), *Towards New Media Paradigms*, 11 COST A20 International Conference Proceedings, University of Navarre, Garrasi, Avd Baranain, 52, Pamplona.

- GANS, HERBERT (1980) *Deciding What's News: a study of CBS Evening News, NBC Nightly News, "Newsweek" and "Time"*, London: Constable.
- HALL, STUART (1978) *Policing the Crisis: mugging, the state, and law and order*, London: Macmillan.
- HALLS, JONATHON (2001) "Writing for Multimedia", *European Content Summit*, <http://www.contentsummit.com/workshops/masterseries01.php#1336>, accessed 10 November 2005.
- HEINONEN, ARI (1999) *Journalism in the Age of the Net*, Tampere: Acta Universitatis Tamperensis.
- JOHNSON, THOMAS and KELLY, JAMES (2003) "Have New Media Editors Abandoned the Old Media Ideals? The journalistic values of online newspaper editors", *New Jersey Journal of Communication* 11(2), pp. 115–34.
- KATZ, ELIHU (1992) "The End of Journalism? Notes on watching the war", *Journal of Communication* 42(3), pp. 5–13.
- LIGHT, ANN (2003) "Participants, Consumers and the Politics of Online: British Information Providers on the Web Reconfigure the Audience", in: Roman Salaverria and Charo Sadaba (Eds), *Towards New Media Paradigms*, 11 COST A20 International Conference Proceedings, University of Navarre, Garrasi, Avd Baranain, 52, Pamplona, pp. 619–34.
- LINDLOF, THOMAS (1995) *Qualitative Communication Research Methods*, Thousand Oaks, CA: Sage Publications.
- MASSEY, BRIAN and LEVY, MARK (1999) "Interactivity, Online Journalism, and English-language Web Newspapers in Asia", *Journalism & Mass Communication Quarterly* 20(2), pp. 138–51.
- MCMANUS, JOHN (1994) *Market Driven Journalism*, London and Thousand Oaks, CA: Sage.
- MORRIS, JOHN (2003) "Audience Identifies Types of Reporters, News Consumers", *Newspaper Research Journal* 24(2), pp. 36–49.
- NEWHAGEN, JOHN and RAFAELI, SHEIZAF (1996) "Why Communication Researchers Should Study the Internet: a dialog. Symposium", *The Net. Journal of Computer-Mediated Communication* 1(4), <http://jcmc.indiana.edu/vol1/issue4/rafaeli.html>
- OUTING, STEVE (2005) "Sophisticated Web Stats Give Editors Better Idea of Reader Interests", *Editor and Publisher*, http://www.editorandpublisher.com/eandp/search/article_display.jsp?vnu_content_id=1000992026.
- O'SULLIVAN, JOHN (2005) "Delivering Ireland", *International Journal for Communication Studies* 67(1), pp. 45–68.
- PAULUSSEN, STEVE (2004) "Online News Production in Flanders: how Flemish online journalists perceive and explore the Internet's potential", *Journal of Computer-Mediated Communication* 9(4), <http://jcmc.indiana.edu/vol19/issue4/paulussen.html>
- POGASH, CAROL (1996) "Cyberspace Journalism", *American Journalism Review*, June, pp. 25–31.
- PURDEY, HEATHER (2000) "Radio Journalism Training and the Future of Radio News in the UK", *Journalism* 1(3), pp. 329–52.
- QUINN, GARY and TRENCH, BRIAN (2002) "Online News Media and Their Audiences", *MEDIA (Multimedia in the Digital Age)*, <http://www.ecdc.info/publications/index.php>.
- SCHLESINGER, PHILIP (1987) *Putting "Reality" Together: BBC News*, London: Methuen.
- SCHULTZ, TANJEV (1999) "Interactive Options in Online Journalism: a content analysis of 1000 US newspapers", *Journal of Computer-Mediated Communication* 5(1), <http://www.ascusc.org/jcmc/vol5/issue1/schultz.html>.
- SINGER, JANE (1997) "Changes and Consistencies: newspaper journalists contemplate online future", *Newspaper Research Journal* 18, pp. 2–18.
- SPLICHAL, SLAVKO and SPARKS, COLIN (1994) *Journalists for the 21st Century*, Norwood, NJ: Ablex.
- SUMPTER, RANDALL (2000) "Daily newspaper editors' audience construction routines: A case study", *Critical Studies in Media Communication* 17(3), pp. 335–47.

- SUNDAR, SHYAM (2000) "The Internet—multimedia effects on processing and perception of online news: a study of picture, audio, and video downloads", *Journalism Quarterly* 77(3), pp. 480–99.
- TUNSTALL, JEREMY (1971) *Journalists at Work: specialist correspondents, their news organizations, news sources, and competitor–colleagues*, London: Constable.
- WESTERSTAHL, JORGEN and JOHANSSON, FOLKE (1994) "Foreign News: news values and ideologies", *European Journal of Communication* 9, pp. 71–89.

Phil MacGregor, The Media School, Bournemouth University, Weymouth House Talbot Campus, Poole BH12 5BB, UK. E-mail: pmacgreg@bournemouth.ac.uk