

On the Ethics of Management Research: An Exploratory Investigation

Douglas C. Frechtling · Soyoung Boo

Received: 18 July 2011 / Accepted: 20 July 2011 / Published online: 7 August 2011
© Springer Science+Business Media B.V. 2011

Abstract While there is an abundant academic literature on professional codes of ethics, there appears to be few devoted to assessing the compliance of management research with such codes. This article presents the results of applying the *World Association for Public Opinion Research (WAPOR) Code of Professional Ethics and Practices* to research articles based on probability sample surveys in the top three academic journals covering tourism, hospitality, and related fields. Four research questions are posed to focus application of the WAPOR Code to nearly 200 articles published in three recent years. Content analysis of these articles, documented by a measure of intercoder reliability, indicates that it is feasible for multiple coders to accurately apply the WAPOR Code to such articles. None of the articles examined complied with all WAPOR standards, and fewer than half of them complied with half of the standards. Finally, we find that there is some difference among the three journals in compliance, but this difference is relatively small. In sum, there is very little compliance with ethical standards in the field of management research studied here.

Keywords Code of ethics · Management research · Content analysis · Research ethics · Sample surveys · Intercoder reliability

Introduction

There is abundance of corporate codes of ethics, and the published research on them is plentiful (Fennell 2000; Kaptein and Schwartz 2008; Langlois and Schlegelmilch 1990; Long and Driscoll 2008; Svensson and Wood 2008). There is also a class of ethical codes developed by professional associations to guide their members toward ethical behavior (Christian and Gumbus 2009; Coughlan 2001; Gaumnitz and Lere 2002; Groves et al. 2006; Pater and Van Gils 2003; Skubik and Stening 2009; Wiley 2000). Adopting from Pater and Van Gils (2003, p. 765), we define “professional codes of ethics” as “written, distinct and formal documents, issued by professional associations, that attempt to guide the professional behaviour of their members.” Skubik and Stening (2009) maintain that “the most important role of a code is to explain the underlying professional values and principles” to guide association members (p. 520). They further note that these may be developed as “an aspirational guide and education tool for members” (p. 515) and may include “enforceable standards” (p. 520).

Purpose

Motivated by Coughlan’s (2001) recommendation that “additional studies are needed that explore the relevance and effectiveness of existing professional codes” (p. 157), we focus here on the ethical guidance provided for the conduct of research for management. Chia (2002) distinguishes management research as dealing “fundamentally with the production and legitimization of the various forms of knowledge associated with the practices of management” (p. 1). These practices of management include

D. C. Frechtling (✉) · S. Boo
Department of Tourism and Hospitality Management,
School of Business, The George Washington University,
2201 G Street, N. W., Suite 301, Washington, DC 20052, USA
e-mail: frechtli@gwu.edu

S. Boo
e-mail: soyoungb@gwu.edu

human resource management (Wiley 2000), marketing (Rau and Kane 1999), research, finance, and operations (Datar et al. 2010). We adopt this definition of management research in this study.

We choose a specific field of management research and identify a professional code of ethics that pertains to that field. We operationalize the standards in the code with statements that can be applied to articles published in academic journals to indicate compliance or non-compliance with each. To shed light on the ethics of management research, we compose four research questions. We then identify nearly 200 articles published in the top academic journals in that field (management research) and apply the professional code to them. Finally, we propose answers to the research questions and state conclusions about the ethics of management research.

Background on Management Research Ethics

Relatively little has been published on the ethics of management research. Rau and Kane (1999) address the ethical issues that can arise in marketing research. They conclude that the establishment of “codes of ethics governing marketing research practice” (p. 144) is worthy of consideration. Payne (2000) explores the assumptions, values, ideologies, and other influences that affect the choice of business research practice, often at an implicit level. Ryan (2005) discusses duties of scientific inquiry in the field of tourism research and concludes that researchers must act with honesty and integrity while pursuing justice. Perdue (1991) examines the field of visitor surveys to determine the economic impact of tourists on a geographic area and provides a list of potential ethical problems. He suggests that presenting results from a convenience sample of visitors as being the same as those derived from a probability sample is unethical behavior.

Chia (2002) observes that management researchers are “governed by a code of practice established by a community of scholars” (p. 4). While some such codes may be implicitly understood, others take the form of formal professional codes of ethical conduct (Groves et al. 2009; Korac-Kakabadse et al. 2002).

One such formal professional code applicable to management research is promulgated by the World Association for Public Opinion Research (WAPOR) as the *WAPOR Code of Professional Ethics and Practices*. WAPOR was founded in 1947 to “(a) promote in each country of the world the right to conduct and publish scientific research on what the people and its groups think and how this thinking is influenced by various factors, (b) promote the knowledge and application of scientific methods in this objective, (c) assist and promote the development

and publication of public opinion research worldwide, (d) promote international cooperation and exchange among academic and commercial researchers, journalists and political actors, as well as between the representatives of the different scientific disciplines” (World Association for Public Opinion Research 2010, p. 1).

The WAPOR Code “defines professional ethics and practices in the field of public opinion research” (p. 1) and explains that the standards within it are promulgated in order

- “to advance the use of science in the field of public opinion research;
- to protect the public from misrepresentation and exploitation in the name of research;
- to maintain confidence that researchers in this field are bound by a set of sound and basic principles” (World Association for Public Opinion Research 2010, p. 1).

The “instrument of public opinion” referred to here is the scientific opinion poll defined by three characteristics:

1. Designed to measure the views of a specific group of humans;
2. Respondents are chosen according to explicit criteria in order to ensure representation of the group;
3. Survey questions are “worded in a balanced way” (ESOMAR 2008, p. 5).

Scientific opinion polls, also called probability sample surveys, gather information for dealing with a number of management issues, such as market segmentation, customer satisfaction, and product planning (Groves et al. 2009).

Structure of the WAPOR Code

The *WAPOR Code of Professional Ethics and Practices* “prescribes principles of ethical practices for the guidance of its members, and a framework of professional standards that should be acceptable to users of research and to the public at large” (World Association for Public Opinion Research 2010, p. 1). Employing the classification scheme proposed by Gaumnitz and Lere (2004), the *WAPOR Code* contains 44 statements in five thematic areas:

1. Responsibilities of Researchers;
2. Responsibilities of Sponsor;
3. Rules of Practice Regarding Reports and Survey Results (i.e., disclosure of methods);
4. Responsibility to Informants;
5. Practice between Researchers.

In Gaumnitz/Lere terms, the WAPOR Code is a horizontal five-statement code. Its shape is 12, 5, 15, 7, 5, disclosure, responsibilities of researchers. In terms of tone,

it is overwhelmingly positive (“thou shalt” = 35 statements) rather than negative (“thou shalt not” = 9 statements). Finally, it appears to be aspirational rather than legal since the words “enforceable” do not appear and no mechanism is stated for reporting and investigating alleged violations (in contrast to Skubik and Stening 2009 and Academy of Management, n.d., p. 6).

WAPOR Code Article II, section C, specifies 14 “Rules of Practice Regarding Reports and Survey Results,” stating, “Every complete report on a survey should contain an adequate explanation of the following relevant points” (p. 3). These points are listed verbatim in Table 1.

The WAPOR Code specifies the ethical obligations of survey researchers toward the public, including their clients. This appears to be congruent with one of the “two broad aspects of ethical practice especially relevant for survey research” that academic and professional survey researchers recognize (Groves et al. 2009, p. 371). The other aspect of ethical practice required from survey

researchers—procedures directly affecting survey respondents—is not addressed here. Researchers’ ethical obligations to respondents are often embodied in law (Institutional Review Boards and other procedures) and deal with minimizing potential harm to respondents and maximizing benefits to them, including respect for persons and informed consent of respondents before their participation.

Groves et al. (2009) maintain that a broad aspect of ethical practice in survey research regards general standards of scientific conduct. These standards include following procedures that yield valid conclusions, as well as avoiding “plagiarism, falsification or fabrication in proposing, performing, reviewing research or on reporting research results” (p. 372). This area of ethical survey practice also requires disclosure of certain information about a survey and its conduct when the findings are publicly released.

The overall objective of this ethical practice is to encourage transparency in survey research, that is, complete disclosure of survey methods. This objective derives from what biologist Glass (1965) calls “the *ought* of science”: “a full and true report is the hallmark of the scientist, a report as accurate and faithful as he can make it in every detail. The process of verification depends upon the ability of another scientist who wishes to repeat a procedure and to confirm an observation” (p. 83). It is noteworthy that the *Academy of Management Code of Ethics “Professional Principles”* for research and publications state similar objectives (Academy of Management, n.d., p. 4). Moreover, Michalos (1991, p. 416), in a different context, proposes eight characteristics that publishers of results of public opinion polls of the electorate during election campaigns should provide so as “to maintain and even increase the benefits of public opinion polling while significantly reducing the costs.”

To provide a focus for our research, and to build upon the knowledge of a distinct field of management that we have acquired, we focus on sample surveys as management research in tourism, hospitality, recreation, and related fields. We investigate compliance with the WAPOR rules of a set of articles published in specific academic journals in these fields in recent years. We do so by defining Research Questions that indicate compliance with the WAPOR rules and applying them to the set of articles through content. After careful consideration of the consensus of the content coders, we determine whether and how the articles comply with the WAPOR principles. We draw conclusions from these findings and recommend approaches that can improve the compliance of management research articles and reports with ethical principles.

“Compliance” relating to codes of ethics includes auditing, verification, and enforceability (Kolk and van

Table 1 WAPOR Rules of Practice Regarding Reports and Survey Results

Every complete report on a survey should contain an adequate explanation of the following relevant points:
For whom the survey was conducted and by whom it was carried out
The purpose of the study
The universe or population to which the results of the survey are projected
The method by which the sample was selected, including both the type of sample (probability, quota, etc.) and the specific procedures by which it was selected
Steps taken to ensure that the sample design would actually be carried out
The degree of success in actually carrying out the design, including the rate
Of non-response and a comparison of the size and characteristics of the actual and anticipated samples
A full description of the estimating procedure used for all results that are reported, including the sample size on which it was based and weighting procedures used to adjust raw data
A full description of the method employed in the survey
The time at which the survey, if any, was done, and the time span covered in collecting data
The findings obtained
(Where the nature and the research demands it) the characteristics of those employed as interviewers and coders and the methods of their training and supervision
A copy of the interview schedule or questionnaire and instructions
Which results are based on parts of the sample, rather than the whole sample
A description of the precision of the findings, including, if applicable, estimates of sampling error

Source World Association for Public Opinion Research (2010)

Tulder 2002; referenced in Fennell and Malloy 2007, p. 77). Examples of effective compliance factors in the published literature are unavailable, according to these authors. However, as at least one concrete example of compliance procedures in a professional code, we note The Academy of Management Code of Ethics includes a set of “Technical Standards [that] set forth enforceable rules of conduct for AOM members” (Academy of Management, n.d., p. 1).

To be effective, codes of conduct require explicit investigations of compliance and identification of instances of non-compliance, defined as behavior that does not conform to the prescriptions in a code of ethics (Fennell and Malloy 2007, p. 15). Wiley (2000) maintains that without an enforcement mechanism, professional codes degenerate into public relations tools.

Research Questions

We do not believe there is enough “conceptual development and concomitant empirical support” to justify presenting formal hypotheses here (Somers 2001, p. 187). Rather, we propose several Research Questions relating to published survey research for management examined here. Our findings regarding these Questions may lead to formal hypotheses that may be tested in later research.

Research Question 1: Is it practicable to determine compliance of published management research survey articles with WAPOR rules with an acceptable degree of reliability? This question addresses Compliance Assessment Feasibility: can coders with little coding experience consistently apply the rules to published journal articles? If there is little agreement among coders as to whether individual articles comply with individual standards, then there is little point in trying to apply the WAPOR standards to the published articles on probability sample surveys. On the other hand, if coders evidence a high level of agreement, then we can conclude it is practicable to apply the WAPOR standards to published articles to ascertain compliance with ethical standards for management research.

Research Question 2: Do a majority of these articles comply with most of the WAPOR rules? This addresses General Compliance of authors of management research articles with the WAPOR standards. If we find that most articles comply with most of the WAPOR standards, then we can fairly affirm that management research is ethical. On the other hand, if most of the articles fail to comply with most of the principles, then we can fairly deduce that management research is not ethical.

Research Question 3: Are most of the WAPOR rules widely observed in the articles while a few are not? This question addresses Specific Compliance with the WAPOR rules. If we find there are several principles that are widely ignored in published management research, but that many of the others are generally observed, then we can conclude that non-compliance is limited to a few specific standards. We could then conclude that while management research is ethical in general, there are a few areas of research ethics, which need to be observed for management research to completely comply with ethical standards for management sample survey research. Conversely, we might find that most of the standards are widely ignored, while only a few are generally followed, suggesting that management research is ethical only with regard to a few rules.

Research Question 4: Does the degree of compliance with the WAPOR Code differ significantly among the journals providing the articles? This last question addresses Publication Compliance. How it is answered indicates how widely management research ethics is observed among the journals. If we find that only one journal is the source of a majority of the non-compliance, then we cannot fairly conclude that management research is unethical, only that lack of compliance is centered in one source. The other journals can then be labeled sources of ethical management research.

Method

Content Analysis

Content analysis is a method of codifying the content of a selection of writing into various categories depending on specified criteria (Weber 1990). Holsti (1969) offers a broad definition of content analysis as any technique for making inferences by objectively and systematically identifying specified characteristics of messages. Although the term “content analysis” was first used in the field of communication, the practice of such methodology has been widely employed in exploratory research, theory development, hypothesis testing and applied research (Smith 2000).

Krippendorff (2004) viewed content analysis as a research technique for making replicable and valid inferences from data according to their context. Content analysis entails a systematic recording of a body of units, images, and symbolic matter, though not necessarily from the author’s perspective. The overall goal of content

analysis is to identify and record relatively objective characteristics of messages (Stemler 2001). Tucker et al. (1999) and Gaumnitz and Lere (2002) apply this technique to the analysis of professional codes of U.S. associations.

Hayes and Krippendorff (2007) maintain that generating data may take the form of judgments of kind (in which category the unit belongs), of magnitude (how prominent an attribute is within a unit), or of frequency (how often something occurs). We apply content analysis for the third purpose in this article.

In general, manifest content analysis (i.e., surface elements that are physically present) and latent content analysis (i.e., coders' subjective interpretations) are the two distinguishable areas central in the application of content analysis. Initially, content analysis dealt with quantitative descriptions of the manifest content of communications (Krippendorff 2004). Its application has been later expanded to include the study of latent content analysis for quantitative measures in various fields including tourism management (Choi et al. 2007; Malloy and Fennell 1998; Murphy 2001; Pan et al. 2007) and business ethics (Gaumnitz and Lere 2002; Stohl et al. 2009).

Latent content analysis is subjective and complex because the coders' own mental schema intervene (Potter and Levine-Donnerstein 1999). Owing to this potential difficulty, a more systematic coding scheme may be required. On the other hand, manifest content analysis is simple and direct, and can provide more objective information. Inferences about latent meanings of messages are therefore permitted (Holsti 1969). As recommended by Holsti (1969), our study employs blended manifest and latent content analysis and relies on observers' judgments regarding interpretation of textual matter (Hayes and Krippendorff 2007) to achieve our purposes.

Intercoder Reliability

"Intercoder reliability" is the term widely used for the extent to which independent coders evaluating a characteristic of a message reach the same conclusion (Kolbe and Burnett 1991). Neuendorf (2002) suggests that when human coders are used in content analysis, intercoder reliability quantifies the amount of agreement among two or more coders. Although intercoder reliability is often perceived as a standard measure of research quality (Kolbe and Burnett 1991), researchers have noted that most articles using content analysis do not provide information on intercoder reliability clearly or in significant detail (Riffe and Freitag 1997).

Nevertheless, researchers have emphasized the importance of measuring intercoder reliability in content analysis. For example, Neuendorf (2002) notes that content analysis exercises are useless without a measure of

reliability. Tinsley and Weiss (2000) discuss the necessity of intercoder agreement in content analysis, concluding that interpretations of the data cannot be considered valid without a measure of consistency. Kolbe and Burnett (1991) also emphasize the importance of measuring intercoder reliability, indicating high levels of disagreement among judges suggest weaknesses in the method. Overall, intercoder reliability is considered as a critical component of content analysis, and the key to reliability is the agreement of the opinions found among independent observers (Hayes and Krippendorff 2007).

Although a number of measures of intercoder reliability have been proposed, there seems to be no universally agreed-upon single measure (Hayes and Krippendorff 2007; Holsti 1969; Lombard et al. 2002). Instead, it appears that researchers select an index of intercoder reliability based on research assumptions and the characteristics of the data (e.g., the level of measurement of each variable). Among the various indices of intercoder reliability proposed by researchers, we choose to employ Cohen's (1960) kappa to provide the measure of agreement among coders in a content analysis.

One widely used method of the agreement between the pairs of observations is the simple percentage of agreement (Stohl et al. 2009). However, such percentages do not take into account the likelihood of chance agreement between coders that we would expect even if the two observations were unrelated (Grayson and Rust 2001). Cohen (1960) offered kappa to correct for such chance agreement by comparing the observed proportional agreement of two coders to the amount of agreement that would be expected entirely by chance. We employ Cohen's kappa in this article to indicate the amount by which the observed intercoder agreement exceeds the agreement which could be expected by chance alone, divided by the maximum that this difference could be. Specifically,

$$k = \frac{(p_a - p_c)}{(1 - p_c)}$$

where k is Cohen's kappa, p_a is the proportion of agreed on judgments, and p_c is the proportion of agreement one would expect by chance.

Although Cohen's kappa has drawbacks, it is considered to be one of the most reliable and useful measures of intercoder reliability by researchers (Lombard et al. 2002; Neuendorf 2002). Kappa is generally used only for measuring intercoder reliability for nominal level variables (Lombard et al. 2002). We have chosen to use Cohen's kappa as a measure of agreement because the content to which coding was applied consists of nominally scaled variables (i.e., compliance or non-compliance with the WAPOR standards, or non-applicability) and it is designed to measure the agreement of a pairing of coders. Kappa

coefficients for this study were calculated using the Statistical Package for the Social Sciences (SPSS).

Coding Categories

Creating categories is a core feature of content analysis. Categories may be thought of as compartments into which content units are placed. Holsti (1969) states that the definition of categories requires that they actually represent the elements of the investigator's study, so that items relevant to the study can be classified. Well-specified decision categories in conjunction with well-specified decision rules

will produce fewer discrepancies, even when used by relatively inexperienced coders (Krippendorff 2004).

Our intent is to apply the auditing function in compliance, "the assessment of performance against [a] goal or other stated criteria" (Fennell and Malloy 2007, p. 84). Through content analysis, we apply the "Rules of Practice Regarding Reports and Survey Results" from the WAPOR "Code of Professional Ethics and Practices" (World Association for Public Opinion Research 2010) to identify compliance and non-compliance of specific articles published in the top three academic journals in the fields related to tourism and hospitality. Table 2 shows how

Table 2 WAPOR "Rules of Practice Regarding Reports and Survey Results" and criteria for coders

Coding categories from WAPOR "Rules of Practice Regarding Reports and Survey Results": reports must state	Criteria for coders
(a) For whom the survey was conducted and by whom it was carried out	(a.1) Compliance = names of those who carried out the survey are stated (a.2) Non-compliance = states that the survey was conducted for or funded by an unidentified sponsor or client; otherwise Not Applicable
(b) The purpose of the study	(b) Compliance = purpose and/or objectives are stated
(c) The universe or population to which the results of the survey are projected	(c) Compliance = target population is defined
(d) The method by which the sample was selected, including both the type of sample (probability, quota, etc.) and the specific procedures by which it was selected	(d.1.) Compliance = sampling frame is specified (d.2.) Compliance = how sample was selected from the population is stated
(e) Steps taken to ensure that the sample design would actually be carried out	(e) Not applied because this statement is an intent prior to survey conduct that is better expressed by (f) below
(f) The degree of success in actually carrying out the design, including the rate of non-response and a comparison of the size and characteristics of the actual and anticipated samples	(f.1.) Non-compliance = statement is made that conduct of the survey differed from the survey plan without explanation of differences (f.2.) Compliance = final response rate is stated or both the initial sample size and the effective sample size are reported, allowing response rate to be computed
(g) A full description of the estimating procedure used for all results that are reported, including the sample size on which it was based and weighting procedures used to adjust raw data	(g) Non-compliance = statement is made that responses were weighted prior to analysis but weighting procedures are not explained
(h) A full description of the method employed in the survey	(h) Not applied because items d, f, g, and i are considered to comprise this method
(i) The time at which the survey, if any, was done, and the time span covered in collecting data	(i) Compliance = when respondents were first contacted and when researchers ceased collecting data from them are specified
(j) The findings obtained	(j) Compliance = one or more tables, charts OR graphs of the results is/are presented, OR the text reports salient findings
(k) (Where the nature and the research demands it) the characteristics of those employed as interviewers and coders and the methods of their training and supervision	(k) Compliance = survey conducted face-to-face or by telephone AND the interviewers are identified and how they were trained is stated; otherwise Not Applicable
(l) A copy of the interview schedule or questionnaire and instructions	(l) Compliance = Invitation and verbatim questions are stated OR data collection instrument is presented OR a web URL is presented where this instrument can be reviewed
(m) Which results are based on parts of the sample, rather than the whole sample	(m) Not applied due to difficulty finding this in a survey article that highlights results
(n) A description of the precision of the findings, including, if applicable, estimates of sampling error	(n) Compliance = sampling error for at least one significant question is stated

Sources World Association for Public Opinion Research (2010), and authors

eleven of the 14 rules were defined for application to journal articles presenting survey results in our chosen field. Two of the rules (e and h in Table 2) were not applied as they were deemed covered by other rules. One of the rules (m) was not applied because of the difficulty of determining if results reported were “based on parts of the sample, rather than the whole sample.” Unless an author explicitly states that results analyzed were based only on part of the entire sample, we would not know this. Hence, compliance with this standard can never be disproved.

For most rules in Table 2, the criterion of compliance can be clearly stated and, if not met by the journal article, then non-compliance is recorded for that rule. In other words, for most rules there are only two possible, mutually exclusive states: compliance or non-compliance. In three cases (rules a.2., f.1. and g.), the criterion is only applicable under certain conditions. If the condition for such a criterion is met in a study, then compliance or non-compliance can be determined. Otherwise, such assessment cannot be made. Therefore, three states are defined for these three specific rules: compliance, non-compliance (i.e., information is provided indicating the rule does apply, but the requirements of the rule are not satisfied) and not applicable because the required condition is not met and compliance or non-compliance cannot be determined. Finally, the reader should note that rules (a), (d), and (f) actually contain two criteria each that can be applied to the articles. Consequently, we have defined 14 criteria from the 14 WAPOR rules for application here.

Data Collection

There are more than 80 English language academic journals devoted to tourism, hospitality, and related fields, including recreation, sport management, meeting management, and event management (Frechtling 2010). Despite the large number of journals covering these fields, there has been considerable agreement about the three leading English-language tourism and hospitality journals over the years among academics (Howey et al. 1999; Marketing in Asia Group 2009; McKercher et al. 2006; Pechlaner et al. 2004) and practitioners (Frechtling 2004). These three journals are the *Annals of Tourism Research*, *Journal of Travel Research*, and *Tourism Management*. All these three journals employ double-blind peer review processes. All three have been published continuously for more than 30 years. Two of them are published in the United States, while *Tourism Management* is published in the United Kingdom. We think it is fair to conclude that these three journals provide a reasonably representative sample of probability sample survey research articles published in all journals covering tourism, hospitality, and related fields, and may even demand higher standards for publishing research than the other journals in the field.

In this study, we consider only full research articles published in these three journals from 2007 through 2009. Editorials, research notes, and book reviews are excluded from this analysis. Between 2007 and 2009 inclusive, *Annals of Tourism Research* published a total of 126 full research articles, and *Journal of Travel Research* published 123. On the other hand, *Tourism Management* published 284 full research articles during this period, more than the number of research articles published by the other two journals combined.

The articles examined in this study were obtained through complete enumeration. All the 533 articles published by the three selected journals during 2007–2009 were read independently by three coders to identify those articles that presented research using a probability sample survey methodology. Of the 533 articles, 200 (36% of the total) were found to use a probability sample survey methodology. Among these, 13 were found to address non-management issues, such as student satisfaction, curriculum design, and sociological issues. These were deleted from our population, as being not part of management research. The final sample of management-related articles employing probability sample surveys consists of 187 articles published in the *Annals of Tourism Research* ($n = 36$), *Journal of Travel Research* ($n = 54$), and *Tourism Management* ($n = 97$) from 2007 through 2009.

Coding Procedures

Two graduate students plus one author of the current study applied the coding process to the 187 articles. The two graduate students had completed a graduate course on probability sample survey research, but had little prior coding experience in content analysis. Since previous studies stress the importance of coder training (Holsti 1969; Krippendorff 2004; Lombard et al. 2002; Neuendorf 2002), the authors conducted specific training sessions for the student coders. This training session familiarized all coders with the purpose of this research and the content to be coded, as well as coding procedures and their application. This training included practice coding of a small number of articles not selected for this study. Each coder read all the articles and coded each according to the rules in Table 2 independently and without consultation from March 15 to April 30, 2010.

Results

Main Findings

Table 3 arrays the WAPOR Rules of Practice in descending order of percentage of compliance found in the articles.

Table 3 Compliance and non-compliance of journal articles with WAPOR Rules of Practice Regarding Reports and Survey Results, 2006–2009

Coding criteria from “Rules of Practice Regarding Reports and Survey Results”	Compliance	Non-compliance	Not applicable
a.1. States who carried out the survey	100%	0%	–
b. Purpose of the study is specified	100	0	–
j. Findings obtained are shown	100	0	–
a.2. Sponsor or client indicated but not identified	–	–	100%
g. Weighting procedures to adjust raw data explained	–	–	100
f.1. Survey conduct differs from survey plan	–	–	100
c. Target population is specified	87	13	–
d.1. Sampling frame is defined	75	25	–
f.2. Response rate or initial and effective sample sizes is presented	66	34	–
i. Time period of the survey is stated	61	39	–
d.2. Sampling selection method is identified	45	56	–
k. Interviewer characteristics, training, and supervision are presented	39*	61*	59
n. Estimates of sampling error are presented	16	85	–
l. Copy of the data collection instrument and instructions are presented or web source provided	12	88	–

Sources Table 1 and authors

* Proportions based on articles where this rule is applicable

All of the 187 articles analyzed in this study complied with specifying who carried out the survey (rule a.1), the purpose of the study (rule b), and the survey findings (rule j). Three of the rules were not applicable to any of these articles (a.2, g, and f.1), leaving eleven criteria as being applicable to the articles.

Overall, the content analysis results demonstrate that the articles reviewed comply universally with some criteria regarding study description and findings, while few comply with sampling- and survey-related standards. More than three-quarters of the articles specified the target population and sampling frame. About three-fifths of the articles specified the response rate and time period the survey was

Table 4 Intercoder reliability in assessing compliance and non-compliance of journal articles to WAPOR Rules of Practice Regarding Reports and Survey Results, 2006–2009

Coding criteria from “Rules of Practice Regarding Reports and Survey Results”	Cohen’s kappa between coders		
	A and B	A and C	B and C
a.1. States who carried out the survey	1.00	1.00	1.00
b. Purpose of the study is specified	1.00	1.00	1.00
j. Findings obtained are shown	1.00	1.00	1.00
a.2. Sponsor or client indicated but not identified	1.00	1.00	1.00
g. Weighting procedures to adjust raw data explained	1.00	1.00	1.00
f.1. Survey conduct differs from survey plan	1.00	1.00	1.00
c. Target population is specified	0.604	0.823	0.454
d.1. Sampling frame is defined	0.817	0.602	0.532
f.2. Response rate or initial and effective sample sizes is presented	0.862	0.632	0.749
i. Time period of the survey is stated	0.955	0.865	0.821
d.2. Sampling selection method is identified	0.794	0.645	0.460
k. Interviewer characteristics, training, and supervision are presented	0.788	0.775	0.773
n. Estimates of sampling error are presented	0.890	0.743	0.709
l. Copy of the data collection instrument and instructions are presented or web source provided	0.786	0.841	0.637

conducted. At the extreme non-compliance end of the scale, only about one in ten of the articles presented estimates of sampling error or presented the data collection instrument questions and instructions.

Results for the Research Questions

Research Question 1: *Is it practicable to determine compliance of published management research survey articles with WAPOR rules at an acceptable degree of reliability (addresses Compliance Assessment Feasibility)?* In order to address this question, we examine Cohen’s kappa measures of intercoder reliability for applying the 14 criteria to the 187 articles by the three coders.

Table 4 displays the kappa measures among the three coders for each of the WAPOR standards. While there is no objective standard indicating acceptable levels of intercoder agreement for Cohen’s kappa, Landis and Koch (1977) suggest that values greater than 0.75 indicate excellent agreement, while values between 0.40 and 0.75 indicate fair-to-good agreement.

As shown in Table 4, Cohen’s kappas for the coder A–coder B pair were higher than those of other pairs

(i.e., coder A–coder C, or coder B–coder C). Table 4 indicates that for the 42 possible kappas, 73% exceeded 0.75, the threshold for excellent agreement. The other 37% of the kappa scores all exceeded 0.40, placing in the fair to good agreement range. Therefore, we conclude that this study indicates that the answer to Research Question 1 is yes, it is practicable to determine compliance of published management research survey articles with WAPOR rules to an acceptable degree of reliability.

Before discussing the results regarding Research Questions 2, 3, and 4, we note that wherever intercoder disagreement was found in this study regarding a particular WAPOR criteria in a particular journal article, we jointly examined the disagreement and jointly settled on a single code for the instance. Therefore, the following discussion covers all the 187 articles assessed on all the 14 WAPOR criteria without disagreement.

Research Question 2: *Do a majority of these articles address all of the WAPOR rules (addresses General Compliance)?* None of the 187 articles examined adhered to all the eleven of the WAPOR standards examined here, indicating there is no General Compliance with the WAPOR standards in this population. Therefore, the answer to Research Question 2 is an unconditional negative. Indeed, there is no evidence that the authors or journal editors have any familiarity with WAPOR's rules or any set of standards regarding presentation of probability sample survey methodologies.

Research Question 3: *Are there some WAPOR rules that are widely observed in the articles while others are not (addresses Specific Compliance)?* As indicated in Table 4, all the 187 probability sample survey articles analyzed observed three of the WAPOR rules: who carried out the survey, the purpose of the survey, and presentation of findings. Eighty-seven percent defined the target population. Failure to do so prevents generalization of the survey findings to a known population. Three-quarters of the articles defined the sampling frame, the list of all members of the survey population, or rules for simulating such a list. Definition of the sampling frame is essential for assessing survey coverage error, that is, to what extent, members of the target population were excluded from being selected for the survey (Dillman et al. 2009).

Nearly two-thirds of the articles reported response rates or the information necessary to compute response rates, or designated the time period of the survey. A low response rate suggests the presence of nonresponse bias in survey results (Groves et al. 2009). Almost as many articles reported the time period the survey was carried out.

Nearly one half of the studies reported the sample selection method for the survey. A minority of articles reported interviewer characteristics and training or reported estimates of sampling error. Sampling error indicates

the precision of the survey results in describing the target population (Dillman et al. 2009). Without this information, survey precision cannot be assessed. Only one in eight articles displayed the data collection instrument or provided an address where it could be examined. While not all the journals may be able to devote space to presenting data collection instruments, some do (Byrd et al. 2009). As an alternative, when space is unavailable, an article may report the URL for a website displaying the instrument.

In sum, we note that there is a substantial degree of Specific Compliance for five of the WAPOR rules but less than two-thirds compliance with six of them. We find that the answer to Research Question 2 is yes, there is a substantial variety of compliance among the WAPOR rules. Some are widely recognized and reported, while others appear unknown among the majority of authors.

Research Question 4: *Does the degree of compliance differ significantly among the three journals (addresses Publication Compliance)?* To examine differences in compliance across the journals, chi-square tests were conducted. The chi-square test is useful for determining if there is a statistical difference between groups of nominal variables (Easterby-Smith et al. 2002). The critical level of statistical significance was set at 0.05, and the statistical test was two-tailed. This test found significant differences for only three out of the 14 rules:

- d.2. Sampling selection method is identified ($\chi^2 = 6.729$, $df = 2$, $p < .05$);
- i. Time period of the survey is stated ($\chi^2 = 6.216$, $df = 2$, $p < .05$);
- n. Estimates of sampling error are presented ($\chi^2 = 18.470$, $df = 2$, $p < .0001$).

Among the three rules, “Estimates of sampling error are presented” shows the largest difference in compliance across the three journals. In each case *Annals of Tourism Research* showed a significantly higher degree of compliance than *Journal of Travel Research* or *Tourism Management*. Between the latter two, the *Journal of Travel Research* shows compliance statistically greater than *Tourism Management* on two of these three, while falling behind *Tourism Management* on one.

We conclude that this study answers Research Question 4 with a qualified affirmative. There is weak evidence that *Annals of Tourism Research* shows significantly higher compliance in applying the WAPOR rules, but only for three of the 11 criteria applied. From our studies, we cannot conclude that this difference is because of to a conscious policy on the part of the editors of that journal. Rather, this difference may be due to randomness in the amount of survey information authors chose to submit with their articles in our sample.

Discussion and Conclusions

This study attempts to shed light on a research topic in business ethics that has received relatively little attention: is management research ethical? We focused on a particular type of management research: articles based on probability sample surveys published in the three top peer-reviewed journals in tourism and hospitality fields. We investigated compliance with a specific professional code of ethics, another area lacking significant research attention.

In order to focus our analysis, we examined four Research Questions. We concluded that journal articles' compliance with specific ethical standards for management research employing probability sample surveys can be ascertained through content analysis by two or more coders (Research Question 1). We further found that a majority of these articles do not comply with relevant ethical standards (Research Question 2). Indeed, very few of the standards are addressed in even a minority of the articles (Research Question 3). Finally, we found this lack of compliance with ethical standards characterizes all the three journals serving as sources of the articles (Research Question 4).

Overall, we conclude that published research for the distinct area that we have investigated, tourism management research utilizing probability sample surveys, is not ethical. This finding applies to both the authors and the editors of the journals examined. Both groups need to seek out ethical standards such as the WAPOR Code and ensure that their published studies abide by them.

However, our findings have another implication, as well. The *WAPOR Code of Professional Ethics and Practice* is not effective in securing compliance with standards of ethical practice in at least one area of management research. While it states, "Adherence to this code is deemed necessary to maintaining confidence that researchers in this field are bound by a set of sound and basic principles," (World Association of Public Opinion Researchers, n.d., p. 1), this code incorporates no enforcement mechanism. This might be as simple as an annual report of compliance in various fields such as we have done here for a branch of management research, or aggressive condemnation of noncompliance as practiced by the American Association for Public Opinion Research (Bohannon 2009).

This study of the ethics of management research has a number of limitations. It focused on articles published in academic journals, while there are other significant outlets for management research, such as books, reports, and postings on internet websites. It examined only articles based on probability sample surveys. Management research also employs qualitative research techniques, such as convenience sample surveys, focus group interviews,

content analysis, and participant observation (Easterby-Smith et al. 2002), which were not examined in this study.

In addition, this study examined articles published in three academic journals focused on tourism, hospitality, and related fields. There are many other journals on other subject areas, which publish management research. The study period was limited to 2007–2009, leaving other periods unexplored. Finally, the study applied a single code of ethics when there might be others that could be applied equally as well. Indeed, there may be ways of judging the ethics of management research other than applying professional codes.

These limitations suggest fruitful areas of further research. Focusing on the probability sample survey methodology, are there any among the 80 journals now published in tourism, hospitality, and related fields, which do conform to the WAPOR Code or other similar code? If so, how instructive can their review procedures be for other journals in these fields? How well do studies in management fields outside of tourism, hospitality, and related fields comply? Looking forward, will the journals singled out for analysis in this study require higher standards of compliance with the WAPOR standards in the future? If so, what costs and benefits can be discerned and documented for researchers and research users? In a broader context, what are the ethical standards for management research methodologies other than probability sample surveys? How do such management studies comply with these standards? What can be done to encourage compliance with published ethical standards for management research? Addressing these and other pertinent research topics can enhance our overall understanding of the degree of ethical compliance in management research and suggest improvements.

References

- Academy of Management. (n.d.). *Code of ethics*. Accessed April 26, 2010 from http://www.aonline.org/aomPrint.asp?ID=268&page_ID=240.
- Adams, J. S., Tashchian, A., & Shore, T. H. (2001). Codes of ethics as signals for ethical behavior. *Journal of Business Ethics*, 29(3), 199–211.
- Bohannon, J. (2009). Author of Iraqi deaths study sanctioned. *Science*, 323, 1278.
- Byrd, E. T., Bosley, H. E., & Dronberger, M. G. (2009). Comparisons of stakeholder perceptions of tourism impacts in rural eastern North Carolina. *Tourism Management*, 30(5), 693–703.
- Chia, R. (2002). The production of management knowledge: Philosophical underpinnings of research design. In D. Partington (Ed.), *Essential skills for management research* (pp. 1–19). London, UK: Sage Publications.
- Choi, S., Lehto, X., & Morrison, A. (2007). Destination image representation on the web: Content analysis of Macau travel related websites. *Tourism Management*, 28(1), 118–129.

- Christian, V., & Gumbus, A. (2009). Shades of gray: Applying professional codes of ethics to workplace dilemmas. *Organizational Management Journal*, 6, 178–199.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Coughlan, R. (2001). An analysis of professional codes of ethics in the hospitality industry. *International Journal of Hospitality Management*, 20(2), 147–162.
- Crompton, J. L., Lee, S., & Shuster, T. J. (2001). A guide for undertaking economic impact studies: The Springfest example. *Journal of Travel Research*, 40(1), 79–87.
- Datar, S. M., Garvin, D. A., & Cullen, P. G. (2010). *Rethinking the MBA*. Boston, MA: Harvard Business Press.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method* (3rd ed.). Hoboken, NJ: Wiley.
- Easterby-Smith, Thorpe, M. R., & Lowe, A. (2002). *Management research: An introduction* (2nd ed.). Newbury Park, CA: Sage Publications.
- European Society for Opinion and Marketing Research (ESOMAR). (2008). *ESOMAR/WAPOR guide to opinion polls including the ESOMAR international code of practice for the publication of public opinion poll results*. Accessed February 11, 2010 from http://www.esomar.org/uploads/pdf/ESOMAR_Codes&Guidelines_OpinionPolling_v5.pdf.
- Fennell, D. A. (2000). Tourism and applied ethics. *Tourism Recreation Research*, 25(1), 59–69.
- Fennell, D. A., & Malloy, D. C. (2007). *Codes of ethics in tourism: Practice, theory, synthesis*. Clevedon, UK: Channel View Publications.
- Frechtling, D. C. (2004). Assessment of tourism/hospitality journals' role in knowledge transfer: An exploratory study. *Journal of Travel Research*, 43(2), 100–107.
- Frechtling, D. C. (2010). *List of active, peer-reviewed journals in tourism, hospitality, sport and event subject areas*. Accessed June 10, 2010 from <http://home.gwu.edu/~frechtli/resources.html>.
- Gaumnitz, B. R., & Lere, J. C. (2002). Contents of codes of ethics of professional business organizations in the United States. *Journal of Business Ethics*, 35(1), 35–49.
- Gaumnitz, B. R., & Lere, J. C. (2004). A classification scheme for codes of business ethics. *Journal of Business Ethics*, 49(4), 329–335.
- Glass, B. (1965). *Science and ethical values*. Chapel Hill, NC: University of North Carolina Press.
- Grayson, K., & Rust, R. (2001). Interrater reliability. *Journal of Consumer Psychology*, 10(1&2), 71–73.
- Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: Wiley.
- Groves, B. W., Kucharewski, R., & Longsdorf, E. L. (2006). Ethical issues in recreation research as perceived by recreation, park resources, and leisure services faculty. In J. G. Peden & R. M. Schuster (Eds.), *Proceedings of the 2005 northeastern recreation research symposium* (pp. 120–129). Newtown Square, PA: U.S. Forest Service.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77–89.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Menlo Park, CA: Addison Wesley.
- Howey, M., Savage, K. S., Verbeeten, M. J., & Van Hoof, H. B. (1999). Tourism and hospitality research journals: Cross-citations among research communities. *Tourism Management*, 20(1), 133–139.
- Kaptein, M., & Schwartz, M. (2008). The effectiveness of business codes: A critical examination of existing studies and the development of an integrated research model. *Journal of Business Ethics*, 77(2), 111–127.
- Kolbe, R. H., & Burnett, M. S. (1991). Content-analysis research: An examination of applications with directives for improving research reliability and objectivity. *Journal of Consumer Research*, 18(2), 243–250.
- Kolk, A., & van Tulder, R. (2002). *International codes of conduct: Trends, sectors, issues and effectiveness*. Accessed October 4, 2009 from <http://www.ib-sm.org/internationalcodesconduct.pdf>.
- Kolk, A., van Tulder, R., & Welters, C. (2007). International codes of conduct and corporate social responsibility: Can transnational corporations regulate themselves? In A. Crane & D. Matten (Eds.), *Corporate social responsibility. 2: Managing and implementing corporate social responsibility*. Thousand Oaks, CA: Sage Publications.
- Korac-Kakabadse, N., Kakabadse, A., & Kouzmin, A. (2002). Ethical considerations in management research. In D. Partington (Ed.), *Essential skills for management research* (pp. 20–44). London, UK: Sage Publications.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage Publications.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Langlois, C. C., & Schlegelmilch, B. B. (1990). Do corporate codes of ethics reflect national character? Evidence from Europe and the United States. *Journal of International Business Studies*, 21(4), 519–539.
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587–604.
- Long, B., & Driscoll, C. (2008). Codes of ethics and the pursuit of organizational legitimacy: Theoretical and empirical contributions. *Journal of Business Ethics*, 77(2), 173–189.
- Malloy, D. C., & Fennell, D. A. (1998). Codes of ethics and tourism: An exploratory content analysis. *Tourism Management*, 19(5), 453–461.
- Marketing in Asia Group. (2009). *The MAG scholar journal ranking list 2009*. Accessed September 9, 2009 from <http://www.mag-scholar.com/news.htm>.
- McKercher, B., Law, R., & Lam, T. (2006). Rating tourism and hospitality journals. *Tourism Management*, 27(6), 1235–1252.
- Michalos, A. C. (1991). Ethical considerations regarding public opinion polling during election campaigns. *Journal of Business Ethics*, 10, 403–422.
- Murphy, L. (2001). Exploring social interactions of backpackers. *Annals of Tourism Research*, 28(1), 50–67.
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage Publications.
- Pan, B., MacLaurin, T., & Crotts, J. C. (2007). Travel blogs and the implications for destination marketing. *Journal of Travel Research*, 46(1), 35–45.
- Pater, A., & Van Gils, A. (2003). Stimulating ethical decision-making in a business context: Effects of ethical and professional codes. *European Management Journal*, 21(6), 762–772.
- Payne, S. L. (2000). Challenges for research ethics and moral knowledge construction in the applied social sciences. *Journal of Business Ethics*, 26(4), 307–318.
- Pechlaner, H., Zehrer, A., Matzler, K., & Abfalter, D. (2004). A ranking of international tourism and hospitality journals. *Journal of Travel Research*, 42(4), 328–332.
- Perdue, R. R. (1991). Ethical issues in tourism accountability research: Some case studies. In *Proceedings of the travel and tourism research association twenty-first annual conference*,

- June 10–13, 1990 (pp. 311–314). New Orleans, LA: The Fairmont Hotel.
- Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, 27(3), 258–284.
- Rau, P., & Kane, D. (1999). Marketing research ethics. In S. K. Chakraborty & S. R. Chatterjee (Eds.), *Applied ethics in management: Towards new perspectives*. New York, NY: Springer-Verlag.
- Riffe, D., & Freitag, A. A. (1997). A content analysis of content analyses: Twenty-five years of journalism quarterly. *Journalism & Mass Communication Quarterly*, 74(4), 873–882.
- Ryan, C. (2005). Ethics in tourism research: Objectivities and personal perspectives. In B. W. Ritchie & P. Burns (Eds.), *Tourism research methods: Integrating theory with practice* (pp. 9–19). Cambridge, MA: CABI Publishing.
- Skubik, D. W., & Stening, B. W. (2009). What's in a credo? A critique of the academy of management's code of ethical conduct and code of ethics. *Journal of Business Ethics*, 85(4), 515–525.
- Smith, C. P. (2000). Content analysis and narrative analysis. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 313–335). Cambridge, UK: Cambridge University Press.
- Somers, M. J. (2001). Ethical codes of conduct and organizational context: A study of the relationship between codes of conduct, employee behavior and organizational values. *Journal of Business Ethics*, 30(2), 185–195.
- Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluation*, 7(17). Accessed June 28, 2010 from <http://PAREonline.net/getvn.asp?v=7&n=17>.
- Stohl, C., Stohl, M., & Popova, L. (2009). A new generation of corporate codes of ethics. *Journal of Business Ethics*, 90(4), 607–622.
- Svensson, G., & Wood, G. (2008). International standards of business conduct: Framework and illustration. *European Business Review*, 20(3), 260–274.
- Tinsley, H. E. A., & Weiss, D. J. (2000). Interrater reliability and agreement. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 95–124). San Diego, CA: Academic Press.
- Tseng, H., Duan, C., Tung, H., & Kung, H. (2010). Modern business ethics research: Concepts, theories, and relationships. *Journal of Business Ethics*, 91(4), 587–597.
- Tucker, L. R., Stathakopolous, V., & Patti, C. H. (1999). A multidimensional assessment of ethical codes: The professional business association perspective. *Journal of Business Ethics*, 19(3), 287–300.
- Weber, R. P. (1990). *Basic content analysis* (2nd ed.). London, UK: Sage Publications.
- Wiley, C. (2000). Ethical standards for human resource management professionals: A comparative analysis of five major codes. *Journal of Business Ethics*, 25(2), 93–114.
- World Association for Public Opinion Research. (2010). *Constitution of the World Association for Public Opinion Research*. Accessed February 11, 2010 from <http://www.unl.edu/wapor/constitution.html>.
- World Association of Public Opinion Researchers. (n.d.). Accessed January 19, 2010 from <http://www.unl.edu/wapor/ethics.html>.