What Makes Journalism “Excellent”?  
Criteria Identified by Judges in Two Leading Awards Programs

Ivor Shapiro  
Patrizia Albanese  
Leigh Doyle  
Ryerson University

Abstract: What does “excellence” mean in journalism? The literature reveals no universally agreed set of standards, and awards guidelines are often unclear. We interviewed judges in two leading Canadian print journalism awards programs, using a sequence of open-ended and ranking questions to probe their criteria of excellence in a way calculated to elicit not just the standards they felt should be applied but the standards they actually did apply. Judges mentioned a wide variety of criteria, including the social importance and impact of works of journalism. But only two values were affirmed consistently: writing style and reporting rigour.

Résumé : Comment définit-on un « excellent » journalisme ? Un parcours de la bibliographie à ce sujet révèle que les règlements relatifs aux prix décernés aux journalistes sont souvent vagues et qu’il n’y a pas de standards d’évaluation unaniment établis. Pour ce projet de recherche, nous avons interviewé des juges de deux programmes de remise des prix dans le domaine du journalisme écrit au Canada. Nous avons posé des questions ouvertes et des questions fermées dans le but de recueillir non seulement les critères qui devraient être appliqués, mais aussi ceux que les juges ont déjà appliqués dans leur pratique. Les juges ont mentionné un vaste éventail de critères, y compris l’impact social des œuvres journalistiques. Toutefois, seulement deux valeurs se sont imposées invariablement : le style d’écriture et la rigueur du reportage.

Keywords: Research methods; Publishing/print journalism: magazines; Newspapers; Ethics; Awards

Ivor Shapiro is Associate Professor, Director of Magazine Journalism: School of Journalism, Ryerson University, 350 Victoria Street, Toronto, ON M5B 2K3. He is also Editor-in-Chief: The Canadian Journalism Project (JournalismProject.ca). E-mail: ishapiro@ryerson.ca. Patrizia Albanese is Assistant Professor in the Department of Sociology, Ryerson University. E-mail: palbanes@ryerson.ca. Leigh Doyle is a recent graduate and former Research Assistant, School of Journalism, Ryerson University. E-mail: leigh.doyle@gmail.com.

©2006 Canadian Journal of Communication Corporation

425
Introduction

When Seymour Topping was asked about the concentration of winners of the Pulitzer Prizes shortly before his retirement as the prizes’ administrator in 2002, he replied:

> When the Pulitzer Board reviews an entry, it doesn’t discuss the circulation of the newspaper . . . , the location of the newspaper, the ownership of the newspaper, or whether or not the newspaper has won any previous Pulitzer prizes . . . . The judging is based solely on excellence, comparative excellence.

(Cunningham, 2002)

The gravitas attached to leading national awards (Shepard, 2000) makes it likely that those selected for the juries would take their responsibility with a high degree of seriousness, and would therefore be able to list and prioritize criteria of excellence. How important, for example, is the subject matter of a story in determining its excellence? How much weight should be placed on the number of interviews or rigour of reporting, or on writing style, or on the originality of the theme? To what degree do issues of fairness, balance, and independence from sources come into play? And how important is the story’s public impact or benefit—the degree to which it serves the interests of democratic society (Rosen, 1999)?

As Bogart points out, awards “are, surely, an indicator of how . . . quality is assessed by colleagues,” even if “the subject of awards often carries in its train the epithet of elitism” (Bogart, 2004). “Prizes are the only way we have to keep score,” according to Thomas Kunkel, dean of the College of Journalism at the University of Maryland. “Every journalist you ever talk with will say our obsession with prizes is criminal . . . . But it’s the only quantifiable way of the industry recognizing you as a player” (Shepard, 2000, p.24). It is true that even the most venerable awards come under attack from time to time. After the return of the 1981 Pulitzer Prize awarded to the Washington Post’s Janet Cooke for what turned out to be a fabricated story (Green, 1981), one critic called the incident “the ultimate insult” to the Pulitzer Prize, which had already become “controversial and flawed, hurt by whispers and suspicions that some major newspapers organize their coverage to try to win Pulitzers and that some years the judges play politics in picking the winners” (O’Donnell, 1982). In 1994, Seymour Topping, the Pulitzer Prizes’ administrator, found himself in the position of denying, for the record, that the competition was a “crapshoot” (“Discussing the Pulitzer Prizes,” 1994). But in the main, the attention that continues to be given by journalists and news organizations to leading national awards is ample evidence of their prestige in the industry.

To gain insight into how excellence is measured in journalism, we surveyed judges in two leading Canadian print journalism awards programs about how they approached the task of assessing submissions and with what criteria in mind. Most of the judges were quick to emphasize subjective factors such as intuition and experience, rather than specific criteria. But when presented with lists of possible criteria drawn from the literature, they were able to rate and rank these criteria. We used a sequential method of supplementing rating and ranking questions with
open-ended questions and then compared the results. After comparing the answers to these various types of questions, we concluded that while judges acknowledge a variety of values, including social impact and importance, originality, analysis, and integrity, two values were emphasized consistently: writing style and research rigour.

**Exploring excellence: A limited literature**

Only recently has the study of excellence in journalism emerged as a discrete field of interest, and nothing close to an agreed list of consensus standards has yet emerged in the literature. The first book-length scholarly work to look specifically at the nature of excellence in journalism and the prospects for its achievement was published in 2001 (Gardner, Csikszentmihalyi, & Damon, 2001). Its authors compared the way journalists talk about their field to the way geneticists talk about theirs, and painted a bleak picture of the news business. After extensive interviews with reporters, editors, their audiences, scholars, and shareholders of media corporations, the authors reported that the field of journalism was “wracked with tension.” The stakeholding groups “differ sharply in their aspirations,” the authors wrote. Though most journalists are attuned to broad notions such as truthfulness and fairness, agreement on standards goes little further than that, while other professions have achieved a much higher degree of alignment between the generally accepted values of the profession and those of the culture in which they work. Under these circumstances, the authors concluded, the idea of achieving excellence was “but a distant dream.”

It is true that a substantial literature has existed for some time on the measurement of “quality” in journalism, using value-based criteria such as accuracy, impartiality in reporting, and investigative enterprise (Bogart, 2004) as well as quantifiable measurements such as the sources (wire copy versus local reporting) and contents of stories (Bogart, 1989; Lacy and Fico, 1991). This field, which has often been tied to a study of the link between quality journalism and business success (Meyer, 2004; Overholser, 2004), seems similar to the idea of research into standards of excellence, but the relationship between “quality” and “excellence” is murky. Gladney, for example, uses “excellence” to describe his research into how editors and readers rank journalistic standards (Gladney, 1990, 1996), but those standards are very similar to the “quality” criteria employed by Bogart, who himself appears to use the two terms interchangeably (Bogart, 2004). In any case, Gladney’s research—consisting in the main of mail surveys of newspaper editors in the United States—produced seminal rankings of nine “content standards” for excellence in newspapers (including news interpretation, lack of sensationalism, strong local coverage, visual appeal, accuracy, strong editorial page, comprehensive coverage, and good writing) and nine “organizational standards” (including integrity, staff enterprise, editorial independence and courage, and decency) (Gladney, 1990).

It is possible to see excellence as, essentially, quality in abundance. That is, “excellence” would refer, in effect, to high scores for “quality.” This notion, which seems to underlie Gladney’s work, seems entirely appropriate to his and others’
studies of how excellence is measured within a publication or organization. But excellence in individual works of journalism is another matter—both because many of the criteria do not apply (e.g., the amount of wire copy) and because some of the others (e.g., accuracy) would be seen by most journalists not as criteria of excellence for particular works but as bare minimum requirements.

In 2001, U.S. authors Kovach and Rosenstiel, reporting on the work of the Committee of Concerned Journalists, made a start on the road toward establishing consensus standards of excellence by listing, explaining, and illustrating ten “principles that journalists agree on.” The first is a statement of the purpose of journalism (“to provide people with the information they need to be free and self-governing”), and the other nine are statements about what is needed for journalists to achieve this goal. Those nine statements address journalists’ obligations of truthfulness and verification, their duties of loyalty to citizens and independence from those they cover, their role as monitors of power and providers of “a forum for public criticism and compromise,” the need for journalism to be interesting, relevant, comprehensive, and proportional, and the need for journalists to exercise freedom of conscience (Kovach & Rosenstiel, 2001). While these “elements of journalism” do not masquerade as agreed standards of excellence, they have reinvigorated a discussion on these standards and, generally, been well received by practising journalists and scholars alike.

But the most visible measure of excellence in journalism remains the major national awards programs. In a 1974 study of U.S. managing editors, Allen and Blankenburg reported that 68.4% of respondents held a “favourable” or “very favourable” view of journalism contests, and that 91.2% of their newspapers or staff had entered news or editorial contests in the past year (Blankenburg & Allen, 1974). In a follow-up study 15 years later, Coulson reported that 90% of editors considered awards valuable. Valuable in what way? For 80%, awards bolstered journalistic prestige; only 50% said they provided a measure of achievement, while 37% believed winning created a false standard of excellence (Coulson, 1989). Another study, in 1986, found that winners of journalism awards are more likely than their colleagues to enjoy organizational and occupational prestige (Beam, Dunwoody, & Kosicki, 1986). There have not been comparable studies in the years since, but the unabated energy devoted by news organizations and magazines to pursuing prizes and celebrating victories suggests little reason to suspect a substantial change.

The authors of the above studies, among others, acknowledge differences between two broad types of awards programs. The first type is judged by leading journalists and by people chosen and respected by journalists. The second group consists of dozens of less well-known awards—including many that are awarded by interest groups for stories that cover those groups’ fields, which raises obvious potential conflicts of interest (Zinman, 1970).
A more limited list of 43 popular awards, posted at Journalismjobs.com, runs from the American Association for the Advancement of Science Awards to the Women's Economic Round Table, Inc. prize for entrepreneurship journalism. Attitudes to the lesser awards will naturally skew journalists' responses when asked about journalism awards in general (Journalism Awards/Contests, 2005). Some hold the view, regarding interest-group awards, that many journalists "enter contests they know are jokes" (Ingle, 1982).

In this article, the awards under discussion are those national programs in which journalists in general hold a high degree of interest, rather than those given out by groups seeking to influence the media. This does not deny the fact that journalists remain skeptical, to some degree, about even major prizes, but there is no longer a serious question about the prestige associated with such awards. At least one pair of researchers considered the Pulitzer Prize so synonymous with quality that they used prize-winning to test the possibility that competition among newspapers leads to improvement in the quality of journalism (White & Andsager, 1990).

If, therefore, one is seeking a list of characteristics that reflect a de facto consensus on standards of excellence, an obvious place to look would be the guidelines developed by journalism's most respected awards programs.

Recognizing excellence: Awards programs' benchmarks

What, then, are the criteria of excellence according to which the arbiters of journalism's best work compile their annual lists of winners? We were surprised to find that the answer is as sketchy as the literature on quality standards. We searched for and requested lists of criteria applied by judges in several leading awards programs, including the Pulitzer Prizes, the George Polk awards, and national awards programs in Canada, Britain, and Australia. The result? Many programs simply don't have judging guidelines, while those of others consist mostly of lists of terms, without explanation or illustration.

Some of the Pulitzer Prizes' one-sentence category definitions provide terse clues as to what jurors are expected to look for. ("For a distinguished example of explanatory reporting that illuminates a significant and complex subject, demonstrating mastery of the subject, lucid writing and clear presentation, ten thousand dollars.") Others do not. ("For a distinguished example of reporting on national affairs, ten thousand dollars.") The Pulitzer Prizes website includes a list of frequently asked questions, of which Number 19 is: "What are the criteria for the judging of The Pulitzer Prizes?" The answer: "There are no set criteria for the judging of the Prizes. The definitions of each category . . . are the only guidelines. It is left up to the Nominating Juries and The Pulitzer Prize Board to determine exactly what makes a work 'distinguished' " ("The Pulitzer Prizes," 2005).

Isolated first-person accounts by members of Pulitzer juries, and published interviews with jurors, shed some light on the jurors' de facto criteria. In one of these published accounts, a juror—the managing editor of the Wall Street Journal—says: "I found myself using the same standards I use in evaluating Journal job prospects and whether to print stories: Are the stories factual? Are there good
explanations? Is there an absence of hype? Is the writing lively? Is there evidence of enterprise and initiative?” (O’Donnell, 1982). Most of these accounts tend to focus on the judging process, especially its intensity and pace (Cunningham, 2002; Dedman, 1991; McCormally, 1972). McCormally also alludes to the juror’s search for “cutting-edge” journalism, reportorial initiative, and “beautiful writing.” One story quoted 1999 Pulitzer jurors as saying they had asked questions such as “Is the subject significant? Is this entry really original and breaking new ground? Did the story make a difference?” One juror said: “I found myself looking for results as a way to separate competing entries . . . . If an entry showed good work but had no consequences, I would find myself gravitating to another entry that had impact.” Another said that he and his fellow jurors were drawn to “good, old fashioned reporting—finding something and pursuing it,” instead of pre-planning a major project (“Inside the Pulitzers,” 1999). A content analysis suggested that Pulitzer juries favour “information richness”—that is, they reward stories that used more, and more diverse, sources (Hansen, 1990).

Some awards do publish criteria. Criteria frequently highlighted in awards such as Australia’s Walkley Awards and the National Magazine Awards of the United States include enterprise and courage in reporting, intelligence in analysis, originality in theme, stylishness and storytelling skills, and the public impact and benefit of the submitted works (National Magazine Awards [USA], 2005; “The Walkley Awards—Categories,” 2005). In Canada, the National Magazine Awards (NMA) program offers judges brief explanations for each category, but these are supplemented by a general list of four criteria (each worth 25% of judges’ final score), without explanation or illustration. The four criteria are writing style, content, overall impact, and “how well the article engages the reader for whom it was intended” (Instructions for First-Tier Text Juries, 2004). Judges in the Canadian National Newspaper Awards (NNA) are asked to follow exceptionally detailed criteria (described as “guidelines”) for each category of the awards. For example, the guidelines for the Investigation category emphasize “enterprise and depth” and include a list of 13 questions to consider, including “Was this work a worthwhile allocation of this newspaper’s resources—does the subject involve a matter of reasonable importance to the public? Is this a significant exposé? Is the public interest or the rights of individuals at stake? Does this work emanate primarily from the initiative of the reporter/newspaper? Does this work expose secrets and/or wrongdoing? Does fact-gathering go beyond routine, drawing on computer databases, analysis, public records and authoritative (perhaps reluctant?) sources for its information?” In all categories, NNA judges are also provided with a set of notes listing the following elements for evaluation:

Idea: Significance (Was it worth reporter and reader spending time on?); Newsworthiness; Timeliness; Originality and creativity; Humorous; Initiative.

Reporting: Depth and breadth; Context and background; Accuracy; Fairness and balance; Comprehensive, relevant sources (officials and real people); Detail that engages reader; Answers reader’s questions; Enterprise and effort.
Writing: Language (precision of usage, elegance); Style, tone, mood (appropriate to content); Credibility/authority; Compelling lead/opening; Clarity; Strong focus/theme—what is this story about?; Structure and organization; Effective anecdotes quotes and examples; Narrative and description; Accuracy and fairness; Creativity/Risk-taking; Reader interest.

Overall impression: Excellent; Good; Indifferent. (“National Newspaper Awards,” 2005).

Drawing together the criteria of quality or excellence cited in all the sources mentioned above, it is possible to distinguish several types of criteria. Clearly, such a list should include the quality of reporting—that is, the depth, rigour, and volume of research and investigative enterprise. An excellent work of journalism should provide a benefit to society in terms of public impact, relevance, or service provided by the story. At its best, journalism is expected to provide thoughtful and logical analysis of reported material, to bring clarity to complicated subjects, and to place the facts in the context of their background and of the “how” and “why” questions behind the news. Excellent journalism should not just be original in content (breaking news stories or providing unique angles on the news) but should be innovative and engaging in the technique used to tell those stories. Journalists are expected to conduct their craft with integrity (that is, behaving with fairness and independence from their sources), and to many observers of the craft, it is important that their work be transparent in method—that is, the audience should be able to understand where the information comes from. Finally, the style of writing should demonstrate exemplary storytelling technique.

Investigating excellence: A method of inquiry

We set out to discover the extent to which this derived list of criteria was reflected in the thinking of judges in national awards programs. We not only wanted to see how judges were inclined to rank items on our list, but to gain an insight into the judges’ own ways of thinking about journalistic excellence and their routines in forming judgments about the relative merits of submissions. To do this, we needed to come up with a method of inquiry that would allow us to get beyond values to which judges might pay lip service and to glimpse their methods of prioritizing competing values when making their findings.

Survey sample

To pursue this research goal, our first challenge was to find a representative sample of judges to survey. We decided to interview judges in two leading Canadian national awards programs. The choice of these two programs for our first foray into this work was straightforward, and drawn directly from our research interests. We were seeking Canadian data, as a base for international studies. We wanted to study national programs offered by independent organizations (not companies or lobby groups) that offer awards for individual works of journalism for a variety of types and fields of journalism, rather than one particular type or purpose (such as the Michener Award for public service or the Canadian Association of Journalists award for investigative journalism). We needed to focus on
print, rather than broadcast journalism, to draw on the existing literature on quality standards. In Canada, only two awards programs fit these parameters, so we decided to interview 25 judges in various categories of the National Newspaper Awards (NNA) and 25 judges in various categories of the National Magazine Awards (NMA).2

To obtain our sample, we selected judges’ names randomly from lists provided by the awards programs’ organizers; we then set out to contact the selected judges and interviewed (by telephone) the first 25 reached in each program. To achieve this number, we contacted 61 NNA judges (of whom two declined to be interviewed and 34 did not respond) and 67 NMA judges (of whom four declined and 38 did not respond). All 50 interviewees were judges within the years 2001 to 2003, and all judged text entries, as opposed to visuals. Overwhelmingly, they judged categories involving features or explanatory writing, or local or beat or spot-news reporting, or reporting on politics, business, science, medicine, or international affairs. Of the 50 respondents, 32 had participated in more than one year, and 26 had adjudicated more than one category. In total, our respondents reported having adjudicated 17 NMA categories and 22 NNA categories. Some of them had been doing so for many years, even decades, but the median number of times respondents had adjudicated was two. In all cases where a judge had adjudicated more than once, we asked them to focus on the most recent experience in answering our questions. Depending on the category and program, they reported having read between 21 and 160 stories (median 45) in the course of adjudicating their category, and taking between five and 140 hours (median 22) to fulfill the adjudication task.

Qualitative approach
Having defined our sample, we needed an interviewing method that would allow us to secure the most accurate picture of judges’ values, priorities, and methods, as opposed to mere rankings of predetermined criteria. We chose to pursue a qualitative method involving in-depth interviews with this relatively small number of judges, rather than a mail-out quantitative survey. We made this choice because we were breaking new ground in this research. Journalism awards judges have never been formally surveyed about their criteria, and although the literature did suggest the existence of implicit criteria for excellence in journalism, these had never been formulated as a list of standards that are applied to individual works of journalism. We felt that we should allow judges to tell us about their considerations in an open-ended way, rather than having to fit their responses into a predetermined framework that might or might not correlate well to actual practice. We therefore set out to help establish a foundation for future research, not to arrive at definitive answers on judges’ criteria.

Although the earlier studies on editors’ criteria of quality, cited above, had relied on mailed surveys, we adopted telephone interviewing because it has some practical advantages over other types of surveys, namely reduced cost, quick turnaround time, and the possibility of greater standardization of interviews (Holbrook, Green, & Krosnick, 2003). Response rates are generally higher in face-to-
face interviews compared to telephone (Holbrook et al., 2003; Oksenberg, Coleman, & Cannell, 1986); however, our population was geographically dispersed across Canada, making face-to-face interviewing difficult and expensive. At the same time, telephone interviews generally have higher response rates than mailed surveys (Czaja & Blair, 1996).

As is the case with all survey techniques, there are shortcomings to conducting telephone interviews. It has been argued that when respondents are asked to use audio-only communication to make judgments, they tend to express less confidence in their judgment compared to face-to-face settings where show-cards or response cards (a written list of response options) may supplement oral questioning (Groves, 1990; Reid, 1977). However, the lack of visual stimuli attributed to telephone interviews is easily offset by having a cognitively sophisticated, articulate, and highly educated survey population (Alwin & Krosnick, 1991; McClendon, 1991). Our study participants were subject specialists, who had been selected as judges in part due to their status and expertise within their given professions. The questions asked were directly related to their areas of expertise, and terms used in the questions were drawn directly from the literature about journalism quality standards and, in many cases, from criteria used by awards programs, including the programs in which these respondents had served as judges. Furthermore, all response options in closed-ended questions involved professionally related terms, which judges already used in the context of evaluating excellence (part of criteria given to them by awarding bodies). The literature of telephone word-list recall tests clearly indicates better recall rates for more educated respondents (Hogervorst, Bandelow, Hart, & Henderson, 2004).

As for response rates, we were confident that respondents would be willing to participate, because informal conversations and a round of advance testing of our methodology showed the topic we were researching to be of interest to a sampling of judges in awards program; it was relevant to their own work, and the invitation to participate affirmed their expertise. In other words, we were not conducting “cold calls” to a general population. Our calls came from an institution that is recognized and respected for its work in the field of journalism—a factor which, according to Groves, is very important when predicting response rates (Groves, 1990).

When we constructed questions, considered question order, and provided response categories, we anticipated that because our respondents possessed topic-specific knowledge (item-specific expertise), acquiescence and recency response-order effects would not pose a significant problem in telephone interviews (McClendon, 1991). On the other hand, we were somewhat concerned that if the respondents were allowed to complete a self-administered, mailed questionnaire, they might be more inclined to provide “socially desirable” and more consistent responses between close-ended and open-ended items, to demonstrate their professionalism.

In other words, we feared that in a mailed survey, judges would be more inclined to describe how they should judge excellence, rather than how they did
judge it. Our goal was to understand what criteria were actually considered by those charged with judging excellence.

**Sequence of questions**

To achieve this goal, our interview questionnaire was built on a mix of closed and open-ended questions. The complete questionnaire began with questions about the judge’s background and the experience of being a judge (such as how many times the judge had participated in the awards program and why they believed they were selected as a judge). Subsequent questions probed aspects of the work done by the judge, including the number of pieces read, the method used to identify winners, the time judging took, and the quality of the entries. These process questions were followed by four questions about criteria. Other questions (the responses to which are not reported here) explored subject matter favoured by judges and the effect of knowing or not knowing the authorship of articles. Finally, the respondents were offered an opportunity to provide additional comments.

We were well aware that responses to attitudinal questions vary depending on preceding items in surveys (Krosnick & Alwin, 1987; Tourangeau, Rasinski, Bradburn, & D’Andrade, 1989). This is known as response-order effect or context effect (Rockwood, Sangster, & Dillman, 1997), where answers to specific questions are produced in stages, and outcomes of the different stages can be affected by previous items (Krosnick & Alwin, 1987). The questions about standards of excellence started some 10 minutes into the interview. The first of these asked about the judge’s criteria in an open-ended way:

**Question 4:** There are many ways to measure excellence in journalism and you probably had some particular criteria in mind as you tried to identify the best of the nominated pieces. What were those criteria? [Or, if the respondent could not remember or was unable to say: Generally, what do you think are the most important criteria of excellence in a piece of journalism?]

It was followed by two specific questions prompting the respondents to evaluate the 12 predefined criteria of excellence we had drawn from the literature and awards materials (see literature review, above):

**Question 6a:** How important would you say each of the following criteria of excellence are on a scale of one to five, with one being somewhat important and five being extremely important? (1) The originality or uniqueness of the subject matter. (2) The depth and amount of reporting or investigative rigour. (3) Transparency of method (that is, the extent to which the audience will understand where the information comes from). (4) Integrity, fairness, and independence from the story’s sources. (5) Thoroughness in exploring the context of the story (that is, asking the “how” and “why” questions). (6) Thoughtful and logical analysis of the subject matter. (7) Superior writing style. (8) Innovative and engaging technique. (9) Bringing clarity to a complicated subject. (10) Taking on difficult subject matter. (11) Breaking news. (12) The story is a benefit to society (public service).

**Question 6b:** I’m going to read the list [from Question 6 (a)] again. Can you tell me which two or three are the most important overall?
A later, open-ended question tried to draw out the judge’s highest benchmarks of excellence in light of the reflection induced by the foregoing questions:

Question 8: Thinking just about the best piece you ever judged, what did you especially like about it?

In this way, context items (including the first open-ended criteria question) were used to trigger memory. We then intentionally placed blocked items (list of 12 predefined criteria of excellence) before the open-ended “best piece” question, to be able to make further comparisons.

**Coding of responses**

After looking at the responses to open-ended questions, we identified 44 keywords indicating common responses and then coded each response for the keywords or clear synonyms. (Two student coders independently verified the classification of responses, and coding discrepancies were resolved case by case. No limit was set on the number of criteria coded in this way.) The criteria codes were then grouped under seven headings. In alphabetical order, these seven criteria groupings were as follows:

- **Balance or fairness.**
- **Insight, analysis, or examination of the context of the story.** (This grouping included references to the journalist having probed the questions “why” and “how” behind a story.)
- **Mention of criteria supplied to judges by the awards program.**
- **Originality** (included originality of idea and/or its importance, originality of subject, originality of execution, initiative, surprising angle, interesting subject).
- **Relevance and public impact or benefit of the story** (included relevance to community, public service impact, effect on society or reader, “leads to action on part of readers,” “challenges the reader,” value/utility for reader).
- **Reporting rigour** (included references to information gathering, depth of research or number of sources, thoroughness of reporting, imaginative research, new information, and factual detail).
- **Writing and/or storytelling techniques** (included writing quality, engagement and emotional impact, vivid presentation, compelling/captured attention, structure/architecture, creativity, style or flair, voice, and narrative technique).

The coders then recorded, as a single mention of a criteria grouping, each instance where a respondent had referred to one or more elements associated with that grouping. By setting no limit to the number of groupings recorded as having been mentioned by any respondent, we gave equal weight to each heading, regardless of how many times it was alluded to within a judge’s comments. Where a judge’s comment could be understood as fitting under more than one code, we recorded citations for each criterion (and the grouping containing each). In cases where a respondent used words that were not directly related to the criteria descriptions, we made assumptions as to what they were describing and recorded
for any connection or reference to the criteria—again, erring on the side of references to more criteria rather than fewer, so as to avoid the possibility that we might be stereotyping the respondent’s attitude, perhaps according to our own biases.

For example, one respondent described the best story they had ever read like this: “It took the reader into the story and I was able to visualize the problems. It was fair and balanced. The reporter interviewed all the important people in the story. The piece addressed a specific problem in the community that normally wouldn’t get addressed because people in the story were powerless. And the reporter found the story himself.” Here we recorded citations under the groupings writing style, fairness, reporting, relevance and public impact, and originality. Another respondent said: “Great writing, passionate engagement with subject matter. The writer had a burning, sweating, heaving need to convey his thoughts on the matter.” We coded only for writing style, because all of the descriptions clearly related to criteria keywords under that heading, such as passionate, engaging, and emotional.

After completing our coding and grouping, we tabulated and assessed our findings, as follows.

**Determining excellence: Judges’ criteria**

When asked about their criteria in an open-ended way, many judges tended to avoid naming specific standards. Instead, they would refer to their professional experience or tastes, or the overall experience itself. “I can’t really describe it. We don’t deconstruct it. I don’t look at it that way. Do they compel me? Do they draw me in?” said one judge. Another said: “If it gets my attention, keeps me reading, and I forget that I’m judging, then it’s worthwhile. It comes down to how the writer speaks to me.” However, when pressed for specifics, all the respondents were able to name criteria.

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Newspaper</th>
<th>Magazine</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing and storytelling</td>
<td>21</td>
<td>24</td>
<td>45</td>
</tr>
<tr>
<td>Rigour in reporting</td>
<td>14</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td>Originality</td>
<td>11</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>Relevance, public impact</td>
<td>9</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Insight/analysis/context</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Awards criteria</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Balance/fairness</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

*Note: Asterisk (*) indicates tie.*

Table 1 shows the number of judges who mentioned criteria in each grouping in response to the initial, open-ended question on judging standards. The groupings are then ranked according to the number of judges who mentioned each. Results are presented for the newspaper and magazine awards separately, and then
for the total sample. It seems fair to characterize judges’ responses to this preliminary and open-ended question about criteria as “top-of-mind” responses. By this we mean that judges may have a relatively small set of values uppermost in their minds before being prompted to consider a range of standards drawn from awards materials and the literature. Above all, these top-of-mind criteria for excellence seem to emphasize writing and storytelling style, followed (some way behind) by reporting rigour and then by originality in subject or approach, with relevance or public impact the only other theme to win mentions in the double digits.

Table 2: Average scores for twelve predefined criteria of excellence (Question 6a)

<table>
<thead>
<tr>
<th>Newspaper</th>
<th>Magazine</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity to complicated subject</td>
<td>4.5</td>
<td>Clarity to complicated subject</td>
</tr>
<tr>
<td>Context</td>
<td>4.5</td>
<td>Writing</td>
</tr>
<tr>
<td>Reporting</td>
<td>4.3</td>
<td>Analysis</td>
</tr>
<tr>
<td>Independence, fairness</td>
<td>4.2</td>
<td>Clarity to complicated subject</td>
</tr>
<tr>
<td>Writing</td>
<td>4.1</td>
<td>Reporting</td>
</tr>
<tr>
<td>Analysis</td>
<td>4.0</td>
<td>Context</td>
</tr>
<tr>
<td>Originality</td>
<td>4.0</td>
<td>Technique</td>
</tr>
<tr>
<td>Benefit to society</td>
<td>3.6</td>
<td>Independence, fairness</td>
</tr>
<tr>
<td>Technique</td>
<td>3.6</td>
<td>Originality</td>
</tr>
<tr>
<td>Transparency</td>
<td>3.5</td>
<td>Difficult subject</td>
</tr>
<tr>
<td>Difficult subject</td>
<td>3.4</td>
<td>Benefit to society</td>
</tr>
<tr>
<td>Breaking news</td>
<td>2.8</td>
<td>Transparency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Breaking news</td>
</tr>
</tbody>
</table>

*Note: Judges’ average scores, on a scale of zero to five, are rounded to one decimal point.*

While the two groups of respondents (magazine and newspaper judges) answered the above question in very similar ways, some differences between the two cohorts emerged in the next question about criteria. Table 2 shows how judges scored the 12 predefined criteria of excellence we had drawn from the literature, on a scale of one to five. The two scores that stand out for newspaper judges are context and clarity to complicated subject, while magazine judges seem to gravitate toward writing (reflecting their top-of-mind leaning) as well as analysis and, as with the newspaper judges, clarity to complicated subject.

These results were very close: rather than homing in on a few criteria, judges were ready to give ratings averaging 3.5 or higher not just in the areas of interest that reflect their own top-of-mind values, but to as many as 10 different criteria (with at least six criteria garnering ratings of 4.0 or higher). Because we had anticipated such a result, we moved on to ask respondents to name the three most important of the 12 criteria, and, as Table 3 shows, this indeed served to clarify matters. It seems fair to characterize these forced top-three responses as “tie-
break” rankings—that is, when push came to shove, the top three ranked standards would likely influence judges to prefer one contender over another. Seen this way, both groups of judges’ tie-break criteria for excellence clearly emphasize writing style and reporting rigour. However, some differences remained between newspaper and magazine judging. Newspaper judges gave strongest emphasis to independence and fairness along with reporting. Magazine judges chose writing most often (by far), with reporting in second place and other values substantially farther down the scale. Originality, analysis, and clarity to complicated subject are given some emphasis by both groups; context rivals these values among the newspaper judges, and technique among the magazine judges. Social benefit and other values receive relatively little emphasis among both groups.

**Table 3: “Most important” criteria (number of mentions) (Question 6b)**

<table>
<thead>
<tr>
<th>Newspaper</th>
<th>Magazine</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence, fairness</td>
<td>Writing</td>
<td>19 Writing</td>
</tr>
<tr>
<td>Reporting</td>
<td>11 Reporting</td>
<td>10 Reporting</td>
</tr>
<tr>
<td>Writing</td>
<td>9 Analysis</td>
<td>8 Independence, fairness</td>
</tr>
<tr>
<td>Analysis</td>
<td>8 Technique</td>
<td>8 Analysis</td>
</tr>
<tr>
<td>Originality</td>
<td>8 Clarity to complicated subject</td>
<td>7 Clarity to complicated subject</td>
</tr>
<tr>
<td>Clarity to complicated subject</td>
<td>6 Independence, fairness</td>
<td>6 Originality</td>
</tr>
<tr>
<td>Context</td>
<td>6 Originality</td>
<td>5 Technique</td>
</tr>
<tr>
<td>Benefit to society</td>
<td>4 Benefit to society</td>
<td>2 Context</td>
</tr>
<tr>
<td>Difficult subject</td>
<td>2 Context</td>
<td>2 Benefit to society</td>
</tr>
<tr>
<td>Transparency</td>
<td>2 Transparency</td>
<td>2 Transparency</td>
</tr>
<tr>
<td>Technique</td>
<td>1 Breaking news</td>
<td>1 Difficult subject</td>
</tr>
<tr>
<td>Breaking news</td>
<td>0 Difficult subject</td>
<td>0 Breaking news</td>
</tr>
</tbody>
</table>

Going into the survey, we would not have hazarded a guess as to what our fourth and final criteria question—open-ended, like the first—would reveal. It seemed possible that when asked to specify the qualities of the best story they’d ever judged, our respondents would throw us a curve and name entirely new considerations—casting earlier results into doubt or putting them into perspective. Instead, these results, reported in Table 4, unmistakably confirmed the leading chief finding from all the preceding questions: both newspaper and magazine judges continued to emphasize writing style followed by reporting rigour, though the gap between the top two criteria had narrowed.

**Weighing excellence: An iterative process**
When first asked about their criteria, judges tended to emphasize their intuition and experience, rather than specific standards. But when pressed, they could identify criteria, though the weight placed on individual criteria appeared to vary
depending on the how the question was asked. How may these varying responses be understood?

We began to see a clue toward answering this question when we juxtaposed the respondents’ answers to all four criteria-related questions. In order to do this, we used a system of equivalencies to compare coded responses to the two open-ended questions against scores given to the two questions listing predefined criteria. This allowed us to come up with eight general headings: quality of writing (including storytelling technique), research depth and rigour of reporting, evidence of integrity (fairness, balance, or independence), the social relevance (public benefit or impact) of the story, the originality of the subject matter, the quality of analysis, the attention to context, and the story’s clarity.\(^3\)

Only when we juxtaposed responses in this way did we start hearing the data telling a story—a story that is illustrated in Table 5. We suggest that our order of questions gave the respondents an opportunity to rethink their “top-of-mind” thoughts (Question 4) in light of a more wide-ranging list of accepted notions about journalistic excellence, to which they were able to give “initial grades” (Question 6a).\(^4\) The prompted list may have reminded individual judges of criteria of which they were aware (or perhaps, at least, of which they now felt they should have been aware), whether explicitly or not, at the time of judging—but which were not currently “top of mind” at the time of our survey. Then, when asked to name their “tie-break” (top-three) criteria, they were obliged to restate their understandings of the chief elements of excellence in terms of established values (Question 6b). And after that, when we asked them to identify the qualities of the “best-ever” story (Question 8), they received an opportunity to do so in a way that relied not only on their own top-of-mind standards but on the full spectrum of received standards.

It is abundantly clear that two themes—the quality of writing and the rigour of reporting—stood fast among judges’ top-rated elements, no matter which question about criteria was asked (as the dark shading in Table 5 illustrates). But a subtler, more dynamic picture of judges’ evaluations emerges from other juxtaposed responses (as the light shading illustrates). Apart from the two dominant themes, judges in both newspaper and magazine awards clearly have both origi-
Table 5: Juxtaposed criteria responses

<table>
<thead>
<tr>
<th></th>
<th>Writing</th>
<th>Reporting</th>
<th>Relevance</th>
<th>Originality</th>
<th>Clarity</th>
<th>Analysis</th>
<th>Context</th>
<th>Integrity</th>
<th>Final Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Newspaper</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Best Ever</em></td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>A</td>
</tr>
<tr>
<td><em>Top of Mind</em></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td><em>Initial Grade</em></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td><em>Tie-Break</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Magazine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Best Ever</em></td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>A</td>
</tr>
<tr>
<td><em>Top of Mind</em></td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td><em>Initial Grade</em></td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td><em>Tie-Break</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: In this table, the headings are sorted in order of the total sample's "best ever" ranking. The resulting tie for fifth place is broken first by A+ grades and then by top-three rankings. We have used shading to highlight what we see as the significance of the various juxtapositions. Dark shading highlights a theme that scored an A or A+ grade, and that was one of the three most commonly mentioned themes in response to all three of the other criteria questions. In other words, these themes held fast in judges' responses on criteria, no matter what questions they were asked. Light shading highlights a theme that scored an A or A+ grade but was not one of the three most commonly mentioned themes. In other words, these themes held fast in judges' responses to the criteria question but were not as significant as the themes highlighted with dark shading. The resulting to the top three is determined by A+ grades, then by top-three rankings.
\textit{Integrity} and \textit{analysis} on their minds when they think about journalistic excellence in an open-ended way. It is likely that such top-of-mind considerations will dominate when individual judges assess submissions alone. Then, when judging panels come together to compare and discuss the individual “picks” (which is the normal procedure, as other questions in our survey confirmed), the communal standards of the journalistic community might command more attention. If our data are telling an accurate story, it is likely that \textit{integrity} for the newspaper judges, and \textit{analysis} for the magazine judges, will be important considerations at this stage, since these themes were rated highly as “tie-break” values. In the end, though, if personal preference returns to the fore (as our respondents freely and repeatedly told us it does) in determining which story shines out as “best” among those that have made the final cut, judges will once again focus strongly on \textit{writing} and \textit{reporting}, though \textit{relevance} (in newspapers) or \textit{originality} (in magazines) will also come into play.

\textbf{Defining excellence: Does form trump content?}

In some respects, we were surprised by the results of our survey in light of our review of the literature and of awards programs’ guidelines. Certainly, the amount and rigour of \textit{reporting} underlies important elements in all the reviewed awards guidelines and is clearly and pre-eminently emphasized throughout the literature, though not necessarily in those words. For instance, Kovach & Rosenstiel (2001) refer to the comprehensiveness of the news and to reporting as a “discipline of verification,” and Gladney’s “content standards” include “strong local coverage” and “comprehensive coverage” (Gladney, 1996). Similarly, \textit{originality} is explicitly emphasized in all the reviewed awards guidelines and receives fair to strong emphasis from our respondents.

But why is \textit{writing} found in so pre- eminent a place in judges’ minds? Writing style is mentioned by Gladney as one of his nine “content standards” of editorial excellence, and it is an important aspect of one of the 10 “elements” of Kovach & Rosenstiel: the striving to “make the significant interesting and relevant.” But “literary style” finished last among seven attributes of editorial quality rated in Bogart’s survey of U.S. newspaper editors and managing editors (Bogart, 2004), and none of the annual awards guidelines studied by us gives particular prominence to writing style. A similar disconnect is seen with regard to \textit{relevance or public impact}, which is emphasized throughout the literature on the importance and purpose of journalism, and specifically in the works cited above by Kovach & Rosenstiel and by Rosen as well as many of the cited awards guidelines. Although this theme does seem to weigh on some of our respondents’ minds (especially among newspaper judges), it does so to a considerably lesser extent than writing and reporting.

Finally, it seems significant that both newspaper and magazine judges mention \textit{writing style} much more often than \textit{reporting rigour} when they describe qualities of excellence in an open-ended way (as shown by the number of mentions for these criteria in Table 1 and Table 4). We suggest, therefore, that \textit{writing} is likely to carry the most weight at the end of the judging process—when winners
get determined. In other words, where more than one story demonstrates excellence in reporting and in other aspects, judges will probably favour the one that they consider better written. In this sense, form seems to trump content.

Conclusion
Our study clearly suggests that writing style slightly outweighs reporting rigour as a factor of excellence in journalism awards judges’ minds, followed by originality and social relevance. Despite the widespread view that journalism’s primary social purpose is to serve the interests of democracy, judges were relatively slow to mention or emphasize the social benefit or public service of works of journalism; indeed, the content of journalism seems clearly secondary to more ideologically neutral (and, perhaps, more technical) aspects, such as storytelling skill and the number of interviews.

Given that our study is exploratory research, future forays and our own ongoing analysis may lead to other interpretations. The maturation of judges’ responses through the course of interviews could, for instance, be seen more psychologically. It is conceivable that, as one reviewer of our study has suggested, respondents adjusted their answers to meet their own perception that their initial comments did not fully answer our questions, thus provoking further responses that sought to “justify, amplify, and occasionally rethink” the initial responses. That said, the respondents did not give much indication of discomfort with the interview process, and though several respondents indicated (as reported above) that they themselves thought of their decisions as based more on subjective taste than concrete standards, they were generally expansive in their responses and affirmative in their concluding comments. This seems to indicate a positive and undefensive attitude to our questions.

Further investigation and analysis could include a larger, quantifiable study of jurors in North America’s leading print journalism awards, applying and testing the standards isolated in this paper. Such a study would allow for an exploration of differences among judges in various awards categories as well as compare Canadian and U.S. judges’ attitudes to, for example, the relative importance of the social purpose of journalism. Meanwhile, we hope we have made a modest contribution toward a more developed understanding of the process and criteria applied in evaluating excellence in journalism. We also hope to have provided a promising survey method for exploring these questions further.

Notes
1. The authors acknowledge with gratitude the assistance of Bryan Cantley and staff of the Canadian Newspaper Association (which administers the National Newspaper Awards) and of Pat Kendall, Barbara Gould and Terry Sellwood of the National Magazine Awards Foundation. A version of this paper was presented at the conference of the Association of Educators in Journalism and Mass Communication in San Antonio, Texas, August 11, 2005. Selected highlights of the study were informally presented at a conference on “The Best Australian Journalism of the 20th Century” hosted by the Royal Melbourne Institute of Technology in Melbourne, Australia, on November 26, 2004.
2. The annual National Newspaper Awards, established in 1949, are administered by a board of governors representing both daily newspapers and the public. The awards are administered by the
Canadian Daily Newspaper Awards Programme Administration Corporation, which is associated with the Canadian Newspaper Association. The NNAs recognize excellence in 22 categories of journalism published in Canadian daily newspapers, of which 15 categories are for text as opposed to visuals. In 1995, there were 884 submissions. Category winners receive $1,500 and runners-up receive $250. The submission fee is $35 per entry in each category, but $25 for smaller newspapers (“National Newspaper Awards,” 2005). The annual National Magazine Awards, which were founded in 1977, are run by the National Magazine Awards Foundation, a non-profit group governed by a board and drawn primarily from people actively involved in creating Canadian-owned and operated magazines. The NMAs recognize excellence in 35 categories of journalism published in Canadian magazines, of which 22 are for text. Winners receive Gold, Silver, and Honourable Mention awards; Gold awards are currently accompanied by cheques of $1,500, Silver by $500. Each year over 120 judges assess more than 2,000 submissions. The submission fee is normally $75 per entry in each category (National Magazine Awards [Canada], 2005). In both programs, the honoured work is in both English and French. Judges are drawn from the ranks of leading active and retired journalists, journalism educators, and members of the public with special expertise or interest in a particular field. Judges work in panels of three and generally adjudicate in just one category’s panel in any given year. Panels are provided with criteria specific to the program and the category, as described above. (Information on awards programs not drawn from cited publications was obtained directly from the awards organizers.)

3. Because of the interrelatedness of descriptors of context, clarity, and analysis, we juxtaposed open-ended responses to any of these three themes with each of these three predefined criteria, to allow any of these three to have equal weight against rival themes. Not juxtaposed—because juxtapositions were not intuitively suggested and because they won weak support from respondents—were the awards criteria group of open-ended responses, and the following scored criteria: transparency of method, taking on difficult subject matter, breaking news. The criterion innovative technique is not included in this juxtaposition because we see a high degree of overlap with the concepts of writing style and storytelling. In retrospect, we would not have included technique as a separate criterion, or would have defined it more rigorously in the ranking lists to distinguish it clearly from writing style. That said, none of the respondents baulked at or questioned a distinction between the two descriptors.

4. For the purposes of this table, we converted the original one-to-five scores to a crude grading scale from C to A+, using the following equivalences: C=3.0-3.4; B=3.5-3.9; A=4.0-4.4; A+=4.5-5.0.

5. Some judging panels elected to split the nominated pieces among themselves in order to identify potential winners. In these cases, the whole panel read only the high-potential entries.

6. Respondents were also generous with their time: the questionnaire took an average of 26 minutes to administer.

References


Cunningham, Brent. (2002). Excellence is the only guidepost. *Columbia Journalism Review, 41*(1), 40.


Discussing the Pulitzer Prizes. (1994). *Editor & Publisher, 127*(27), 58-60.


Shapiro, Albanese, & Doyle / What Makes Journalism “Excellent”? 445


