ABACUS, Vol. 37, No. 1, 2001

CHARLES D. BAILEY, JAMES R. HASSELBACK
AND JULIA N. KARCHER

Research Misconduct in Accounting Literature: A Survey of the Most Prolific Researchers' Actions and Beliefs

Incidents of research misconduct, especially falsification, in the hard sciences and medicine have been widely reported. Motives for misconduct include professional advancement and personal recognition. Accounting academics are under the same tenure, promotion and recognition pressures as other academics; and, without the life-and-death issues of medical research, rationalizing misconduct in accounting research seems much easier. It is reasonable to assume that dishonesty has occurred, even if little evidence exists about its prevalence. Further, accounting research can influence tax policy and market conditions, affecting people's lives. This study is, to our knowledge, the first attempt to survey accounting researchers directly as to their ethical practices; it uses the randomized response technique of addressing sensitive questions. To maximize its relevance to the most respected accounting research, this questionnaire was sent only to the most prolific researchers at U.S. colleges and universities—those who have published in the 'top thirty' accounting journals. The results indicate that serious misconduct has occurred among these established accounting researchers. The estimated percentage of seriously tainted articles in the top thirty accounting journals, based on self-reporting, is about 4 per cent, while the respondents on average *believe* that about 21 per cent of the literature is tainted. Faculty who were tenured more recently provide higher estimates of the falsification rate, and attribute more of the cause to external factors like tenure pressure. The article concludes by discussing implications and offering policy recommendations.

Key words: Accounting; Ethics; Fraud; Research; Scientific method.

Both the popular and academic press have noted the presence of falsified research in the hard sciences and medicine. Recently, front-page stories in major U.S. newspapers have reported that a cell biologist at Lawrence Berkeley Laboratories resigned after investigators with the U.S. Office of Research Integrity of the Department of Health and Human Services concluded he falsified key data tying cancer to electrical power lines (Broad, 1999; Schevitz, 1999). In another dramatic instance, a panel of scientists called falsifications by two German biomedical researchers

CHARLES N. BAILEY is a Professor in the School of Accounting, University of Central Florida, JAMES R. HASSELBACK is a Professor in the Department of Accounting, Florida State University, and JULIA N. KARCHER is an Associate Professor in the School of Accountancy, University of Louisville.

'unprecedented' and alleged that they either manipulated or made up data (Koenig, 1997). Desruisseaux (1999) concludes that fraud permeates all areas of academia and has become a worldwide epidemic.

Some scientists complain that a few rare instances of misconduct have received too much attention, and that the very nature of scientific experimentation results in the self-correction of errors in earlier research. In response to charges that scientific misconduct is overblown, Gunsalus (1997) notes that little data are available about the accuracy of scientific literature. Similarly, Davis (1999, p. 55) recognizes the newness of the field of research ethics and the importance of determining, first, whether there is a 'problem of research conduct or only the impression of one . . . [and if one exists] its extent and cause'.

This same state of knowledge clearly exists concerning accounting research, and we cannot afford overconfidence.[1] Accounting academics face the same internal and external pressures as other academics. They, too, wish to be recognized by their peers or, at a minimum, be promoted and granted tenure. However, scientists researching medical treatments have another, offsetting external pressure—the potential for their research to cure or kill—that isn't present in accounting research. Therefore, given that medical scientists have falsified research, we can expect the same of accounting researchers. In addition, although accounting research will not cure cancer, it can have real economic consequences and affect people's lives through its influence on tax policy and market conditions, for example.

Why would a person who feels called to become a scholar subvert the process itself? Kohn (1986) notes that a variety of factors play a part in the choice of an academic career—including mere chance as well as the obvious family/peer influence and the lure of worldly success. Citing psychiatrist R. A. Kubie, Kohn notes that scientific researchers (including, it is argued here, accounting academicians) undergo a long period of intense training, during which their own emotional problems may be overlooked. Hence they may complete their training with a variety of unresolved neuroses. Kubie reported several instances of neurotically induced fraud from his clinical experience with scientists. Furthermore, after much preparation for a career in academia, success is not assured. 'Rewards for industry, perseverance, imagination, and intelligence are often accidental: success or failure are determined by chance', as whether one's research results are statistically significant. 'Because of the long educational period candidates for a research career have to undergo, they have few opportunities to confront themselves with external reality. The educational process may thus select people who subconsciously desire to escape the

[1] Davis and Ketz (1995, p. 318) call the notion of the scientific method as a self-correcting mechanism 'patently false'. Broad and Wade (1982), who provide a chilling account of the history of scientific fraud, report that during the 1981 U.S. Congressional hearings, '[Al] Gore and his fellow Congressmen were moved to visible amazement and then anger at the attitudes of the senior scientists they had called as witnesses' (p. 11). The scientists were 'confident that their existing self-correcting mechanisms made fraud a no-win venture, [and] could not acknowledge it as a problem that might go beyond the mental imbalance of a few individuals . . . Yet their stance was undercut that day by the very rational testimony of . . . a Harvard medical researcher who had confessed to inventing an experiment . . . He was clearly not irrational' (p. 13).

external reality . . . [W]hen a hard working and even brilliant [researcher] encounters a series of unsuccessful experiments, which delay his planned and expected progress, he may be tempted to cut corners' (Kohn 1986, pp. 194–5).

Davis (1999) notes that 'Work on the ethics of scientific research has a relatively short history . . . I think we may usefully date [the formal study] of research ethics from 1982 . . . Not much was written about the ethics of scientific research until a series of scandals in the 1970s aroused interest in the faking of research' (pp. 48–9). Davis also notes that 'discussion of research ethics has so far been limited to empirical research' (p. 47), and largely in the biomedical disciplines. The attention to empirical research is understandable because the methods are more susceptible to manipulation and falsification than analytical research.

Previous surveys have identified the ethical practices that accounting researchers (and others) consider to be severe violations of ethics. Our question concerns the prevalence in actual practice of the subset of such violations that would result in false or unsubstantiated 'knowledge' being introduced to the literature. We use the well-established 'randomized response' technique, which ensures respondents' anonymity, to ask about acts that would lead to reporting false research results.

PREVIOUS RESEARCH

Consciousness of research ethics in accounting has risen in recent years, paralleling that in other disciplines. Keys and Hendricks (1984) proposed a code of ethics for accounting research that was built on those of the American Psychological Association, the American Sociological Association and the American Marketing Association.[2] Crain and Carruth (1992) discuss two broad categories of unethical practices: (a) activities that affect the validity of the results, from fraud and misrepresentation to careless errors; and (b) reporting and publishing practices, such as dividing one research project into a number of different publications and using the same data more than once. In the first category, they mention 'the slight modification of data to more closely resemble the expected outcome [and] the use of a wide variety of statistical models . . . to obtain attractive results' (p. 30). Davis and Ketz (1991) discuss Babbage's classification of scientific misconduct into four areas: hoaxing—the attempt to trick someone else; forging—the production of false data; trimming—the manipulation of data to ensure a favourable result; and cooking—the suppression of data. Several accounting studies have attempted to measure attitudes towards such practices and to assess their prevalence in accounting literature.

Attitudes Towards Research Falsification

Borkowski and Welsh (1998) reported the results of a survey of ninety-five accounting-journal editors about the ethics of author, editor and reviewer practices. Editors rated various practices on a five-point scale. As expected, 'deliberate

[2] For a recent discussion of expanding the marketing code of conduct to include marketing educators, see Sirgy (1999).

falsification or fabrication of data or information' was rated as the most unethical practice. It was viewed as more damaging than plagiarism, perhaps because it introduces false information, whereas plagiarism might not.

Engle and Smith (1990) sent a questionnaire to 700 randomly selected faculty members, eliciting their opinions on the ethical propriety of faculty participation in twenty-nine questionable activities, including falsification of research data. Of the 245 respondents, 97 per cent felt that falsifying data warrants extreme reprimand or dismissal. These attitudes closely parallel those of medical researchers surveyed by Korenman *et al.* (1998).

Prevalence of Serious Research Misconduct

Crain and Carruth (1992) mention the general lack of hard evidence documenting unethical accounting research activities. They discuss the motivations for such activities, including the pressure for promotion and tenure, the competition to get into the top journals, and the pressure for accreditation. Bamber *et al.* (2000) discuss several cognitive biases that cause reviewers to look more favourably on manuscripts that reject the null hypothesis and that are consistent with prevailing beliefs. Researchers generally are aware of these biases and may be motivated to alter research that does not meet these criteria. Lindsay (1998) suggests that accounting researchers often engage in activities which inflate statistical significance (understate α level)—ignoring the effect of multiple tests, arbitrary selection of variables in multiple regression analysis, and data mining. An example of such behaviour appears in the comments of one editor surveyed by Davis and Ketz (1995, p. 324), who mentioned that he had asked an author to perform transformations on data that appeared to be autocorrelated. The editor later saw the paper published in another journal with exactly the same regressions, but with no mention of the autocorrelation coefficients. Another editor in the same study (p. 323) reported seeing at least five cases where authors determined the theory and hypotheses after running the tests so they could report the strongest possible results.[3] Davis and Ketz (1995) surveyed editors, editorial board members and *ad hoc* reviewers of top tier accounting journals, asking if they had detected incidents of research fraud, and if they had not, to what extent they believed it was happening. Following the classifications from their 1991 article, their results showed at least some observation of forging (5 per cent), trimming (7 per cent), and cooking (24 per cent). The respondents' belief that such misconduct occurred was even higher. Twenty-nine per cent of respondents believed that forging took place, 50 per cent trimming, and 78 per cent cooking. Even if these subjects suspected

[3] We do not intend to demean exploratory studies, and accounting research is a fairly young discipline in which unexpected results deserve follow-up. However, one cannot treat the statistical values from multiple tests the same way as those from single tests. Lindsay (1998) goes on to state how little attention accounting researchers pay to this problem, which has generated much discussion in other disciplines, such as psychology. Along similar lines, accounting researchers have not paid sufficient attention to the issue of statistical power (Borkowski *et al.*, forthcoming). Furthermore, Bamber *et al.* (2000) note that accounting researchers may have overgeneralized early research, ignoring the sensitivity of that research to the selected design choices.

these actions were taking place in only a limited number of submissions to the accounting literature, the problem is serious.

Choo *et al*. (1993) tested compliance with the data availability requirements of *The Accounting Review*. Their results indicate a low willingness to comply, from which Choo *et al*. infer low reliability of the research results. They also note that the vagueness of the disclosure policy, in which authors are merely *encouraged* to make their data available, allows much latitude to circumvent its intention.

The Engle and Smith (1990) study also asked for the subjects' perception of faculty involvement in research misconduct. Although the majority believed that very few of their colleagues falsified data, the perception that at least some do so is disturbing. Similarly, Borkowski and Welsh (1998) had their ninety-five journal-editor respondents estimate frequency of occurrence, and obtained an average frequency estimate of 2.26 for falsification and fabrication acts, where 1 = often, 2 = sometimes, and 3 = never.

The Randomized Response Technique

Because many types of research misconduct are difficult to detect by direct observation or indirect inference, some means of self-reporting is a valuable approach to investigating the problem. Even with the strongest assurances of anonymity, however, fear that survey instruments can be traced to respondents may lead to under-reporting. The randomized response technique (RRT) can reduce this bias in many situations. Under the RRT, a subject is given two questions—one sensitive or incriminating and the other non-sensitive. The RRT uses a random number generated by the subject (e.g., from a coin toss or the serial number on a piece of currency) to determine which of the questions the subject should answer. In this way, only the subject knows which question they answered, but the researcher can derive estimates concerning the sensitive question. Warner (1965) originally developed the RRT for use with dichotomous responses, such as yes or no. Greenberg *et al*. (1971) expanded the technique to quantitative data, allowing responses on a continuous scale.[4]

The RRT has been used to elicit responses to sensitive questions in several areas of accounting and auditing. Buchman (1983) studied the reliability of auditors' working papers. Buchman and Tracy (1982) studied premature sign-offs by auditors using both conventional direct-response questionnaires and the RRT. They found more honest responses to sensitive questions in the RRT condition. Reckers *et al*. (1997) also found that auditors reported a higher incidence of early sign-offs using the RRT than using a direct method, lending further evidence that the RRT elicits true responses. The RRT also has been used to detect the levels of false sign-offs of auditing procedures by all levels of governmental auditors (Berry *et al*., 1987) to measure the perceptions of managers about the deterrent effect of external and internal auditors (Schneider and Wilner, 1990) and to estimate violations of the AICPA Code of Conduct by practising accountants (Gibson and Frakes, 1997).

[4] Fox and Tracy (1986) provide a good summary of randomized response techniques.

Attribution Theory

The factors prompting researchers to perform dishonest acts are clearly of interest, and would figure prominently in efforts to discourage such acts. Attribution theory recognizes a human tendency to attribute causality along a spectrum from 'internal' to 'external' (see, e.g., Arrington *et al*., 1985), and this *internality* dimension seems useful for our early investigation. Furthermore, it may be particularly easy to conceptualize and access in memory (Channouf *et al*., 1999). Tenure pressure, which we view as an external force, often is cited as a cause of ethical violation. Internal drives, such as a desire for recognition, also exist, and were the motivation for some of the more spectacular cases of fraud reported by Broad and Wade (1982).

RESEARCH DESIGN

This study builds on the previous work on accounting research misconduct by using the RRT technique to ask sensitive questions about personal behaviour. To maximize its relevance to the most respected accounting research, the questionnaire was sent only to the most prolific researchers employed by U.S. universities, based on the number of publications in the top thirty accounting journals.

Selection of Subjects

As a first effort to investigate the question of serious ethical violations, we chose to survey the most productive researchers for two reasons. First, if violations have occurred among this group, then we can have some confidence in their substantive importance, because these researchers' contributions have been substantial and influential. Second, this group represents a sort of 'acid test', because 'if the gold rusts, what will the iron do?' Accordingly, our first sample consisted of accounting academicians who have published eight or more articles in the top thirty accounting journals, based on the journal ranking that Hasselback *et al*. (1996) summarized from a review of past surveys of accounting faculty. Following the successful collection of that sample, we decided to send a second 'wave' to persons who have published five to seven articles in these journals.

A total of 330 accounting faculty members met the eight-publication criterion and were employed by U.S. universities during fall 1997. A similar number (333) met the criterion of five to seven publications. We omitted persons located outside the U.S. (about twenty persons) because business-reply envelopes are limited to the U.S., and other forms of response would disclose subjects' identity. Each subject received a package like the one shown in Appendix A, but with two manipulations described below.

Choice of 'Violations'

We chose to study only intentional violations that would affect the truthfulness of research reports. Thus, based on the violations enumerated by Engle and Smith (1990), Davis and Ketz (1991), and Resnik (1998), as well as discussions with colleagues, we arrived at the list shown on the questionnaire (called 'Instructions' in Appendix A).

Survey Design

The Instructions sheet requires the respondent to flip a coin three times. If the coin lands 'heads' all three times (probability = 1/8), then they are to respond to one question; if a 'tail' appears on any of the three tosses (probability = 7/8), then they respond to an alternative question. Thus, even if we could identify the respondent (which we cannot) we could not know whether any individual is answering the sensitive question (about their own behaviour) or the nonsensitive question (about behaviour of the general population). The statistical methods of Greenberg *et al*. (1971), however, allow the estimation of means and standard deviations for *both* questions.

In addition to the main question about ethical violations, we directly asked subjects how long they had been tenured (to investigate changes across time, and as a proxy for the time period of their research productivity), and for their opinion as to why a person might engage in such unethical practices (by allocating 100 percentage points between internal causes and external causes).

Manipulations in the Survey

Following Greenberg *et al*. (1971), we randomly split the sample into two groups, based on the probability that the respondent will be answering the sensitive question. Greenberg *et al*. recommend that the split be in approximately the same proportions as the two groups' probabilities of answering the sensitive questions. Thus for 7/8 of the subjects, the sensitive question was to be answered if *any* of the three coin-tosses produced a tail; and for 1/8 of the subjects, the sensitive question was to be answered if *none* of three tosses produced a tail. Hereafter, we refer to these as the 'low protection' and 'high protection' groups, respectively, because of the probability that each is answering the question about *themselves*. That is, subjects who realize the researchers would know that the probability is 7/8 that they are responding about themselves will feel less protected when answering the question.

The second manipulation concerned the question about attribution of causality. To avoid a position bias towards either internal or external causes, we split each of the probability–treatment groups in half and reversed the order of the two causal attribution questions.

Mailing and Response-Improvement Techniques

The first wave was mailed on 18 November 1997, and the second on 17 March 1998, by first-class mail. We expected that the high relevance and interest-level for the recipients would enhance response rates. Conversely, the unusual and highly sensitive nature of the survey could reduce responses. We employed the following response-improvement techniques: The cover letter was individually addressed and signed, and envelopes were addressed directly, not with labels. A dime was included for the coin toss, for convenience (and to make disposal of the survey an awkward option). To ease the response burden and ensure confidentiality, we refrained from asking a litany of demographic questions. Business reply envelopes were included. We sent postcard reminders to first-wave recipients one week after the original

mailing, then email reminders on 13 January 1998, after allowing for holiday-season delays. For the second-wave recipients, we sent an advance-notification email message on 17 March and a similar reminder on 8 April 1998.

RESULTS

As shown in Table 1, responses came from 107 of the 330 higher output (first-wave) subjects, for a response rate of 32 per cent. Interestingly, only 25 per cent of the subjects in the high-protection group initially responded, increasing to 30 per cent after follow-up requests with a special appeal to this group. This gives some assurance that response rates in the low-protection group were not diminished because of respondents' fears of discovery. Response rates for the second wave (lower output) subjects were somewhat higher, probably as a result of the pre-notifications and the avoidance of holiday distractions. The rate was 143/333 = 43%, consisting of 42 per cent of the low-protection group and 48 per cent of the high-protection group.

The overall response rates show no indication of bias. They were 214/573 = 37% for the low-protection group and 36/90 = 40% for the high-protection group, a non-significant difference ($\chi^2 = 0.23$, $p = 0.63$). The proportions of responses received from the low- and high-protection groups also is consistent with the proportions of 7/8 and 1/8, respectively, sent out ($\chi^2 = 0.83$, $p = 0.36$). Follow-up requests brought very few additional responses. We speculate that subjects originally understood the request and made a firm decision whether or not to respond; the time requirement was minimal, giving little reason to set it aside until later.

The position bias that we guarded against by manipulating the order of the internal and external attributions did not materialize, probably because respondents had no difficulty in dealing with a two-item list and assigning percentages. The attributions to internal causes were almost identical for the two sequences, at 36 and 35 per cent.

All but seven respondents were tenured. With few exceptions, the comments were favourable and supportive of the research effort. Twenty response forms had missing data; each analysis discussed below used all observations that included the relevant variables.

TABLE 1

RESPONSE RATES

	Higher output group		Lower output group		Combined		
	n	%	*n*	%	Rec'd	Sent	%
Low-protection	95	32.8	119	42.0	214	573	37.3
High-protection	12	30.0	24	48.0	36	90	40.0
Total	107	32.4	143	42.9	250	663	37.7

TABLE 2

MEAN RESPONSES TO 'PERCENTAGE OF ARTICLES AFFECTED' AND *t*-TESTS OF DIFFERENCE FROM ZERO

	Mean	*SD* of mean	*t*	*p*
Panel A: Respondents with ≥ 8 publications				
Self-reported cheating	5.08	1.81	2.81	0.0029
Others' cheating	17.14	5.30	3.24	0.0008
Panel B: Respondents with 5–7 publications				
Self-reported cheating	2.73	1.83	1.49	0.0693
Others' cheating	22.77	3.03	7.52	0.0000
Panel C: All respondents combined				
Self-reported cheating	3.68	1.31	2.82	0.0025
Others' cheating	20.92	4.14	5.06	0.0000

Percentages of Articles Affected

Greenberg *et al.* (1971, p. 245) present equations for estimating of means and standard deviations of responses to the two questions (see Appendix B). Applying these equations to our data yields the results summarized in Table 2. As shown in panels A and B, both groups show means for self-reported cheating significantly greater than zero, with the result being more pronounced for the more prolific publishers; the estimate of their percentage of affected articles is 5.08, versus 2.73 per cent for the less prolific group. The more prolific group produces an opinion of 17.14 per cent for the overall body of literature, versus 22.77 for the less-prolific group.

However, these means do not differ significantly between groups for either question ($p > 0.20$). Accordingly, the focus is on the combined results. The overall mean estimate for self-reported (actual) falsification is 3.68 per cent, which is significantly different from zero ($t = 2.82$, $p = 0.0025$, one-tailed). The estimate for the overall percentage of articles affected (in respondents' opinion) is 20.92 per cent, also significantly different from zero ($t = 5.06$, $p < 0.0001$). Thus respondents estimate the extent of research falsification to be about six times as great as the self-reported number.

The distribution of responses to the 'percentage affected' question adds some insight into the statistical results just reported. As expected, many respondents report zero cheating. In the low-protection group, 148/210 = 70 per cent reported zero cheating. These respondents likely are self-reporters, because it is doubtful that many people would be this optimistic about the honesty of the general population. However, 87.5 per cent (7/8) of this treatment group are *expected* to be self-reporting, and thus reporting zero if no cheating exists. This difference between actual and expected proportions is significant ($\chi^2 = 55.6$, $p < 0.0001$). Thus we can

conclude that a substantial number (approximately 87.5 – 70.0 = 17.5% of the sample) are reporting some degree of falsification on their own part.[5]

Figure 1 shows the ninety-one non-zero responses to the 'percentage affected' question. The two waves are shown separately, with the low-protection group at the left and the high-protection group at the right for each wave. Recall that the self-reporters might include any respondent in either group. As noted above, it is believed that most, if not all, of the respondents reporting 'zero cheating' are self-reporting. Our priors regarding the highest percentage numbers (40 to 80) are less strong; but these researchers either are *very* cynical about others or are responding about their own work.

Relationship Between Reported or Perceived Cheating and Tenure Length

Figure 2 shows the estimates of percentage of literature affected versus the respondents' length of tenure. A clear trend is evident, with respondents who have been tenured longer providing smaller estimates. This trend is significant in a regression analysis ($F = 5.32$, $p = 0.022$; deleting the outlier [at twenty years, 80 per cent] increases this to $F = 8.43$, $p = 0.004$).

Relationship Between Causal Attributions and Tenure Length

Figure 3 shows the percentage of attribution to internal motives, versus length of tenure. The subjects with longer tenure attribute more causality to internal motives. Again, although substantial variance is evident in the responses, the positive slope is significant ($F = 13.66$, $p = 0.0003$). Clearly, more of the younger faculty members are willing to attribute cheating to external pressures, while older faculty attribute more causality to internal reasons.

CONCLUSIONS AND RESEARCH IMPLICATIONS

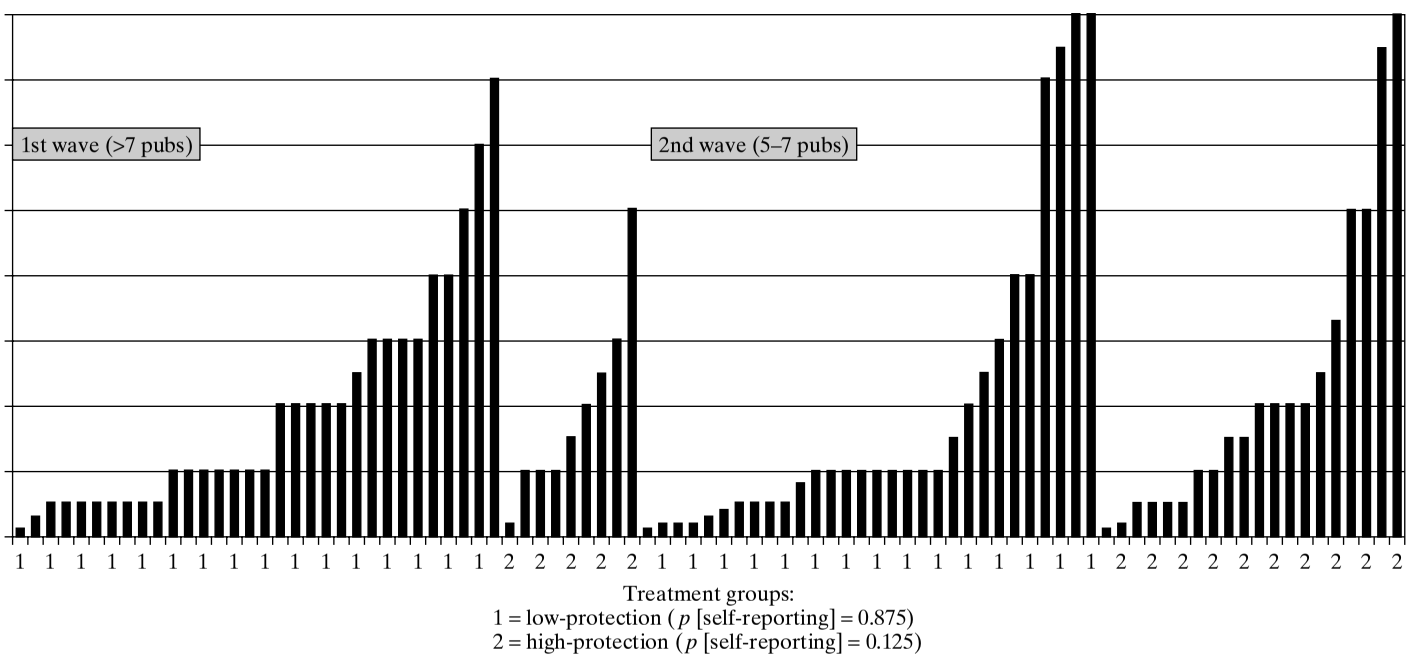
Our findings indicate that falsification has occurred among the most successful and prolific accounting researchers. The estimated percentage of seriously tainted articles in the top thirty accounting journals, based on self-reporting, is about 3.7 per cent, while the respondents on average *believe* that about 21 per cent of the literature is tainted. These findings seem generally consistent with the indirect investigations by Choo *et al*. (1993) and Davis and Ketz (1995), which indicate that some relatively low level of dishonesty occurs. In addition, Bailey *et al*. (2000), using both direct and random response techniques, have found a comparable rate of about 4 per cent for similar violations by academic economists.

Faculty members with longer tenure estimate or report lower rates of falsification. This could result from a real change in the rate of cheating over time, so that the reference point for behaviour (one's own or others') has shifted. Conversely, it could result from suppression or rationalization regarding one's own memory of actions early in a career. The finding does *not* associate increased cynicism with maturity.

[5] A similar test for the high-protection group is not significant, but is based on only thirty-four observations.

FIGURE 1

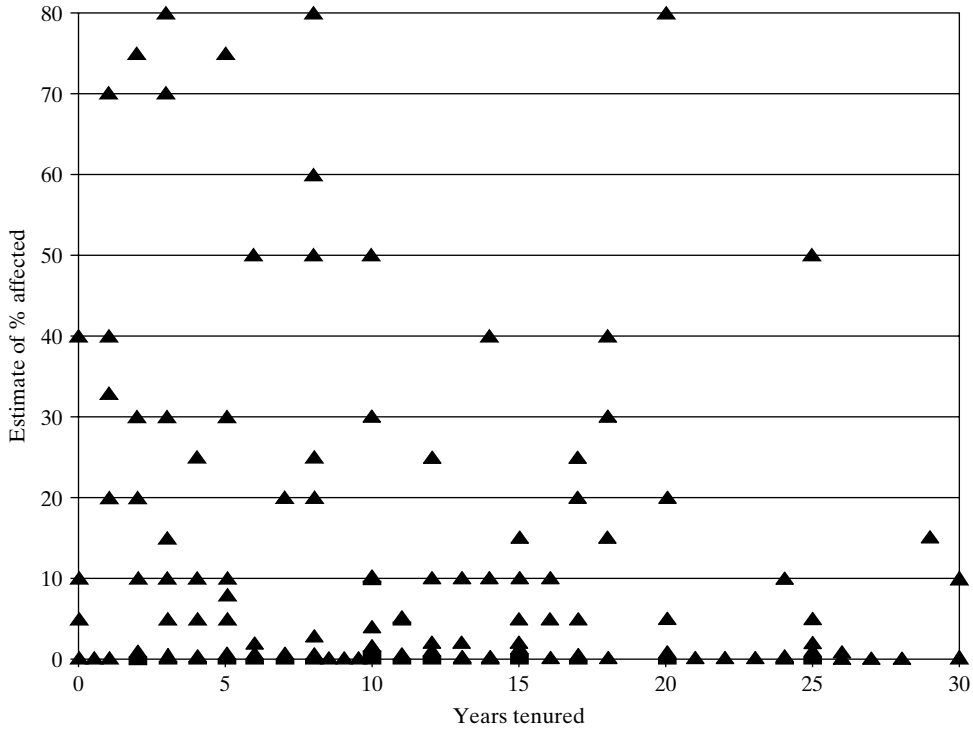
DISTRIBUTION OF THE NON-ZERO RESPONSES TO THE QUESTION OF PERCENTAGE OF ARTICLES AFFECTED BY FALSIFICATION



Note: Each bar represents a single response; there were 91 non-zero responses.

FIGURE 2

ESTIMATES OF PERCENTAGE OF LITERATURE AFFECTED
VERSUS LENGTH OF TENURE

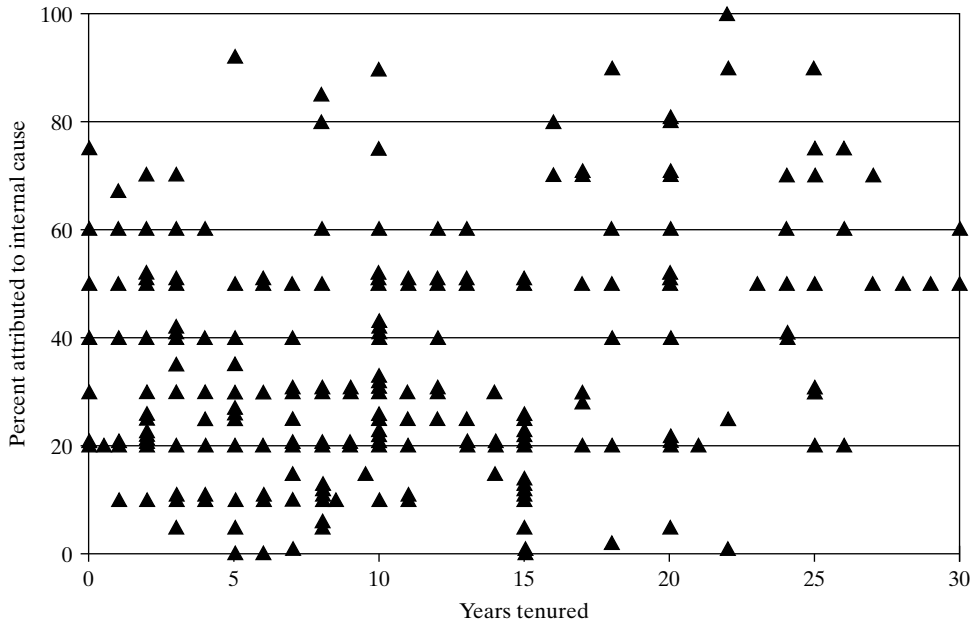


Note: For visibility, multiple data points in identical positions appear 'stacked'. The bottom of a stack denotes the actual value.

Our results also indicate that those with longer tenure tend to attribute more causality for dishonest behaviour to internal factors, while 'younger' faculty attribute it more to external pressures. One possible explanation would be that younger faculty have more recent memories of tenure pressure. However, the rather flat trend which is evident in Figure 3 for about fifteen years after tenure would argue for another cause, perhaps a cultural change or generational difference. Reports of recent trends towards dishonesty (Desruisseaux 1999) would support the latter explanation. Although there is a lack evidence on whether the rate of misconduct is increasing, Davis (1999, p. 54) lists nine reasons suggesting it may actually have increased. Two are particularly relevant to accounting. First, '[t]he increasing expense of particular sorts of research makes replication less likely (thereby weakening one check on a researcher's ability to deceive themselves or others)'. Consider the use of professional subjects like auditors; how likely are the firms to permit

FIGURE 3

INTERNAL ATTRIBUTIONS VERSUS LENGTH OF TENURE



Note: For visibility, multiple data points in identical positions appear 'stacked'. The bottom of a stack denotes the actual value.

replications, and how eager are researchers to consume valuable subjects for this purpose, versus new discovery? Second, 'Scientists increasingly feel that they are competitors in a rough business rather than cooperators in a noble enterprise and therefore feel less responsible for the misconduct of others in their field'.

Pressures on accounting researchers are likely to continue. Promotion, tenure and merit decisions at most institutions will always include publication results. Even with the mission-driven focus for accreditation decisions by the International Association for Management Education (AACSB) and its expanded definition of intellectual contributions to include a wider range of activities, most schools' missions will still require some level of scholarly research. Anecdotal evidence and observation also convince us that competition for journal space and for external grant funding has grown more intense in recent years.

Limitations and Research Implications

This early empirical study has several limitations, some of which suggest lines of future research. First, it is limited to the most prolific publishers, and a comparison with less prolific researchers seems a natural next step. Causal explanations from the study's data are limited; we wanted to minimize the response burden, and so

avoided excessive demographic questions. The technique does not allow assessment of the variance in individual behaviour; have a few established academics falsified *most* of their work, or have *a lot* of academics falsified *small* portions?

Future research might build upon the research on lying behaviour. While self-interest is the prevalent theory explaining dishonesty, role-conflict theory is a more recent competitor. In particular, Grover (1993, 1997) found lying to be predicted by an interaction between moral maturity and pressureful situations, particularly role conflict. Accounting academicians play multiple roles, apportioning their efforts between the demands of teaching, research and service. More principled people (who are less concerned with formal rules or laws and more with a higher set of moral principles) are less likely to lie in response to a role conflict, according to Grover's findings. The importance of role conflict versus self-interest, and ways to reduce role conflict, seem to be appropriate areas of inquiry.

In this work, only ethical breaches that are generally considered to be serious are examined. Future research might study actions where consensus is not so great, such as presenting the same paper at multiple meetings, cutting one research study into multiple parts, giving undeserved co-authorship, or refusal to share data. Also, this study looked only at author practices. Borkowski and Welsh (1998) surveyed journal editors about the ethics of editor and reviewer practices, such as favouritism, which could be another fertile research area.

POLICY RECOMMENDATIONS

Davis (1999, pp. 55–8) identifies five practical responses to misconduct: study, regulation, punishment, education and environmental change. These deserve consideration with respect to their status in the area of accounting.

Study An encouraging growth trend in ethics research already exists, as evidenced by the proliferation of new ethics journals and conferences. World-Wide-Web sites, such as the Online Ethics Center for Engineering and Science, help to coordinate conferences and discussion (see http://onlineethics.org/reseth/research.html). The study of *research* ethics, however, is at an early stage and raises interesting questions that can be addressed by a variety of methodologies. This paper represents an initial attempt to assess the overall scope of serious research misconduct.

Regulation This approach entails official actions by agencies such as professional societies, universities and governments. In 1989, the American Accounting Association adopted a policy on data availability for data-dependent results reported in the Association's journals. However, the policy only *encourages* that data be made available, rather than imposing strict requirements. Universities should also accept responsibility for regulation. Weingartner (1999) notes that

> [i]t can be difficult to determine whether a piece of work was done sloppily or whether there was intentional cheating, especially in research undertakings that are both complex and carried out in an atmosphere of fierce competition. If anyone can cope with the problems raised by research integrity, it is the institution at which the research is conducted. And since someone must, *it* must. (p. 92)

At a still higher level, U.S. Representative John Conyers has proposed making the falsification of scientific research a federal crime, regardless of who does it or where (Davis, 1999, p. 45).

In October 1999, the Office of Science and Technology Policy (OSTP) proposed a policy on research misconduct to be adopted by all U.S. federal government agencies. The OSTP proposal states that 'research institutions bear primary responsibility for prevention and detection of research misconduct, and for the inquiry, investigation, and adjudication of allegation of research misconduct' (OSTP, 1999, p. 5527).

Punishment Regulation entails enforcement and, ultimately, punishment. While a federal criminal law could call for imprisonment, universities are left with fewer options 'between "censure" (which outsiders are likely to regard as a wrist slap) and discharge (which seems too severe for all but the most serious wrongdoing). Research institutions need to experiment with such intermediate penalties as reduction-in-rank, revocation-of-tenure, and suspension-of-research-for-a-term-of-years' (Davis, 1999, p. 57). Crain and Carruth (1992), in a survey to 500 accounting professors to assess agreement about recommendations to improve ethical behaviour in accounting research, found strong consensus (63 per cent) favouring the establishment of a code of conduct, but little agreement about enforcement policies. Although the survey participants in the Davis and Ketz (1995) study reported incidents of research misconduct, they did not report what punishments, if any, were administered. Such information would be valuable in recommending future actions in cases of misconduct—and academics considering 'adjusting' their research results might have second thoughts if they see that others have been caught and punished.

Education The trend toward ethical content in the classroom needs to carry over to doctoral education. Specific discussions about the ethics of research has historically been rare in graduate programs, with students being expected to learn from rumours, anecdotes or occasional reprimands, if at all (Davis, 1999, p. 57). Anderson *et al*. (1994) found that two-thirds of their doctoral student respondents felt their departments did not prepare them to recognize and deal with ethical problems. Loeb (1994) suggests that informal interaction between students and faculty might not be adequate. Although formal inclusion of ethics education within the accounting doctoral curriculum is not without its problems (where it should be taught, who should teach it and what pedagogical methods should be used), 'it would help train future accounting faculty to meet both their own professional responsibilities and the needs of the practising portion of the accounting profession for accounting graduates with appropriate training in ethics' (Loeb, 1994, p. 824).

Students in doctoral seminars and research colloquia should be encouraged to question the behaviour of others. These forums generally provide a thorough grounding in the scientific method and research techniques, and an expansion to include research behaviour could be accommodated. Senior faculty should give serious consideration to ethical concerns as well as question their own research and that of their colleagues. Such open discussions would foster a commitment to the highest values in research behaviour. We see a place for case-study scenarios from

accounting research practice, perhaps developed through confidential survey techniques, to foster discussion and inspire ethical behaviour.

Finally, doctoral students should be encouraged to question the research that has gone before, particularly influential studies. Replication is a critical part of the whole process of scientific discovery. Meticulous replications and critical tests of past studies should be encouraged as valuable, as good training for younger academics and as an accepted part of dissertation studies. Integral to such an environment, of course, would be a greater openness by journal editors to the publication of replications, as we advocate below.

Environmental Change The factors discussed above will help to effect environmental change, as they will foster conscious attention to the subject of ethics. More direct changes to the environment could include modifications to the tenure process, although the course of action is not clear-cut. For example, would a de-emphasis on sheer numbers of articles, by limiting consideration to five, actually reduce pressure or make young researchers more dependent on letters of support from senior researchers (Davis, 1999, p. 57)?

Given this analysis of the current status of research ethics in accounting, two fruitful areas offer much potential for progress. They relate to the editorial peer-review process and the importance of establishing a research code of ethics.

Peer Review and Journal Policies

Given the key role of research-journal editorial processes as gatekeepers of research, the possible actions of editors and reviewers deserve special attention. Swan (1993) stresses the importance of peer review and replication studies in discouraging fraud. Review work, however, is unpaid and unglamorous. Although many reviewers conscientiously provide feedback to authors, they often do not have adequate raw data or time to find fraud, especially when the perpetrator presents internally consistent data or expands existing findings. At the same time, replication studies are viewed as less valuable than innovative work, are rarely published, and thus are unappealing projects. Lindsay (1998) believes that the multiplicity problem (the use of numerous tests, data mining, data selection) can be reduced if replications are done on different data. Clearly, editors need to give more than lip service to replications of key research findings before the findings become entrenched and immovable, as Bamber *et al.* (2000) imply.

Combined with the weak market for replications is what an economist might call

> a *market for statistical significance*, and hence a price to be had for achieving it. This would imply that statistical significance has some intrinsic or substantive value. A more cynical 'supply side' explanation is that statistical significance is easily acquired by more or less trained researchers and then peddled with great fanfare through the most respected academic journals. (Johnstone, 1988, pp. 326–7)

Johnstone further argues that 'the market for statistical significance is not essentially a market for *evidence* . . . [in which case we would see] great attention and journal space given to repetition or "replication" of interesting but *inconclusive* studies, of which of course there are many' (p. 327). The inherent noisiness of the

statistical environment of most accounting studies contributes to the inconclusiveness of the any single study, so that replications of 'interesting' studies should be welcome. Further, a failure to welcome replications implies that nothing of much importance hinges upon a correct or erroneous conclusion.

In looking at science as an economic venture, Wible (1998, pp. 23–42) suggests that 'seemingly replicable articles' (SRAs) and 'genuinely replicable articles' (GRAs) cannot be easily distinguished during peer review. That is, the data analysis in both types of articles appears to be based on solid data and could be replicated if one had the original data. However, a researcher would find an SRA impossible to replicate. For example, short cuts in obtaining or retaining the data might result in an SRA rather than a GRA. Time and other resource constraints motivate scientists to produce both SRAs and GRAs. Although a research agenda of only SRAs might result in greater productivity, GRAs are necessary to retain a scientist's reputation. Accounting academics can encourage GRAs over SRAs by increasing the publication and prestige of replication studies.[6]

In harmony with encouraging replications, journals might also publish more complete descriptions of the research designs, countering the general trend that Davis (1999, p. 55) has noted in scientific literature. More complete descriptions and background would facilitate replications.

Data Management Data should be available for criticism, statistical re-analysis, use in further research, and evidence that the research was in fact done. It is less clear, however, who should store the data or for how long (Resnik, 1998, p. 94), but the decreasing cost of computer storage media makes longer term storage increasingly viable and economical. We believe that journals and sponsoring associations should set standards for data availability, and that the knowledge of such requirements would tend to deter the irregularities studied here.

The Need for a Code of Ethics

Although the code proposed by Keys and Hendricks (1984) mentioned the need for researchers to do their work with integrity, the bulk of the document dealt with the responsibility toward human subjects. Although this code did not receive extensive exposure, a code of conduct is still high on the list of ways to limit research misconduct.

Of the ten 'possible efforts to improve research conduct' that Crain and Carruth (1992) listed, they found a majority agreeing or strongly agreeing that 'The AAA

[6] Wible specifies three approaches to replication. In a *direct replication*, the researcher uses the same methods and materials in an attempt to derive identical results. Such studies are done every day in science labs around the world where students rediscover the laws of nature. Many graduate programs in accounting also use replication studies as a way to teach research methods. However, these replication studies can be expensive and time-consuming for large empirical works. A *design replication* is not a copy of the original study, but uses an alternative research design. The original results are deemed to be replicated, in fact more strongly so, if this alternative methodology yields similar results. The final approach, *conceptual replication*, is more common in accounting. Here, the original research is assumed to be valid, and the researcher uses these results to extend the work. If the work can be extended, the original results are indirectly replicated.

should develop a code of research ethics' (p. 36) and 'Training in research ethics should be a required part of coursework included in Accounting doctoral programs'. Weingartner (1999) agrees with this philosophy, dismissing the 'wishful thinking of academics that . . . students and junior colleagues will somehow pick up [ethical values] by rubbing shoulders with their elders, without anyone's ever needing to *say* anything . . . Academic administrators must insist that . . . research groups spell out the rules appropriate to their work' (p. 92). Loeb (1990) also supports the adoption of a code of conduct for accounting academics as a way to reduce research errors, fraud and plagiarism. While outright falsification may not need to be spelt out as forbidden, the existence of specific ethical guidelines covering more ambiguous areas creates an atmosphere of ethical concern that should deter serious infractions.

A code of conduct could help doctoral students internalize ethical research standards. The Professionalism and Ethics Committee of the AAA held hearings in 1997 at various AAA meetings about a proposed 'Credo for AAA Members: Responsibilities and Rights of Accounting Academics'. The following revised version was presented to the Executive Committee of the AAA as 'American Accounting Association Statement of Responsibilities' in 1999:

> *Pursuit and Advancement of Knowledge*
> Another way members serve society is through scholarship: the discovery, application, and integration of ideas that advance the interests of society. Members fulfill their scholarship obligations in ways that match their personal abilities, passions, positions on the academic career cycle, and their institutions' specific mission. Some aspects of that obligation include
> - Establishing and maintaining a scholarship agenda that is directed to important societal needs and stretches the member's knowledge.
> - Maintaining and strengthening scholarly competence, consistent with the requirements of each member's own research agenda.
> - Accepting criticism of one's own work and providing objective and constructive criticism to colleagues. The criteria for criticism are its contribution to the public good and the quality of the work.
> - Publishing one's scholarship in ways that make the results available both to scholars and to a larger community.
> - Encouraging and promoting a variety of scholarship approaches that contribute to the public good and result in quality work.
> - Publishing research results that may call into question the profession's actions, when warranted.

An earlier version of the Code specifically included 'Intellectual honesty and objectivity' in the responsibilities under Pursuit and Advancement of Knowledge. Although the latest version does not have such an explicit statement in this section, it does include 'Conducting their academic, business, and personal affairs with integrity' within the Enhancement of Learning section. The proposed code was rejected, and the Professionalism and Ethics Committee of the AAA is unsure whether a revised code will be submitted to the Executive Committee in 2000.[7]

[7] Conversation with Mary Beth Armstrong, Chair of the Professionalism and Ethics Committee, March 2000.

Elliot (1995) stresses that the development of knowledge is the central facet of research and that all researchers have a stewardship responsibility to 'conform to the conventions of their research disciplines' (p. 5). Weil and Arzbaecher (1995) state that each research group has responsibility to clarify its own ethical standards, to identify the practices that support those standards and to monitor itself. Gardner (1996) also calls for the enforcement of scientific codes. He discusses the moral character of scientists—their trustworthiness. Empirical research must be performed using standard methods, reflect impartiality in writing, and comply with norms. In order to build on the achievements of others, researchers need to be able to trust their results. Given the evidence that some research misconduct exists in accounting publications, it is necessary for accounting researchers to explicitly agree to a code of conduct and hold themselves accountable for violations of it.

REFERENCES

Anderson, M. S., K. S. Louis and J. Earle, 'Disciplinary and Departmental Effects on Observations of Faculty and Graduate Student Misconduct', *Journal of Higher Education*, May 1994.

Arrington, C. E., C. D. Bailey and W. S. Hopwood. 'An Attribution Analysis of Responsibility Assessment for Audit Performance', *Journal of Accounting Research*, Spring 1985.

Bailey, C. D., P. Euzent, J. A. List and T. L. Martin, 'Economists Behaving Badly? A Survey on Three Areas of Ethical Behavior,' working paper, University of Central Florida, presented at the Southern Economic Association meetings, November 1998; revised April 2000.

Bamber, L. S., T. E. Christensen and K. M. Gaver, 'Do We Really "Know" What We Think We Know? A Case Study of Seminal Research and its Subsequent Overgeneralization', *Accounting, Organizations and Society*, February 2000.

Berry, L. E., G. B. Harwood, J. L. Katz, 'Performance of Auditing Procedures by Governmental Auditors: Some Preliminary Evidence', *The Accounting Review*, Vol. 66, No. 1, 1987.

Borkowski, S. C., and M. J. Welsh, 'Ethics and the Accounting Publishing Process: Author, Reviewer, and Editor Issues', *Journal of Business Ethics*, December 1998.

Borkowski, S. C., M. J. Welsh and Qinke (Michael) Zhang, 'An Analysis of Statistical Power in Behavioral Accounting Research', *Behavioral Research in Accounting*, Vol. 13, (forthcoming).

Broad, W. J., 'Data Tying Cancer to Electric Power Found to Be False', *New York Times*, 24 July 1999.

Broad, W., and N. Wade, *Betrayers of the Truth*, Simon & Schuster, 1982.

Buchman, T. A., 'The Reliability of Internal Auditors' Working Papers', *Auditing: A Journal of Practice and Theory*, Fall 1983.

Buchman, T. A., and J. A. Tracy, 'Obtaining Responses to Sensitive Questions: Conventional Questionnaire Versus Randomized Response Technique', *Journal of Accounting Research*, Spring 1982.

Channouf, A., J. Py and A. Somat, 'Cognitive Processing of Causal Explanations: A Sociocognitive Perspective', *European Journal of Social Psychology*, Vol. 29, Nos 5–6, 1999.

Choo, F., J. Huang, K. Leong and R. Wagner, 'Data Availability in Accounting Research: Some Evidence', *Accountability in Research*, Vol. 2, 1993.

Crain, J. L., and P. J. Carruth, 'Academic Accounting Research: Opinions of Academics on Recommendations for Improving Ethical Behavior', *Accounting Educators' Journal*, Vol. 4, No. 2, 1992.

Davis, M., *Ethics and the University*, Routledge, 1999.

Davis, S. W., and J. E. Ketz, 'Fraud and Accounting Research', *Accounting Horizons*, September 1991.

——, 'An Examination of Possible Occurrence of Fraud in Accounting Research', *Advances in Public Interest Accounting*, Vol. 6, 1995.

Desruisseaux, P., 'Cheating Is Reaching Epidemic Proportions Worldwide, Researchers Say', *Chronicle of Higher Education*, 30 April 1999.

Elliot, D., 'Researchers as Professionals', *Professional Ethics*, Vol. 4, Nos 3–4, 1995.

Engle, T. J., and J. L. Smith, 'The Ethical Standards of Accounting Academics', *Issues in Accounting Education*, Vol. 5, No. 1, 1990.

Fox, J. A., and P. E. Tracy, *Randomized Response: A Method for Sensitive Surveys*, Sage, 1986.

Gardner, W., 'The Enforcement of Professional Ethics by Scientific Societies', *Professional Ethics*, Vol. 5, Nos 1–2, 1996.

Gibson, A. M., and A. H. Frakes, 'Truth or Consequences: A Study of Critical Issues and Decision Making in Accounting', *Journal of Business Ethics*, Vol. 16, No. 2, 1997.

Greenberg, B. G., R. R. Kuebler, Jr, J. R. Abernathy and D. G. Horvitz, 'Application of the Randomized Response Technique in Obtaining Quantitative Data', *Journal of the American Statistical Association*, Vol. 66, Issue 334, 1971.

Grover, S. L., 'Why Professionals Lie: The Impact of Professional Role Conflict on Reporting Accuracy', *Organizational Behavior and Human Decision Processes*, Vol. 55, No. 2, 1993.

——, 'Lying in Organizations: Theory, Research, and Future Directions', in R. A. Giacalone and J. Greenberg (eds), *Antisocial Behavior in Organizations*, Sage, 1997.

Gunsalus, C. K., 'Rethinking Unscientific Attitudes About Scientific Misconduct', *Chronicle of Higher Education*, 28 March 1997.

Hasselback, J. R., A. Reinstein and E. S. Schwan, 'Toward the Development of Benchmarks to Assess the Academic Research Productivity of Accounting Faculty', working paper presented at the 1996 Annual AAA Meeting.

Johnstone, D., 'Comments on Oakes on the Foundations of Statistical Inference in the Social and Behavioral Sciences: The Market for Statistical Significance', *Psychological Reports*, Vol. 63, No. 1, 1988.

Keys, D. E., and J. A. Hendricks, 'The Ethics of Accounting Research', *Journal of Accounting Education*, Vol. 2, No. 2, 1984.

Koenig, R., 'Panel Calls Falsification in German Case "Unprecedented"', *Science*, 15 August 1997.

Kohn, A., *False Prophets*, Basil Blackwell, 1986.

Korenman, S. G., R. Berk, N. S. Wenger and V. Lew, 'Evaluation of the Research Norms of Scientists and Administrators Responsible for Academic Research Integrity', *JAMA, The Journal of the American Medical Association*, Vol. 279, No. 1, 1998.

Lindsay, R. M., 'Lies, Damned Lies and More Statistics: The Neglected Issue of Multiplicity in Accounting Research', *Accounting and Business Research*, Summer 1998.

Loeb, S. E., 'A Code of Ethics for Academic Accountants?', *Issues in Accounting Education*, Spring 1990.

——, 'Ethics and Accounting Doctoral Education', *Journal of Business Ethics*, October 1994.

Office of Science and Technology Policy, 'Proposed Federal Policy on Research Misconduct to Protect the Integrity of the Research Record', *Federal Register*, 14 October 1999.

Reckers, P. M. J., S. W. Wheeler and B. Wong-On-Wing, 'A Comparative Examination of Auditor Premature Sign-Offs Using the Direct and the Randomized Response Methods', *Auditing: A Journal of Practice and Theory*, Vol. 16, No. 1, 1997.

Resnik, D. B., *The Ethics of Science*, Routledge, 1998.

Schevitz, T., 'US Says Fake Data Tied Cancer, Power Lines', *San Francisco Chronicle*, 23 July 1999.

Schneider, A., and N. Wilner, 'A Test of Audit Deterrence to Financial Reporting Irregularities Using Randomized Response Technique', *The Accounting Review*, Vol. 65, No. 3, 1990.

Sirgy, M. J., 'Social Responsibility and the Marketing Educator: A Discussion Document', *Journal of Business Ethics*, Vol. 19, No. 2, 1999.

Swan, N., 'Baron Munchausen at the Lab Bench?', in S. Lock and F. Wells (eds), *Fraud and Misconduct in Medical Research*, BMJ Publishing Group, 1993.

Warner, S. L., 'Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias', *Journal of the American Statistical Association*, Vol. 60, 1965.

Weil, V., and R. Arzbaecher, 'Ethics and Relationships in Laboratories and Research Communities', *Professional Ethics*, Vol. 4, Nos 3–4, 1995.

Weingartner, R. H., *The Moral Dimensions of Academic Administration*, Rowman & Littlefield, 1999.

Wible, J. R., *The Economics of Science: Methodology and Epistemology as if Economics Really Mattered*, Routledge, 1998.

APPENDIX A

EXPERIMENTAL MATERIALS

COVER LETTER

November 18, 1997

Dr Joan Doe
Professor of Accounting
College 'O Bidness
Major Research University
College Town, ST 12345-6789

Dear Professor Doe,

You probably have seen the news reports about medical researchers who have falsified their studies. In accounting research, with no lives in danger, rationalizing falsification seems much easier. We can be sure that more than one person has done it under the pressure of tenure review, and you probably have wondered how common such practices are.

I hope that, as a researcher, you will find the attached questionnaire to be interesting and worthwhile. It is an initial attempt to estimate the proportion of accounting research affected by falsification. The *very brief* questionnaire assures your complete anonymity through the 'randomized response' technique (see references in the enclosed materials).

You were selected because of your substantial contribution to the research literature. You certainly understand the importance of a high response rate in this type of research, and we greatly appreciate your taking a couple of minutes to participate. Any comments would be welcome, too—and we'd be pleased to send you the early results.

Sincerely,

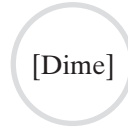
Instructions

The design of this questionnaire ensures your complete anonymity.

The enclosed 'Response Form' calls for a *single number* after you follow these steps:

Step 1

Before performing Step 2, please flip a coin three times. (A dime is attached for your convenience.) Note whether the coin comes up heads all three times. The purpose of these coin flips is to maintain your anonymity, so please don't disclose the outcome.



Step 2

If you do toss three heads, please respond to this question:

What percentage of published empirical articles in the top 30 Accounting journals* do you believe have been affected by at least one of the following?

If you do not toss three heads, please respond to this question:

What percentage of *your* published empirical articles have been affected by at least one of the following?

- Manipulation of data by fabricating it or inappropriately deleting/modifying observations.

OR

- Purposely reporting the results of an inappropriate statistical test because it gave the desired result, while ignoring the results of an appropriate test that gave different results.

OR

- Concealing an important validity problem (for example, an extraneous variable that you really believe drove your results).

OR

- Reporting that key procedures, such as randomization, were performed when they were not.

OR

- Otherwise intentionally reporting false research results.

*See back page of the enclosed 'FYI' sheet if you wish to see the list of journals.

Please answer on the Response Form.

Since you do not indicate the result of the coin flips, no one can know whether you have answered the question about yourself or about others. Drawing conclusions about any individual's response would be foolish (and unethical!).

Thanks very much for your time and cooperation.

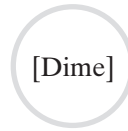
Instructions

The design of this questionnaire ensures your complete anonymity.

The enclosed 'Response Form' calls for a *single number* after you follow these steps:

Step 1

Before performing Step 2, please flip a coin three times. (A dime is attached for your convenience.) Note whether the coin comes up heads all three times. The purpose of these coin flips is to maintain your anonymity, so please don't disclose the outcome.



Step 2

If you <u>do</u> toss three heads, please respond to this question:

What percentage of *your* published empirical articles have been affected by at least one of the following?

↔ *Note reversal of questions in this version of the instrument*

If you do <u>not</u> toss three heads, please respond to this question:

What percentage of published empirical articles in the top 30 Accounting journals* do you believe have been affected by at least one of the following?

- Manipulation of data by fabricating it or inappropriately deleting/modifying observations.

OR

- Purposely reporting the results of an inappropriate statistical test because it gave the desired result, while ignoring the results of an appropriate test that gave different results.

OR

- Concealing an important validity problem (for example, an extraneous variable that you really believe drove your results).

OR

- Reporting that key procedures, such as randomization, were performed when they were not.

OR

- Otherwise intentionally reporting false research results.

*See back page of the enclosed 'FYI' sheet if you wish to see the list of journals.

Please answer on the Response Form.

Since you do not indicate the result of the coin flips, no one can know whether you have answered the question about yourself or about others. Drawing conclusions about any individual's response would be foolish (and unethical!).

Thanks very much for your time and cooperation.

Read Only If You Want
Further Information about this Research

If you or a colleague would like an early report of the results, please e-mail xxx@xxx.edu

or write to Prof. xxxxxxxxx
xxxxxx of Accounting
College of Business
University xxxxxxxx
xxxxxxxxxxxxxxxxx

Other Accounting research that has used the randomized response technique:

Berry, L. E., G. B. Harwood, and J. L. Katz, 'Performance of Auditing Procedures by Governmental Auditors,' *The Accounting Review*, 1987, v62(1), 14–28.

Buchman, T. A., 'The Reliability of Internal Auditors' Working Papers,' *Auditing: A Journal of Practice & Theory*, 1983, v3(1), 92–103

Buchman, T. A., and J. A. Tracy, 'Obtaining Responses To Sensitive Questions: Conventional Questionnaire Versus Randomized Response Technique,' *Journal of Accounting Research*, 1982, v20(1), 263–271.

Reckers, P. M. J., S. W. Wheeler, and B. Wong-On-Wing, 'A Comparative Examination of Auditor Premature Sign-Offs Using the Direct and Randomized Response Methods,' *Auditing: A Journal of Practice & Theory*, 1997, v16(1), 69–78.

Schneider, A., and N. Wilner, 'A Test Of Audit Deterrent To Financial Reporting Irregularities Using The Randomized Response Technique,' *The Accounting Review*, 1990, v65(3), 668–681.

The particular technique in this study is based on Greenberg *et al.*, 'Application of the Randomized Response Technique in Obtaining Quantitative Data,' *Journal of the American Statistical Association*, 1971, v66(334), 243–250.

The 'Top 30 Academic Accounting Journals' (Alphabetical)[1]

Abacus
Accounting and Business Research
Accounting and Finance
Accounting, Auditing and Accountability
Accounting Educators' Journal
Accounting Horizons
Accounting, Organizations and Society
The Accounting Review
Advances in Accounting
Advances in International Accounting
Advances in Taxation
Auditing: A Journal of Practice and Theory
Behavioral Research in Accounting
Contemporary Accounting Research
Critical Perspectives on Accounting
International Journal of Accounting Education and Research
Issues in Accounting Education
Journal of Accounting and Economics
Journal of Accounting and Public Policy
Journal of Accounting, Auditing and Finance
Journal of Accounting Education
Journal of Accounting Literature
Journal of Accounting Research
Journal of Business, Finance and Accounting
The Journal of Information Systems
Journal of Management Accounting Research
Journal of the American Taxation Association
National Tax Journal
Research in Accounting Regulation
Research in Governmental and Nonprofit Accounting

[1] Composite list based on five published studies. See J. R. Hasselback, A. Reinstein, and E. S. Schwan, 'Toward the Development of Benchmarks to Assess the Academic Research Productivity of Accounting Faculty.' Working paper presented at the 1996 Annual AAA Meeting.

Response Form

Response to question about the percentage of work affected: ___%

About how long have you been tenured? Not yet tenured ___(✔). Tenured about ___ years.

What do you believe would motivate researchers to falsify research? Please allocate 100 percentage points between the following two perceived causes of research fabrication:

Internal reasons such as desire for recognition___%
External reasons such as tenure pressure ___%
100%

Thank you very much! ◇

Note: The symbol in the lower right corner identified the four treatment groups for data coding.

APPENDIX B

GREENBERG *ET AL.*'S (1971, P. 245) EQUATIONS

The sample estimate of the mean of the sensitive distribution (in our study, the percentages of the respondent's *own* work actually affected by the misconduct) is

$$\hat{\mu}_A = \frac{(1 - p_2)\overline{Z}_1 - (1 - p_1\overline{Z}_2}{p_1 - p_2}$$

and the sample estimate of the mean of the nonsensitive distribution (in our study, the *overall body of literature* believed to be affected) is

$$\hat{\mu}_Y = \frac{p_2\overline{Z}_1 - p_1\overline{Z}_2}{p_2 - p_1}$$

with respective variances

$$V(\hat{\mu}_A) = [(1 - p_2)^2 V(\overline{Z}_1) + (1 - p_1)^2 V(\overline{Z}_2)] / (p_1 - p_2)^2,$$

and

$$V(\hat{\mu}_Y) = [p_2^2 V(\overline{Z}_1) + p_1^2 V(\overline{Z}_2)] / (p_2 - p_1)^2$$

where

p_i = probability that the sensitive question is selected by the respondent in sample i $(i = 1,2)$, $p_1 \neq p_2$,

$1 - p_i$ = probability that the nonsensitive question is selected by the respondent in sample i $(i = 1,2)$, and

Z_{ij} = response from individual j in sample i $(i = 1,2; j = 1,2, \ldots, n_i)$.