Why and *How* the ASA Ethical Guidelines Should Be Integrated into Every Quantitative Course

Rochelle E. Tractenberg, PhD, MPH, PhD, PStat®, FASA
Collaborative for Research on Outcomes and –Metrics; Departments of Neurology; Biostatistics, Bioinformatics & Biomathematics; and Rehabilitation Medicine; Georgetown University Medical Center, Washington, D.C.
Building D, Suite 207
Georgetown University Medical Center
4000 Reservoir Rd. NW
Washington, DC 20057

Abstract

Interest in the revised ASA Ethical Guidelines for Professional Practice is keen across the ASA membership and leadership, but as of the 2013-14 academic year, only 35% of US universities required any ethics content for *at least some* of their students in statistics and biostatistics programs. Since data analysis is becoming important across disciplines, the Guidelines can serve to introduce all students to critical concepts of responsible data analysis, interpretation, and reporting. The Guideline principles interact, and sometimes must be prioritized. Therefore, neither the simple distribution of –nor an encouragement to memorize- the Guidelines can promote the necessary level of awareness. The Guidelines contain elements that are suitable, and important, components of training for undergraduates and graduates whether or not they are statistics majors, to prepare them for ethical quantitative work. To achieve this preparation, and to support responsibility in the conduct of research involving data and its analysis, the Guidelines should be incorporated into *every* quantitative course. This paper discusses why and how this can be accomplished.

Key Words: ASA Ethical Guidelines; ethics education; curriculum development; ethical reasoning; statistical practice.

1. Introduction: ASA Ethical Guidelines go far beyond "ethics education" for statisticians-in-training

"Graduate instruction in statistics requires the presentation of general frameworks and how to reason from these." (Hubert & Wainer, 2011:62). This claim for two-part instruction is also true for "instruction in ethics for statistical practice", and the purpose of this paper is to articulate and support the argument that **instruction in statistics *also* requires the presentation of general *ethical* frameworks, and instruction and practice in how to reason from *these***. A common conceptualization of ethics education, or "training in the responsible conduct of research" (RCR) is to complete one general module or course – and these courses or sessions typically focus on (or emphasize) the appropriate treatment of human subjects (see Tractenberg, 2016-a, for discussion). Professional statisticians are responsible for maintaining a wholly ethical representation of any research in which they are involved. It may seem an unfair burden,

and that imbalance in responsibility is actually even more critical when the statistician in a project is not actually a "professional" because there is no other "professional" with the specific responsibility for ethical treatment of data and its analysis. Thus, these Guidelines are as important – and useful! - for non-professional data analysts and data scientists as they are for accredited and professional statisticians.

The ASA Ethical Guidelines (American Statistical Association, 2016) comprise 8 core principles, which entail 49 specific elements (See Appendix):
A. Professional Integrity & Accountability (6)
B. Integrity of data and methods (10)
C. Responsibilities to Science/Public/Funder/Client (5)
D. Responsibilities to Research Subjects (6)
E. Responsibilities to Research Team Colleagues (4)
F. Responsibilities to Other Statisticians or Statistics Practitioners (5)
G. Responsibilities Regarding Allegations of Misconduct (6)
H. Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners (7)

The number of specific elements (indicated in parentheses above) do not correspond to the importance or complexity of the core principles, but their number does suggest that memorizing this list of 49 elements is an unwieldy task. Instead, the objective of the Committee on Professional Ethics is that data analysts know and understand the core principles, and also know and understand that these can sometimes be in conflict within a single case or analysis. For this reason, the Committee does not advocate memorization but does advocate learning how to reason with, and prioritize, these core principles and their constituent elements for every person who engages in data analysis.

Many individuals without professional statistician accreditation (PStat®) or even comprehensive training in statistics are asked to carry out statistical analysis in business and research settings –and this is an increasingly common situation as software, applications, and a perceived need for data analysis become ever more ubiquitous. It is untenable to assume that the training and practice that are required for ethical statistical practice would be conveyed within the single, *general-institutional* training module in "responsible conduct of research" that universities in the United States are required to provide for individuals receiving federal funding (possibly even less so within the institutional "ethics" training that many businesses and companies require). It is also nonsensical to assume that one general training opportunity – particularly if it is completed early in a career - can promote the level of professional ethical conduct that the ASA Ethical Guidelines for Statistical Practice describe and promote. Instead, for undergraduate and graduate statistics students alike, if the ASA Ethical Guidelines were introduced early, and reinforced throughout a curriculum to promote an ongoing induction into the profession, this would be a driving force for ethics education that would be obviously and specifically relevant to both students *and* faculty in the discipline of statistics. However, trainees/students who are learning statistics from or for *other disciplinary perspectives* could also learn both how to engage in these same important conversations about the ethical dimensions of statistical research and practice – *and also that such conversations are important*. Preparing quantitative scientists to engage competently in these conversations requires purposeful, widespread, and developmental training that can come from, and support, a culture of ethical research and practice.

1.1 Typical ethics training is not focused on sustainable learning

In November, 2009 the National Institutes of Health (NIH, 2009) outlined rules stipulating that scientists proposing to train new scientists with federal funds in the United States must both document their RCR training plans for these trainees and also be able to describe how training plans change/have changed. There are no requirements –nor are there incentives or guidelines- to either develop or document the capacity to reliably train others in RCR or ethical practice. It is unlikely that the single, common, and static ethics course offered by most institutions can lead PhD scientists to develop new ethics training plans or courses (even if they use ethics education archives and resources). This context supports the recycling of static information even though ethical challenges can arise in new and wholly unexpected situations. Without a framework that promotes the ongoing development of abilities to identify and reason through ethical challenges, the actual utility of the typical ethics training that is available (or required) is intrinsically limited.

Moreover, "sustainable learning" is learning that can continue after, and in different contexts from specific training or educational experiences. When ethics training or education involves static presentation of facts to memorize, the sustainability is near zero – and the implicit message to the learners is that once the "test" is passed, the facts that were memorized can be forgotten to make room for other information that would actually be used/useful going forward. Training in ethics or "responsible conduct" in research that does not promote and encourage sustainable learning is unlikely to have the intended effects of that training – i.e., that the learning stays with the learner, and is applied and transferred in new settings whenever needed (Tractenberg et al., in review).

The Office of Research Integrity in the US Department of Health and Human Services have published an Introduction to The Responsible Conduct of Research (Steneck, 2007) where, in the introductory comments it is stated, "Research is not an organized profession in the same way as law or medicine… The norms for responsible conduct can vary from field to field." (p. xi). However, the American Statistical Association (ASA, 2016), Royal Statistical Society (RSS, 2014)), and International Statistics Institute (ISI, 2010) all have codes of conduct for statistical practice –i.e., "general frameworks" for ethical practice. As pointed out by Steneck (p. 7), "…most codes of ethics are simply general statements about ideals and do not contain the specific guidance researchers need to work responsibly in complex research settings." The ASA, RSS and/or ISI "general frameworks" – **and how to reason from these** – must be integrated into our preparation of quantitative scientists and data analysts for the profession as well as for *any engagement with data –irrespective of field.* This more dynamic perspective on ethical practice has concrete implications for training – the most obvious implication being that the current paradigm is not fit for purpose.

1.2 Reasoning from ethical frameworks ("ethical reasoning") is a learnable, improvable skill set.

"The entire community of scientists and engineers benefits from diverse, ongoing options to engage in conversations about the ethical dimensions of research and (practice)," (Kalichman, 2013: 13). My colleagues and I conceptualize "ethics education" as purposeful engagement in the development and growth of a set of six learnable, improvable types of knowledge, skills or abilities (KSAs): Prerequisite

knowledge; recognizing an ethical issue; identification of decision-making frameworks; identification and evaluation of alternative actions; making and justifying decisions; and reflecting on the decision (Tractenberg & FitzGerald, 2012; Tractenberg, et al, in review; see also Gunaratna & Tractenberg, 2016). This ethical reasoning KSA list is focused on decision-making and reasoning -and not on the mastery of information alone; and learning these KSAs does require – and promote – ongoing conversations about ethical dimensions of research. We have argued elsewhere (Tractenberg & FitzGerald, 2015; Tractenberg et al. 2015; Tractenberg, 2016-b) that all data analysis requires decision-making, and whether or not an individual self-identifies as a "statistician", quantitative analysis –even if it derives from automation or algorithmic pattern matching – also requires decision-making. This is one reason why ethical reasoning is so crucial for the training of all quantitative scientists. Introducing ethical reasoning in any undergraduate or graduate program, and **explicitly** integrating **acknowledgement of the necessity for** *ethical data practice* into all engagement with data (Big, big, or small), is important for improving the reproducibility of science across disciplines (see e.g. Freedman 2010; Collins & Tabak, 2014; McNutt 2014). Preparing all scientists – including those who utilize statistics or "big data" techniques - to engage competently in these conversations requires purposeful, widespread, and developmental training that can come from, and support, a culture of ethical research and practice.

In 1992, the Association of Computing Machinery (ACM) published revisions of its guidelines for professional and ethical practice (Association for Computing Machinery 1992), followed by an article outlining the principles and how to use the guidelines (Anderson et al. 1993). The article outlined the principles, and also included case studies to support student learning of these principles and also how to use them. Anderson et al. (1993) stated that "(c)ommitment to ethical professional conduct is expected of every voting, associate, and student member of ACM." In 2013, computer science curriculum guidelines (UG) were revised http://www.acm.org/education/CS2013-final-report.pdf and these new guidelines include a requirement for "core hours in the social issues and professional practice knowledge area" to help "to promote a greater understanding of the implications of social responsibility among students". And, "(c)urricula must prepare students for lifelong learning and must include professional practice (e.g., communication skills, teamwork, ethics) as components of the undergraduate experience." (p.21) Characteristics of graduates of these programs include "(c)ommitment to professional responsibility. Graduates should recognize the social, legal, ethical, and cultural issues inherent in the discipline of computing. They must further recognize that social, legal, and ethical standards vary internationally. They should be knowledgeable about the interplay of ethical issues, technical problems, and aesthetic values that play an important part in the development of computing systems. Practitioners must understand their individual and collective responsibility and the possible consequences of failure. They must understand their own limitations as well as the limitations of their tools." (p.25). These principles may in fact be (as Steneck (2007) argues), "simply general statements about ideals and do not contain the specific guidance researchers need to work responsibly in complex research setting" – but they go far beyond the typical topic list discussions that tend to dominate training in responsible conduct in research.

Our ethical reasoning paradigm: A) represents ongoing development in the ability to reason ethically, across the scientist's career, encouraging self-regulation by making developmental objectives and performance criteria explicit; B) promotes certification of journeyman- and master-level performance of ethical reasoning; and C) provides for

concrete evaluation and improvement of training opportunities and their consistency with learning objectives. Specifically, our developmental trajectory loosely follows the levels of craft guilds: novice, apprentice, journeyman and master (Tractenberg & FitzGerald, 2012 Table 1). This reasoning paradigm could be implemented within a department or program at any institution to *build and document capacity for faculty* to engage in ongoing ethical development (e.g., Tractenberg & FitzGerald 2015), and for some, to document their mentorship of others in this domain; and to *build and document capacity in students/new data scientists* for ethical reasoning and for ongoing development in these ethical reasoning KSAs (e.g., Tractenberg, 2013). Our recent work has shown *ethical reasoning* to be both learnable and improvable (e.g., Tractenberg et al, in review). It is extremely difficult to teach "ethics", but ethical reasoning has proven to be more evaluable for preparing future professionals who are enrolled in PhD programs. Thus, this ethical reasoning paradigm is an additional resource that can be used to integrate the ASA Ethical Guidelines into courses, sequences, and curricula. Moreover, in more data science/computational programs, the structural features that have been described here can be applied to the ACM guidelines as well, promoting a wider appreciation for the ethical obligations for participation in data analytic applications in academia, government, and industry.

Thus, the dominant paradigm for "training in ethical research" is unsuited to modern statistical and data practice; unfortunately, the creation and publication of professional codes are also self-limited in that *reasoning from these frameworks* has not generally been incorporated into their conceptualization or promulgation as even relevant, much less *required*. Moreover, while professional societies have articulated guidelines that reflect the mindset of expert statisticians relating to their professional practice, *none* has made any formal effort to date to ensure that all being prepared for entry into this profession are trained in the use of these guidelines – neither their content *nor* how to reason from them. The general "training in the responsible conduct of research" paradigm has been shown repeatedly to be ineffective – and it is probable that the failure of the traditional RCR paradigm arises from the failure to teach and promote "how to reason from" ethical principles (see Tractenberg, 2016-a). Moreover, the lack of formal integration of ASA Ethical Guidelines into the preparation of future quantitative professionals also means that there is *no training at all* in these core principles for those who may be learning some statistics but who are *not* (or do not consider themselves) "future quantitative professionals". This underscores the importance of integrating the ASA Guidelines into all quantitative courses – in a manner that supports an appreciation by all quantitative practitioners that the Guidelines exist, apply, and must be actively employed. In the next sections, justification of dimensions and features of this training ("why?" - Section 2) and methods/resources ("how?" -Section 3) are discussed.

2. Sustainable learning and ethical reasoning

If further support than the arguments above are needed, Tractenberg (2016-b) articulated four rationales that might be utilized to support an argument that training with the ASA Ethical Guidelines is important for training graduate and undergraduates in the quantitative sciences. These four objectives or rationales are:

A. Encouraging ethical conduct in (throughout) the practice of science, by pointing out how everyone on a research team has their specific role with its attendant obligations and priorities.

B. Promoting professionalism for all of the research team members, including analysts irrespective of their level of training in statistics.
C. Promoting the consideration, prior to the start of analyses, of the analyses and the qualifications of the analyst to plan, execute, and interpret them.
D. Engaging with principles of professional practice for statisticians, which can promote both appreciation for the statistician as a collaborating research team member and understanding how this team member is accountable and responsible for their work.

Simply articulating these arguments is not going to achieve the desired integration of ASA Ethical Guidelines. The level at or to which the integration is needed must also be considered. Table 1 features seven levels of increasingly active integration – from simply directing students to the Guidelines (by giving the URL, for example), to attaching the Guidelines to the course syllabus, all the way up to fully integrating the Guidelines throughout the entire curriculum. The extent to which each of the four objectives that might be articulated for integrating the Guidelines listed above can be achieved by each of these seven approaches to that integration is shown in Table 1.

Table 1: Options for integrating ASA Ethical Guidelines- and whether they achieve any teaching or learning objectives for doing so.

OBJECTIVES FOR INTEGRATING GUIDELINES:	Encouraging ethical conduct throughout the practice of science	Promoting professionalism for all participants in research/data analysis	Promoting consideration of analysis features/ requirements	Engaging with principles of professional statistical practice
OPTIONS FOR INTEGRATING ASA ETHICAL GUIDELINES: Direct students to the Guidelines website (ASA)				
Attach Guidelines to syllabus				
Discuss in one class meeting				x*
Integrate into existing course (i.e., discussion in at least 1/3 of meetings and assignments that are discussed in class during (some part) of 1/3 of meetings)			x	
Create stand-alone course	x	x	x	x
Integrate across courses in sequence/series (i.e., discussion in at least 1/3 of meetings and assignments that are discussed in class	xx	xx	xx	xx

during (some part) of 1/3 of meetings –for each course)				
Integrate throughout curriculum	xxx	xxx	xxx	xxx

Table 1 shows that directing students to the ASA Guidelines achieves none of the objectives that might motivate the integration of the Guidelines in the first place; attaching the Guidelines to a syllabus similarly achieves no objectives. If the Guidelines are discussed in one class meeting ("x*") there *may* be a chance of achieving one of the four objectives (engagement) – but this is only a *chance* (hence the asterisk) – and is limited to just one of the four objectives of an integration effort. Even integrating discussion of the Guidelines into 1/3 of class meetings can only really support one of the purposes one might have in integrating the Guidelines –and not very strongly ("x"). Creating a stand alone course does achieve all four objectives; but clearly, integrating the Guidelines across courses achieves these objectives more strongly ("xx") and not as strongly as integrating the guidelines fully across and throughout the entire curriculum ("xxx").

The four justifications for integrating the ASA Ethical Guidelines (Tractenberg, 2016-b and Table 1) may resonate with faculty or instructors who are quantitatively trained, but might not have as much persuasion for institutional officials with different training. Table 2 therefore contextualizes the methods, or levels, to which the integration might be considered with national standards for generating actionable evidence from teaching and learning outcomes articulation (National Institute for Learning Outcomes Assessment, NILOA, 2016).

Table 2. Alignment of Principles for documenting and improving assessment (NILOA, 2016; Tractenberg, in review) with methods for teaching the ASA Ethical Guidelines. *Table adapted from Tractenberg (in review) with permission*.

Approach to integration:	Direct students to Guidelines	Minimally engage with/teach Guidelines	Integrate and teach/promote Guidelines
Principles for Learning Outcomes generating actionable evidence:			
Develop/articulate specific actionable learning outcomes		Learning goals minimal	Specific actionable learning goals possible.
Connect learning goals with student work		Only possible if student work relating to Guidelines is assigned.	Specific alignment and engagement/ constructivist approaches by students possible.

Articulate learning outcomes collaboratively	Institution/instructors may collaborate on learning outcomes.	Institution, instructors and students may collaborate on learning outcomes.
Outcomes support assessment that generates actionable evidence	Only by (very specific) design.	Institutions and instructors see explicit alignment of curricular features (courses, assignments/work products) and can use this evidence to support or change the approach.
Outcomes are focused on *improvement*		The curriculum/series of courses must explicitly articulate expected growth and development.
Outcomes document learning and its extent	Learning goals minimal –so extent of documented learning will be, as well.	Instructors/institutions structure training/teaching to generate documentation of learning and the achievement of articulated learning objectives.
Outcomes provide evidence of quality of learning	Assessment opportunities that document the achievement and quality of learning *may be* developed.	A portfolio can be created articulating the extent and quality of learning across the entire (series) integration –over time.
Expectations are explicit in the outcomes		Institutional obligation to provide learning opportunities that can and do promote growth and development in the target KSAs.
Evidence from the outcomes is externally relevant	With minimal integration, outcomes' relevance will also be minimal.	Portfolios documenting learning outcomes throughout the curriculum can be used as evidence of readiness/qualification to proceed or engage.

Not surprisingly, Table 2 shows that simply directing students to the Guidelines, whether by providing the URL or by appending them to a course syllabus, is completely misaligned with NILOA criteria for generating evidence about learning that can be used by students, instructors, or the institution. Minimally engaging students with the Guidelines does provide some evidence that would be actionable, but naturally, the

fuller integration is completely consistent with the NILOA principles and generates evidence of learning that students, instructors, and the institution can see and use in decision making.

3. Resources you can use for decision-making about teaching and assessment

As has been discussed above, the Guidelines should be integrated into all quantitative training, but how exactly to do this can be challenging to conceptualize, much less implement.

Table 3 below is constructed from the perspective that the Guidelines will be integrated into at least 1/3 of class meetings for one course – and it is also relevant for all of the more-fully integrated options in Table 1 (stand-alone course; across a sequence of courses; throughout the curriculum). Whereas Table 1 shows how a choice to integrate the Guidelines can achieve the objectives for this integration, Table 3 shows how the actual integration itself can actually achieve these objectives – this depends on the assessments and assignments that are employed to teach and give practice with the Guidelines.

Table 3 is stratified by two straightforward ways for integrating consideration of the Guidelines: the first and most common approach is to use cases that can be obtained from many different online resources specifically created for teaching research ethics (e.g., the Online Ethics Center, (OEC http://www.onlineethics.org/CMS/profpractice/ppcases.aspx); the Markkula Center for Applied Ethics (https://www.scu.edu/ethics/ethics-resources/ethics-cases/); the International Dimensions of Ethics Education in Science and Engineering (http://www.umass.edu/sts/ethics/online/home.html); and the American Statistical Association's growing case archive (http://www.amstat.org/committees/ethics/cases.cfm). The second approach is to create cases – or simply structure a discussion about the Guidelines – around homework problems. Homework problem-derived cases can be augmented with features from cases that come from these existing resources, or they can simply be used as articulated in Tractenberg (2016 –b).

Table 3. Using learning objectives for integrating the ASA Ethical Guidelines to identify appropriate assessment/assignments for learning.

OBJECTIVES FOR INTEGRATING GUIDELINES:	Encouraging ethical conduct throughout the practice of science	Promoting professionalism for all participants in research/data analysis	Promoting consideration of analysis features/ requirements	Engaging with principles of professional statistical practice
ASSESSMENT/ ASSIGNMENT OPTIONS FOR INTEGRATING ASA ETHICAL GUIDELINES (AEG):				

USING CASES FROM ONLINE RESOURCES OR BOOKS, including the ASA Archive		
Essay to prompt (e.g., "identify at least one AEG principle and one element that are relevant – discuss their implications")	x	xx
Case study	depends	x
In-class discussion*	depends	depends
USING HOMEWORK PROBLEMS		
Essay to prompt (e.g., "identify at least one AEG principle and one element that are relevant – discuss their implications")	xx	xx
Case study	xx	depends
In-class discussion§	xx	depends

§ Discussions in class are facilitated when students have prepared something from which they can discuss. In class discussions can be used to go through the formal case analysis process and/or to give practice with feedback on what exactly is desired in the assignment, as well as serving as actual assessments.

Table 3 also shows three different types of assignments that can be used for any sort of materials: either an essay that students write in response to a prompt (e.g., "what would you do in order to ensure that ASA Ethical Guideline Principle A is applied in this example?" or "Describe biases that investigators, or those providing the data (to you or in the case study), might have. How might you identify bias, or determine if it is present?"); or a formal case analysis – which requires instruction, practice, and feedback in case analysis on top of instruction, practice, and feedback on the ASA Ethical Guidelines; or an in-class discussion. Formal instruction methods using cases – for science – can be adapted for use in teaching ethics or ethical reasoning (e.g., National Center for Case Study Teaching in Science http://sciencecases.lib.buffalo.edu/cs/collection/submit-case.asp). As noted in Table 3, in-class discussions benefit from both practice and from students having something already written, so instructors could assign and essay and then have students use their essays for the in-class discussion.

A crucial feature of Table 3 is that these assignments do not support the achievement of *two* of the four objectives for integrating the Guidelines. Encouraging ethical conduct and promoting professionalism are two important justifications for the integration of these Guidelines. Assignments that target these two objectives *specifically* –either case analyses, essays to prompts, or in class discussion (or all three) throughout the course, sequence, or curriculum are essential for achieving these two critical objectives (e.g., "Explain how the application of the Guidelines in this case encourages ethical conduct in research" or "discuss how the ASA Guidelines promote professionalism"). Individual assignments – as shown in Table 3 - are unlikely to have those desired effects unless they are specifically targeted.

4. Conclusions

"…(E)thics is not a vaccine that can be administered in one dose and have long lasting effects no matter how often, or in what conditions, the subject is exposed to the disease agent" (National Academy of Engineering and National Research Council, 2009 p. 34).

The American Statistical Association (ASA) established its Ethical Guidelines for Statistical Practice in 1981; they were first revised in 1999 and the third revision is in progress as of May 2015. In 2011, when the Professional Statistician Accreditation (PStat®) initiative was started, all applicants to receive the PStat® designation are required to read, and confirm they have read and comply with, the ASA Ethical Guideline. No other ASA member is required to even read them. Similarly, the Royal Statistical Society (RSS) has ethical guidelines; and only those receiving some kind of RSS accreditation (CStat and CSI) must confirm that they have read, and agree to adhere to, these guidelines (no other RSS member is required to read them). Finally, there is no requirement to read or use the content of the Ethics declaration published by the International Statistics Institute (ISI; ISI 2010) for membership in this body. The ASA Ethical Guidelines were originally created specifically for statisticians (e.g., ASA members), but in the current revision (ASA, 2016), the applicability of the Guidelines to any person who engages in data analysis was explicitly integrated. In 1992, the Association of Computing Machinery (ACM) published revisions of its guidelines for professional and ethical practice (Association for Computing Machinery, ACM 1992), followed by an article outlining the principles and how to use the guidelines (Anderson et al. 1993). Anderson et al. (1993) stated that "(c)ommitment to ethical professional conduct is expected of every voting, associate, and student member of ACM." The 2016 Revised ASA Ethical Guidelines can promote engagement with the CS and ACM commitments to ethical and professional conduct with respect to data; hopefully this article can promote interest in – and a commitment to- this integration *across* quantitative training.

A focus on ethical reasoning has yielded preliminary evidence that this training is sustained beyond the course (Tractenberg et al, in review). Linking ethics education to professional identity development for quantitative scientists/statisticians can begin to promote ethical and professional habits of practice throughout an entire training program. The fuller integration of the ASA Ethical Guidelines (as well as those of the ACM or RSS for example) into all quantitative training has the potential to transform "ethics education" from a requirement merely to be checked off a list into an integral part of professional identity formation for quantitative scientists. If it is true that "(g)raduate instruction in statistics requires the presentation of general frameworks and how to reason from these." (Hubert & Wainer, 2011:62), then *all* **instruction in statistics and quantitative science *also* requires the presentation of general *ethical* frameworks, and instruction and practice in how to reason from *these***. Promoting this two-part view of ethics education in the preparation of statisticians and quantitative scientists for ethical practice should lead to "long lasting effects" and a common culture of ethical research and practice. This is a desirable outcome in general, so integrating formal training that can boost awareness and understanding of the ethical considerations –and specific skills to identify these and reason the way to a satisfactory, defensible, decision – is fully justified for all training in quantitative methods, theory, or applications.

APPENDIX: ASA ETHICAL GUIDELINES – REVISED

Ethical Guidelines for Statistical Practice

Prepared by the Committee on Professional Ethics of the American Statistical Association

Approved by ASA Board April 2016

Purpose of the Guidelines

The American Statistical Association's Ethical Guidelines for Statistical Practice are intended to help statistics practitioners make decisions ethically. Additionally, the Ethical Guidelines aim to promote accountability by informing those who rely on statistical analysis of the standards that they should expect. The discipline of statistics links the capacity to observe with the ability to gather evidence and make decisions, providing a foundation for building a more informed society. Because society depends on informed judgments supported by statistical methods, all practitioners of statistics, regardless of training and occupation or job title, have an obligation to work in a professional, competent, and ethical manner and to discourage any type of professional and scientific misconduct.

Good statistical practice is fundamentally based on transparent assumptions, reproducible results, and valid interpretations. In some situations, Guideline principles may conflict, requiring individuals to prioritize principles according to context. However, in all cases, stakeholders have an obligation to act in good faith, to act in a manner that is consistent with these Guidelines, and to encourage others to do the same. Above all, professionalism in statistical practice presumes the goal of advancing knowledge while avoiding harm; using statistics in pursuit of unethical ends is inherently unethical.

The principles expressed here should guide both those whose primary occupation is statistics and those in all other disciplines who use statistical methods in their professional work. Therefore, throughout these Guidelines, the term "statistician" includes all practitioners of statistics and quantitative sciences, regardless of job title or field of degree, comprising statisticians at all levels of the profession and members of other professions who utilize and report statistical analyses and their implications.

A. Professional Integrity and Accountability

The ethical statistician uses methodology and data that are relevant and appropriate, without favoritism or prejudice, and in a manner intended to produce valid, interpretable, and reproducible results. The ethical statistician does not knowingly accept work for which he/she is not sufficiently qualified, is honest with the client about any limitation of expertise, and consults other statisticians when necessary or in doubt.

The ethical statistician:

1. Identifies and mitigates any preferences on the part of the investigators or data providers that might predetermine or influence the analyses/results.
2. Employs selection or sampling methods and analytic approaches appropriate and valid for the specific question to be addressed, so that results extend beyond the sample to a population relevant to the objectives with minimal error under reasonable assumptions.
3. Respects and acknowledges the contributions and intellectual property of others.
4. When establishing authorship order for posters, papers, and other scholarship, strives to make clear the basis for this order, if determined on grounds other than intellectual contribution.
5. Discloses conflicts of interest, financial and otherwise, and manages or resolves them according to established (institutional/regional/local) rules and laws.
6. Accepts full responsibility for his/her professional performance. Provides only expert testimony, written work, and oral presentations that he/she would be willing to have peer reviewed.

B. Integrity of data and methods

The ethical statistician is candid about any known or suspected limitations, defects, or biases in the data that may impact the integrity or reliability of the statistical analysis. Objective and valid interpretation of the results requires that the underlying analysis recognizes and acknowledges the degree of reliability and integrity of the data.

The ethical statistician:

1. Acknowledges statistical and substantive assumptions made in the execution and interpretation of any analysis. When reporting on the validity of data used, acknowledges data editing procedures, including any imputation and missing data mechanisms.
2. Reports the limitations of statistical inference and possible sources of error.
3. In publications, reports, or testimony, identifies who is responsible for the statistical work if it would not otherwise be apparent.
4. Reports the sources and assessed adequacy of the data; accounts for all data considered in a study and explains the sample(s) actually used.
5. Clearly and fully reports the steps taken to preserve data integrity and valid results.
6. Where appropriate, addresses potential confounding variables not included in the study.
7. In publications and reports, conveys the findings in ways that are both honest and meaningful to the user/reader. This includes tables, models, and graphics.
8. In publications or testimony, identifies the ultimate financial sponsor of the study, the stated purpose, and the intended use of the study results.
9. When reporting analyses of volunteer data or other data that may not be representative of a defined population, includes appropriate disclaimers and, if used, appropriate weighting.

10. To aid peer review and replication, shares the data used in the analyses whenever possible/allowable, and exercises due caution to protect proprietary and confidential data, including all data that might inappropriately reveal respondent identities.
11. Strives to promptly correct any errors discovered while producing the final report or after publication. As appropriate, disseminates the correction publicly or to others relying on the results.

C. Responsibilities to Science/Public/Funder/Client

The ethical statistician supports valid inferences, transparency, and good science in general, keeping the interests of the public, funder, client, or customer in mind (as well as professional colleagues, patients, the public, and the scientific community).

The ethical statistician:

1. To the extent possible, presents a client or employer with choices among valid alternative statistical approaches that may vary in scope, cost, or precision.
2. Strives to explain any expected adverse consequences of failure to follow through on an agreed-upon sampling or analytic plan.
3. Applies statistical sampling and analysis procedures scientifically, without predetermining the outcome.
4. Strives to make new statistical knowledge widely available to provide benefits to society at large and beyond his/her own scope of applications.
5. Understands and conforms to confidentiality requirements of data collection, release, and dissemination and any restrictions on its use established by the data provider (to the extent legally required), and protects use and disclosure of data accordingly. Guards privileged information of the employer, client, or funder.

D. Responsibilities to Research Subjects

The ethical statistician protects and respects the rights and interests of human and animal subjects at all stages of their involvement in a project. This includes respondents to the census or to surveys, those whose data are contained in administrative records, and subjects of physically or psychologically invasive research.

The ethical statistician:

1. Keeps informed about and adheres to applicable rules, approvals, and guidelines for the protection and welfare of human and animal subjects.
2. Strives to avoid the use of excessive or inadequate numbers of research subjects, and excessive risk to research subjects (in terms of health, welfare, privacy, and ownership of their own data), by making informed recommendations for study size.
3. Protects the privacy and confidentiality of research subjects and data concerning them, whether obtained from the subjects directly, other persons, or existing records. Anticipates and solicits approval for secondary and indirect uses of the data, including linkage to other data sets, when obtaining approvals from research

subjects, and obtains approvals appropriate to allow for peer review and independent replication of analyses.

4. Knows the legal limitations on privacy and confidentiality assurances and does not over-promise or assume legal privacy and confidentiality protections where they may not apply.
5. Considers whether appropriate research-subject approvals were obtained before participating in a study involving human beings or organizations, before analyzing data from such a study, and while reviewing manuscripts for publication or internal use. The statistician considers the treatment of research subjects (e.g., confidentiality agreements, expectations of privacy, notification, consent, etc.) when evaluating the appropriateness of the data source(s).
6. In contemplating whether to participate in an analysis of data from a particular source, refuses to do so if participating in the analysis could reasonably be interpreted by individuals who provided information as sanctioning a violation of their rights.
7. Recognizes that any statistical descriptions of groups may carry risks of stereotypes and stigmatization. Statisticians should contemplate, and be sensitive to, the manner in which information is framed so as to avoid disproportionate harms to vulnerable groups.

E. Responsibilities to Research Team Colleagues

Science and statistical practice are often conducted in teams made up of professionals with different professional standards. The statistician must know how to work ethically in this environment.

The ethical statistician:

1. Recognizes that other professions have standards and obligations, that research practices and standards can differ across disciplines, and that statisticians do not have obligations to standards of other professions that conflict with these Guidelines.

2. Ensures that all discussion and reporting of statistical design and analysis is consistent with these Guidelines.

3. Avoids compromising scientific validity for expediency.

4. Strives to promote transparency in design, execution, and reporting or presenting of all analyses.

F. Responsibilities to Other Statisticians or Statistics Practitioners

The practice of statistics requires consideration of the entire range of possible explanations for observed phenomena, and distinct observers drawing on their own

unique sets of experiences can arrive at different and potentially diverging judgments about the plausibility of different explanations. Even in adversarial settings, discourse tends to be most successful when statisticians treat one another with mutual respect and focus on scientific principles, methodology and the substance of data interpretations. Out of respect for fellow statistical practitioners, the ethical statistician:

1. Promotes sharing of data and methods as much as possible and as appropriate without compromising propriety. Makes documentation suitable for replicate analyses, metadata studies, and other research by qualified investigators.
2. Helps strengthen the work of others through appropriate peer review; in peer review, respects differences of opinion and assesses methods, not individuals. Strives to complete review assignments thoroughly, thoughtfully, and promptly.
3. Instills in students and non-statisticians an appreciation for the practical value of the concepts and methods they are learning or using.
4. Uses professional qualifications and contributions as the basis for decisions regarding statistical practitioners' hiring, firing, promotion, work assignments, publications and presentations, candidacy for offices and awards, funding or approval of research, and other professional matters.
5. Does not harass or discriminate.

G. Responsibilities Regarding Allegations of Misconduct

The ethical statistician understands the difference between questionable scientific practices and practices that constitute misconduct, avoids both, but knows how each should be handled.

The ethical statistician:

1. Avoids condoning or appearing to condone incompetent or unethical practices in statistical analysis.

2. Recognizes that differences of opinion and honest error do not constitute misconduct; they warrant discussion, but not accusation.

3. Knows the definitions of, and procedures relating to, misconduct. If involved in a misconduct investigation, follows prescribed procedures.

4. Maintains confidentiality during an investigation, but discloses the investigation results honestly to appropriate parties and stakeholders once they are available.

5. Following an investigation of misconduct, supports the appropriate efforts of all involved, including those reporting the possible scientific error or misconduct, to resume their careers in as normal a manner as possible.

6. Avoids, and acts to discourage, retaliation against or damage to the employability of those who responsibly call attention to possible scientific error or misconduct.

H. Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners

Those employing any person to analyze data are implicitly relying on the profession's reputation for objectivity. However, this creates an obligation on the part of the employer to understand and respect statisticians' obligation of objectivity.

Those employing statisticians are expected to:

1. Recognize that the Ethical Guidelines exist, and were instituted, for the protection and support of the statistician and the consumer alike.

2. Recognize that valid findings result from competent work in a moral environment. Employers, funders, or those who commission statistical analysis have an obligation to rely on the expertise and judgment of qualified statisticians for any data analysis. This obligation may be especially relevant in analyses that are known or anticipated to have tangible physical, financial, or psychological impacts.

3. Recognize that the results of valid statistical studies cannot be guaranteed to conform to the expectations or desires of those commissioning the study or the statistical practitioner(s).

4. Recognize that it is contrary to these Guidelines to report or follow only those results that conform to expectations without explicitly acknowledging competing findings and the basis for choices regarding which results to report, use, and/or cite.

5. Recognize that the inclusion of statistical practitioners as authors, or acknowledgement of their contributions to projects or publications, requires their explicit permission because it implies endorsement of the work.

6. Support sound statistical analysis and expose incompetent or corrupt statistical practice.

7. Strive to protect the professional freedom and responsibility of statistical practitioners who comply with these Guidelines.

Acknowledgements

RET was supported by an NSF EESE grant ("A multidimensional and dynamic ethics education paradigm spanning the science career"; award #1237590). RET was the Vice Chair of the ASA Committee on Professional Ethics, chaired the working group on the 2016 Guidelines revision, and becomes the Chair 1 January 2017; apart from these there are no actual or potential conflicts.

References

American Statistical Association (2016). Ethical guidelines for statistical practice - revised. Downloaded from http://www.amstat.org/committees/ethics/ on 15 June 2016.

Anderson RE, Johnson DG, Gotterbarn D, and Perrolle J. (1993). Using the New ACM Code of Ethics in Decision Making. Communications of the ACM 36:2: 98-107.

Association for Computing Machinery (1992). ACM Code of Ethics and Professional Conduct. http://www.acm.org/about/code-of-ethics Accessed 5 December 2013.

ACM/IEEE-CS Joint Task Force. Computer Science Curricula 2013 (2013). Downloaded from http://www.acm.org/education/CS2013-final-report.pdf 20 January 2015.

Collins FS & Tabak LA. (2014). Policy: NIH Plans to enhance reproducibility. *Nature* 505 (30 Jan 2014): 612-613. Downloaded from http://www.nature.com/news/policy-nih-plans-to-enhance-reproducibility-1.14586 on 10 March 2016

Freedman DH. (2010). Lies, damned lies, and medical science. The Atlantic November 2010 issue; downloaded from http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/308269/ on 7 November 2015.

McNutt M. (2014). Editorial: Journals unite for reproducibility. *Science* 7 November 2014: Vol. 346 no. 6210 p. 679. DOI: 10.1126/science.aaa1724

Gunaratna N & Tractenberg RE. (2016- in press). Ethical Reasoning with the 2016 Revised ASA Ethical Guidelines for Statistical Practice. Proceedings of the 2016 Joint Statistical Meetings, Chicago, IL.

Hubert L & Wainer H. (2011). A statistical guide for the ethically perplexed. In AT Panter & SK Sterba (Eds)., Handbook of ethics in quantitative methodology. New York, NY: Routledge. pp 61-124.

International Statistical Institute (2010). Declaration on Professional Ethics (revised). Downloaded from http://www.isi-web.org/about-isi/professional-ethics/43-about/about/296-declarationprofessionalethics-2010uk on 11 Dec 2013.

Kalichman M. (2013). Why teach research ethics? In, National Academy of Engineering (Eds). Practical Guidance on Science and Engineering Ethics Education for Instructors and Administrators. Washington, DC: National Academies Press. Pp. 5-16.

National Institute for Learning Outcomes Assessment (2016, May). Higher education quality: Why documenting learning matters. Urbana, IL: University of Illinois and Indiana University, Author.

National Institutes of Health (2009) Update on the Requirement for Instruction in the Responsible Conduct of Research. NOT-OD-10-019. http://grants1.nih.gov/grants/guide/notice-files/NOT-OD-10-019.html, accessed 25 January 2012.

Royal Statistical Society. (2014). Code of conduct. Downloaded from http://www.rss.org.uk/Images/PDF/join-us/RSS-Code-of-Conduct-2014.pdf 15 January 2015.

Steneck NH (2007). Introduction to the Responsible Conduct of Research – Revised edition. http://ori.dhhs.gov/documents/rcrintro.pdf accessed 1 January 2010. Washington, DC: US Department of Health and Human Services.

Tractenberg RE & FitzGerald KT. (2012). A Mastery Rubric for the design and evaluation of an institutional curriculum in the responsible conduct of research.

Assessment and Evaluation in Higher Education. 37(7-8): 1003-21. DOI 10.1080/02602938.2011.596923

Tractenberg, RE. (2013). Ethical Reasoning for Quantitative Scientists: A Mastery Rubric for Developmental Trajectories, Professional Identity, and Portfolios that Document Both. Proceedings of the 2013 Joint Statistical Meetings, Montreal, Quebec, Canada. Pp. 3959-3973.

Tractenberg RE. (2016-a). Creating a culture of ethics in Biomedical Big Data: adapting 'guidelines for professional practice' to promote ethical use and research practice. In, L Floridi & B Mittelstadt (Eds)., Ethics of Biomedical Big Data. London: Springer. 367-393.

Tractenberg RE. (2016-b). Institutionalizing ethical reasoning: integrating the ASA's Ethical Guidelines for Professional Practice into course, program, and curriculum. In, J. Collmann & S. Matei (Eds)., Ethical Reasoning in Big Data: An Exploratory Analysis. New York: Springer. 115-139.

Tractenberg RE. (in review). How the Mastery Rubric for Statistical Literacy can generate actionable evidence about statistical and quantitative learning outcomes.

Tractenberg RE & FitzGerald KT. (2015). Responsibility in the conduct of Quantitative Sciences: Preparing future practitioners and certifying professionals. (presented at 2014 Joint Statistical Meetings, Boston, MA). Proceedings of the 2015 Joint Statistical Meetings, Seattle, WA. Pp. 4296-4309.

Tractenberg RE, FitzGerald KT, Collmann, J. (in review). Evidence of sustainable learning with the Mastery Rubric for Ethical Reasoning.

Tractenberg RE, Russell A, Morgan G, FitzGerald KT, Collmann J, Vinsel L, Steinmann M, Dolling LM. (2015) Amplifying the reach and resonance of ethical codes of conduct through ethical reasoning: preparation of Big Data users for professional practice. Science and Engineering Ethics 21(6):1485-1507. http://link.springer.com/article/10.1007%2Fs11948-014-9613-1.