Ten Simple Rules for Integrating Ethics into Statistics and Data Science Instruction

Rochelle E. Tractenberg
Georgetown University
Building D, Suite 207, 4000 Reservoir Rd. NW
Washington, DC 20057 USA

Abstract
Ten simple rules for integrating ethics content/training in ethical practice into every/any statistics and data science course are presented. These rules are intended to support instructors who seek to encourage ethical conduct in (throughout) the practice of science, whether it involves statistics, data science, or qualitative analysis; as well as throughout the employment of tools, methods, and techniques from these domains. Truly integrated ethical training can also promote awareness of the fact that every member of a research – or practice - team has a specific role, with attendant obligations and priorities, relating to the use of statistics, data science, and qualitative analytic approaches. Even if individuals are not going to be the 'designated statistician/analyst' on a project, understanding the roles and responsibilities of team members can strengthen the sense of responsibility and accountability of each member of a science or practice team. True integration of ethical training is not simple to achieve, but the ten rules are based on educational and cognitive sciences, as well as a recognition of the fact that additional content, without furthering a course's existing learning objectives, greatly dampens enthusiasm for, and the likelihood of, integration of ethical training into quantitative courses. Assumptions for readers of these ten simple rules are: the instructor wants to have something that can be graded/evaluated after the students engage with the case; and that one objective the reader has is to teach how to reason & make decisions ethically as students go about practicing or using statistics. The overarching message of the ten rules is that true integration can benefit from leveraging existing structural features that both streamline learning outcomes and increase the chance of successfully embedding ethical practice standards into existing courses. Success is defined as the creation of reproducible, gradable work from students that signal whether or not the ethics instruction had its intended effects; and the documentation of ongoing (sustained) engagement with the ethics training beyond the end of the course.

Key Words: Ethical practice; statistics and data science; ASA Ethical Guidelines; teaching professional ethics; ethical statistics; ethical data science.

1. Introduction

According to the National Academies of Sciences, Engineering, and Medicine (NAS, 2018), "data science spans a broad(er) array of activities that involve applying principles for data collection, storage, integration, analysis, inference, communication, and ethics". (p. 1) This report, by the Committee on Envisioning the Data Science Discipline,

> "… underscores the centrality of studying the many ethical considerations that arise as workers engage in data science. These considerations include deciding

> what data to collect, obtaining permissions to use data, crediting the sources of data properly, validating the data's accuracy, taking steps to minimize bias, safeguarding the privacy of individuals referenced in the data, and using the data correctly and without alteration. It is important that students learn to recognize ethical issues and to apply a high ethical standard." (p. 2).

One of the 2018 NAS action items (p.3) follows directly from the Committee's focus on the centrality of ethical professional practice in data science:

> *Recommendation 2.4: Ethics is a topic that, given the nature of data science, students should learn and practice throughout their education. Academic institutions should ensure that ethics is woven into the data science curriculum from the beginning and throughout.*

To support instructors' achieving NAS Recommendation 2.4, I propose ten simple rules for integrating ethical training into every quantitative course:

1. Follow instructional and/or curriculum development guidelines.
2. Use professional ethical practice standards, not issues, to guide instruction/learning objectives.
3. Aim to *equip*, not solely instruct.
4. Leverage – and accommodate - the complexities of practice.
5. Promote the sustainability of learning.
6. Leverage, possibly reconsider/revise, existing assignments.
7. Prioritize observability and actionability of the target behaviors to be taught and learned (and demonstrated).
8. Assess progression towards fluency in terms of communication and self-assessment.
9. Plan to integrate, then evaluate. Be prepared to make additional, evidence-informed, changes that will optimize the chances of the learning outcomes being achieved by most, if not all, students.
10. Engage with colleagues - instructors in the same course or in later courses (or both!) – who are also interested in integrating ethical quantitative practice instruction, so that students will have the opportunities to continue learning, practicing, and becoming more proficient at ethical quantitative practice.

These rules follow from over 10 years of engagement in the training of scientists in ethical research practices involving data, the revision and dissemination of the ASA Ethical Guidelines for Statistical Practice (2014-2020), and over 25 years of engagement in preparing future faculty/supporting effective instruction in higher education. The ten rules are each discussed more fully below, with specific references from my own body of work -which in turn has extensive literature reviews that readers might find useful.

2. Ten Simple Rules

1. Follow instructional and/or curriculum development guidelines.

Curriculum is defined by the European Union education and training policy as the "*inventory of activities related to the design, organisation and planning of an education or training action, including definition of learning objectives, content, methods (including assessment) and material, as well as arrangements for training teachers and trainers*". The design of curriculum and instruction should be formal and systematic. Once the purpose and context of a curriculum has been determined, the design process can begin, including its content, and assessment and evaluation mechanisms, which will determine its composition and duration. In general, the curriculum should specify the knowledge, skills and abilities – expressed in the Learning Outcomes (LOs) – that are proposed to be achieved; how learners will demonstrate their achievement; the lessons and/or assignments the instructors will use to support learner progression towards the LOs; the materials to be used; and the assessments for evaluating student learning and the effectiveness of the instruction.

Curriculum and instructional guidelines were published in 2020 (Tractenberg et al. 2020) and they feature an established five-phase curriculum-design model (suitable also for instructional design, Nicholls, 2002):

1. Select or identify LOs;
2. Select or develop learning experiences that will help students achieve the LOs;
3. Select or identify content that is relevant to LOs;
4. Identify or develop assessments to ensure learner is progressing towards LOs;
5. Evaluate the effectiveness of the learning experiences for leading learners to the LOs.

Seven Guidelines for curriculum and instructional development are outlined in this 2020 paper to support the design and evaluation of successful curriculum and instruction:

A. Identify and follow a formal paradigm for curriculum or instructional design (e.g., the Nicholls five-phase model above).
B. Focus on LOs first, to inform all other decisions about the curriculum and instruction.
B1: Leverage LOs to explore and identify appropriate learning experiences.
B2: Leverage LOs to select content that is appropriate for the learning experiences and promotes the LOs.
B3: Assess learning based on achievement of LOs using formative and summative assessment, as appropriate.
C: Plan and execute an actionable evaluation of the curriculum and instruction.
D: Document and share the features of the curriculum or instruction – including criteria for their success – with learners.

All formal instructional and curriculum development guidelines will emphasize the articulation of learning outcomes as a first (or key) step (see Diamond, 2008; Fink 2013; Nilson 2013). Using a formal paradigm to develop ethics instruction for any quantitative class will facilitate the development of a curriculum, or of any instructional endeavor, by leveraging what is already known about education and learning, and providing structure for decision-making (Tractenberg et al. 2020, Guideline A). The first Guideline principle is to follow a reproducible, formal design paradigm. All formal instructional paradigms will focus on learning outcomes (LOs) as the principal driver of all later decisions – and that is reflected in principles B-C. Principle D reflects what is known about adult learners, but also promotes the engagement by instructors of the learner as a partner, rather than a vessel to be filled with knowledge (consistent with cognitive and educational psychological research).

This contrasts with the common experience of beginning with content or issues (or instructor availability). Beginning with an agreed-upon paradigm for development will ensure that all of the key aspects of instruction will be attended to (not just content), and that new instructional efforts can and will be evaluable. Individual instructors, programs, and institutions should seek to evaluate whether their new instruction achieved learning outcomes (National Institute for Learning Outcomes Assessment (NILOA), 2016) – and not whether or not new content was added to courses.

Adult learners are particularly sensitive to a need to understand why they are learning what is being taught. Moreover, their requirement for higher-complexity employment of what is learned is also usually much greater than it is for younger students. Recognizing and sharing prerequisites with the learners, encouraging them to understand the demands of the instruction and what they should expect as a result of completing it, are all important contributors to the strength and ultimate effectiveness of the instruction that is developed. Also, fulfilling Curriculum Development Guideline principle D can ensure that all instructors in a program facilitate learning across students in a coherent, reproducible and assessable way.

Instructional design - whether for courses or curricula - can be difficult. There are several problems that come up repeatedly in the design process, and these happen to be similar whether the instruction is short term/short form or longer form instruction. For example, specifying prerequisites; ensuring content coverage; managing time; assessing learning. When a formal approach to curriculum or instructional development is employed, then decisions can be more easily made reproducible across instructors or developers, and can also be better justified. Moreover, curriculum and instructional design should benefit from the decades of research that supports optimizing the learning that results from instruction – at whatever scale. The 2020 Guidelines for Curriculum and Instructional Design are motivated by these supportive features, and are theoretically and empirically supported by those decades of research in the education sciences.

2. Use professional ethical practice standards, not issues, to guide instruction/learning objectives.

Professionalism is defined as "the skill, good judgment, and polite behavior that is expected from a person who is trained to do a job well." And "the conduct, aims, or qualities that characterize or mark a profession or a professional person." (Merriam-Webster).

Professional Identity is defined as "…the sense of being a professional…the use of professional judgment and reasoning … critical self evaluation and self-directed learning..." (Paterson et al. 2002:7)

- "**Professional identity formation means becoming aware of … what values and interests shape decision-making**." (Trede, 2012:163)
- "(S)tudents could learn more from their experiences if they were more explicitly guided to look out for certain aspects of professionalism and given further opportunities to discuss and critique their observations and experiences." (Grace & Trede, 2011:12)

The American Statistical Association (ASA) Ethical Guidelines (GLs) for Statistical Practice (American Statistical Association, 2018) include 52 items under eight general areas:

A. Professional Integrity & Accountability (7)
B. Integrity of data and methods **<u>(11)</u>**

C. Responsibilities to Science/Public/Funder/Client (5)
D. Responsibilities to Research Subjects (7)
E. Responsibilities to Research Team Colleagues (4)
F. Responsibilities to Other Statisticians or Statistics Practitioners (4)
G. Responsibilities Regarding Allegations of Misconduct (6)
H. Responsibilities of Employers/Clients Employing Statistical Practitioners (8)

Those who use statistical methods, tools, and techniques have ethical obligations to the practice of statistics and data science (ASA GLs A and D), those who contribute data or from whom data are obtained (ASA GLs B and C), and to those with whom (ASA GLs A, E, and F) or *for* whom (ASA GLs C, D, E, F, and H) data and their analyses are used or applied. These principles highlight the relevance of the ASA Guidelines in any activities that involve **data** - whether it is big data or small data, and also, whether the user is the "statistician"/data analyst or the employer/client of that statistician or data analyst. The ASA Ethical Guidelines are very tightly aligned with (inclusive of – and going beyond) those for the Royal Statistical Society (last updated 2014) and International Statistics Institute (last updated 2010) (discussed fully in Tractenberg, 2020-A). The ASA Guidelines explicitly articulate that, **"(t)he 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... comprising statisticians at all levels of the profession and members of other professions who utilize and report statistical analyses and their implications."** The ASA Guidelines promote ethical science and ethical evidence-driven decisions, and are intended "to help statistics practitioners <including those who use and report statistical analyses and their implications> make decisions ethically and communicate their rationale (s) honestly".

Data science is a discipline that has emerged at the intersection of computing and statistics – two disciplines with long standing guidance for ethical practice that feature professional integrity and responsibility. The Association of Computing Machinery (ACM) revised its Code of Ethics in 2018 (ACM, 2018). Both practice standards represent the perspectives of experienced professionals in their respective domains, but both organizations explicitly state that the guidelines apply to – should be utilized by – all who employ the domain in their work, irrespective of job title or training/professional preparation. Therefore, awareness of the ASA Ethical Guidelines for Statistical Practice – which are explicitly supportive of ethical practices in research that uses data whether "big" or small should be augmented with similar understanding of ethical practices relating to computing. The ACM Code of Ethics was first published in 1992 and most recently revised in 2018. The ACM Code directs "computing professionals" to focus attention on the positive and negative effects of their decisions – and emphasizes decisions that avoid or minimize harms. In the preamble, it is stated that computing professionals act responsibly when "consistently supporting the public good". There are four main areas with 24 specific points for what "a computing professional should" do:

1. General ethical principles (7)
2. Professional responsibilities (9)
3. Professional leadership principles (7)
4. Compliance with the code (2)

Like the ASA Guidelines, the ACM Code seeks to support ethical decision making, rather than describe what is wrong with aspects of practice. Given that both statistics and computing are essential foundations for data science, their ethical guidance should therefore be a starting point for the community as it contemplates what "ethical statistics and data science" looks like. The ASA Ethical Guidelines for Statistical Practice and ACM Code of Ethics are consensus-based guidelines that are updated regularly; they essentially reflect material that all who teach and train in statistics and data science can use as a common core for ethical practice standards (Tractenberg et al. 2015; Tractenberg, 2020-A). Rather than use "interesting" or "compelling" issues to drive what becomes a highly-variable patchwork of "ethics instruction", instructors can benefit from/leverage the ethical practice standards – and identify or create cases that help teach these standards – to bring a less variable, more reproducible, and potentially more sustainable dimension of ethical practice training to statistics and data science courses for all those who would use the tools, techniques, and methods from these disciplines (Tractenberg, in review-A).

Following **Rule 1**, integration will feature realistic and authentic learning outcomes – the target of instruction – that will help the students move towards satisfying the expectations of "a person who is trained to do a job well" (following **Rule 2**), rather than satisfying a requirement of "awareness of ethical practice standards". As articulated by both the ASA and ACM, users of statistics and computing need to understand their responsibilities to practice with these bodies of disciplinary knowledge ethically. If the learning outcomes of integrated ethics within a statistics and data science course feature the ASA Ethical Guidelines, then the result will be new *ethical practitioners*, rather than new statisticians who completed a required ethics course or mastered "ethics content".

Ethical aspects of practice by non-statisticians as well as statisticians are always relevant, and should be normalized. Recognizing that learning about ethical practice is challenging, perhaps especially for the non-practitioner, is the first step to enabling ethical practice just as consistently as we teach actual practical skills. It is every practitioner's responsibility (whether ASA/ACM member or not, whether job-titled-statistician or not) to teach *ethical* statistical practice when they teach statistical practice. The ASA Ethical Guidelines for Statistical Practice represent a formal structure to promote "the conduct, aims, or qualities that characterize or mark *our* profession or professional persons *in our discipline"*. (italics added). These professional practice standards are maintained and updated over time as the profession changes. ASA Guidelines (2018) and the ACM Code of Ethics (2018) support professionalism, *viz* "…the skill, good judgment, and polite behavior that is expected from a person who is trained to do a job well." Failing to inform biomedical researchers who use/learn to use statistics and data science – at all levels - of their obligations to practice ethically does not make the responsibilities go away, but it does ensure that fewer practitioners are able to take responsibility for ethical practice.

3. Aim to equip, not solely instruct.

Knowing that there are Guidelines, and even what they contain, is not enough (see **Rule 7**). Because different situations require different principles and elements of the Guidelines and Code, Ethical Reasoning is an important skill set that can be learned and improved. Beginning this learning and improving should start as early as possible, and ethical reasoning be brought to bear on situations where the Guidelines may be useful, as well as any others not specifically related to statistics and data science.

Ethical reasoning is actually its own set of knowledge, skills, and abilities (**KSA**s (Santa Clara University (no date); Tractenberg & FitzGerald, 2012; Tractenberg et al. 2017)). These KSAs are learnable and improvable, and can be deployed to ensure ethical practice (when there is no/before there is an ethical problem about which a decision has to be made) as well as when a decision about what to do (ethically) is required. Thus, learning to reason ethically – rather than "learning the Ethical Guidelines and/or Code of Ethics" – will promote "…the skill, good judgment, and polite behavior that is expected from a person who is trained to do a job well" more generally, and more universally (see Rios et al. 2019). When an ethical challenge arises, it must be recognized, and a decision must be made. To *make* and then support the decision, one must:

1. Identify/ assess your prerequisite knowledge – use ASA GLs/ACM CE
2. Recognize an ethical issue (and that a decision must be made) –GLs/CE help pinpoint
3. Identify relevant decision-making frameworks (e.g., virtue, like ASA or utilitarianism, like ACM)
4. Identify and evaluate alternative actions
5. Make & justify a decision
6. Reflect on the decision

Note: each of these Ethical Reasoning steps uses an observable – assessable – verb. Specifically, the ER KSAs can be integrated into a course (with the ASA and ACM practice standards) as follows:

1. Identify and 'quantify' your prerequisite knowledge: what elements of the ASA GLs /ACM EC, and stakeholder analysis (Tractenberg, 2019) contribute (and how) to an activity or task (see **Rule 4**).
2. Identify decision-making frameworks. The ASA GLs and ACM EC represent a "virtue ethics" decision making framework, which can generally be summarized as, "what would the ethical practitioner do in this case?" Direct appeal to the GL/CE facilitates making decisions that are consistent with "the ethical practitioner", but only when there is an obvious match between the GL/CE and features of the case. When the guidance of the GL/CE is less clear, another framework can be used: The utilitarian framework. The utilitarian perspective can help sort out the positive and negative effects that a decision can have on stakeholders. This perspective can generally be summarized as, "how can benefits be maximized while harms are minimized?"
3. Identify or recognize the ethical issue. What, in a case, is inconsistent with the GLs or EC? What seems "questionable"? Does something about the vignette represent an undue imbalance in the harms and benefits identified in the stakeholder analysis?
4. Identify and evaluate alternative actions (on the ethical issue). Every ethics case analysis requires that a decision be made. At least three plausible alternative options (decisions, actions) can be made in any circumstance:

a) do nothing.

b) consult or confer with a peer or a supervisor – using the professional guidelines or other resources (e.g., institutional policy) about what to do.

c) report violations of policy, procedure, ethical guidelines, or law.

In each case we consider whether "do nothing" is consistent with the ethical practice standards, but we also explore two variations on "do something" – either consult with others as to how to "do something" or, in some cases, "what <something> to do". In some cases, the ethical issue is clear – and there is a reporting mechanism available, so this is a plausible option ("report"). However, simply notifying someone (in authority) that someone has done something unethical will not necessarily address the harms that the unethical behavior has created, so there may be other actions that are also needed. So, for every case we consider these three options as a starting point, but the case analyses discuss how to determine plausible responses to unethical behaviors.

5. Make and justify a decision. Articulating the decision –what to do in the face of the ethical challenge that was identified, including stakeholder effects.
6. Reflect on the decision. What makes this case "hard"? What additional information would be/would have been helpful? How does this case highlight rigor, reproducibility and responsibility in data science research?

Note that these KSAs – learnable and improvable as they are – require much more than the memorization or awareness of "ethics" or even the Ethical Guidelines for Statistical Practice/Code of Ethics (see **Rule 7**). For those instructors who seek to teach their students how to reason & make decisions ethically as they go about practicing or using statistics and data science, the cognitive sophistication that ethical reasoning requires demonstrates that "learning the guidelines" is insufficient. Practice, with feedback, is an essential component of the development of ethical reasoning; learning outcomes should recognize this.

4. Leverage – and accommodate - the complexities of practice.

There are seven tasks that all statistics and data science can/do follow/recognize (Tractenberg in review A, in review B):
1. plan/design
2. collect/munge/wrangle data
3. analysis – literal for statistics & data science, "evaluation" for computing
4. interpret – always for statistics & data science, never for computing
5. document
6. report & communicate
7. work on a team

These tasks are essential in the practice of statistics and data science, and the ASA Ethical Guidelines pertain in each of these tasks. By focusing on one task at a time, an authentic, but focused, case can be created. If there is an existing case, then the Guidelines can be used to make a decision about what to do (i.e., the 'solution'). The key is to focus on *the learning* so that it, and assessment of whether it happened, will be objective and reproducible. With structure provided by Curriculum Development Guidelines (specifically, learning outcomes, **Rule 1**), these tasks, and the Ethical Guidelines and/or Code of Ethics (**Rule 2**), ethical reasoning KSAs can be targeted, better integrated/taught, and better assessed in any quantitative course. Instruction and practice with utilizing the Guidelines/Code can start at superficial, and then move on to deeper levels:

Superficial – look for words in the Guidelines that match what you want the vignette to help you teach (i.e., keep your LO in mind, construct a low-Bloom's activity around Guideline elements).

Deeper – the overall Guideline principles may also help you focus the case analysis, and what you think a meaningful analysis can contribute to a class, or an assignment (i.e., with a LO involving higher-Bloom's level performance, construct an analysis or justification type activity around Guideline elements that might conflict or require prioritization).

Asking for one "more obvious" and one "less obvious" Guideline Principle or element, with a brief narrative about what makes them obvious (more and less), is a more complex assignment than one asking for more superficial word- or idea-matching. This can be assigned (and/or discussed) as students learn each of these tasks. The level of sophistication will remain the same, but as students see the Guidelines/Code of Ethics repeatedly, but in the context of different aspects of statistical and data science practice, the embedding of this material will be deeper and the likelihood of establishing mental habits where these professional practice standards are utilized increases. Asking students to formulate an argument for one decision over another is a *much* more complex task, and it is essential to ensure that instruction supports increasing complexity of thinking and reasoning, and assessments also propel the student to demonstrate this increasing complexity over time. Since argument and justification with the ethical reasoning KSAs can be repeated with different material and different expectations as the students learn and improve their ethical reasoning abilities, the learning outcomes should match this progression, but need not require all new material at each stage of progression. The same prompt can be assigned with increasing requirement of sophistication in the response as instruction promotes this increase over time -and across these seven tasks.

If an instructor has an assignment that has neither a superficial nor a deeper match to the Guidelines, tweak the vignette – keeping the LOs in mind – so that the Guidelines are featured. Remember, the purpose of using cases or case analysis is not to "share exciting stories about statistics and data science", it is to provide authentic opportunities for new practitioners to learn how to make ethical decisions about statistical and quantitative practice. This is an important instructional objective for students in statistics, data science, and the research, government, *and business* contexts.

5. Promote the sustainability of learning.

"Sustainable learning" is defined as learning that continues beyond the end of formal instruction (Knapper, 2006; Schwänke, 2009); related to "transfer", which is defined as the application of learned skills or knowledge <u>from</u> the learned-in context <u>to</u> other contexts (Barnett & Ceci, 2002). Clearly, whatever we teach, we hope our students will be able (enabled) to utilize as they move forward to other courses and into their professional lives. This is particularly important for key attributes, such as a sense of professional identity, and the acknowledgement and acceptance of responsibility to use statistics, data science, and computing technologies and methods ethically. It would be ideal if instruction in, and opportunities to develop independence in, ethical practice was explicitly and coherently woven throughout all curricula, degrees, and programs. However, any individual instructor can more realistically strive for sustainability in the learning that occurs in their course/class.

Schwänke (2009) identified four dimensions to sustainable learning:

- *Lifelong learning*: an additional level of depth, or dimension, that you bring to a course or experience unrelated to the (primary) topic.
- *Changing your learning behaviour as a result of (some) specific learning*
- *A process of personal development* continuing beyond the (instruction)
- Deconstruction of incorrect conceptualization/reconstruction w/correct information.

Recognizing that *metacognition* is a known, learnable/improvable entity (National Research Council, 1999; National Research Council, 2001; see **Rule 8**), some colleagues and I (teaching ethical reasoning to graduate students in biological/biomedical programs outside of our departments) wondered if students could perceive *sustainability* –when we explicitly did *not* teach it. We never mentioned the words "transfer" (Barnett & Ceci 2002) or "sustainability" in the course. The course did feature "metacognition" explicitly, and we emphasized that this is a learnable, improvable set of knowledge, skills, and abilities (KSAs) in our semester long graduate course on ethical reasoning. The course and this study/these results are reported in Tractenberg et al. (2017).

Between 3-24 months after the course ended for our master's and doctoral students (none of whom were cognitive science students, nor were they enrolled in our programs or departments), we sent a four-item survey asking a) whether or not (Yes/No) the individual experienced each of the 4 dimensions of sustainability; and b) whether or not they believed it was due to the course. We also asked students to describe some evidence of how they recognized the dimension (because this is explicitly the type of metacognitive reasoning we taught/asked them to practice and develop all semester long). All nine (of 12) respondents perceived (yes) each of the four elements of sustainability. All respondents but one responded that they recognized that feature of sustainability as originating in this course specifically. The one time this did not happen, the student noted that they were not sure where that feature of sustainability (deconstruction/reconstruction) actually became available to them. Most importantly, though, is that every student was able to provide evidence from their own (current) work of each of the four dimensions of sustainable learning. Thus, it is feasible to conceptualize the training, instruction, and practice that one intends to provide with ethical reasoning and the ethical practice standards as the start of sustainable learning. In the 2017 paper discussed above, we did not mention any of the four features that were included in our survey (see Appendix); instructors will have a greater chance of succeeding at an objective to initiate *sustainable* learning if they are explicit about their intentions. (see also **Rule 8**). Metacognition is an essential aspect of lifelong learning (National Research Council, 2001; see also Ambrose et al. 2010).

6. Leverage, possibly reconsider/revise, existing assignments.

This list of ten rules is (clearly!) not simple. In fact, these represent over 20 years of my own work with cognitive science, educational psychology, and adult learning as I have tried to improve the sustainability, and effectiveness of learning in statistics and statistical literacy as well as in ethical practice (in research and statistics and data science). While this list may seem to create an impossible burden or make it seem like all instructors need to start over, what it is intended to do is offer a few different lenses through which each instructor can examine the materials they use to teach, and to assess learning, already. Anyone using this list of rules should sit down with their current syllabus and -applying **Rule 1**, consider the learning outcomes (Nicholls' Phase 1) and then determining how best to ensure that the teaching/in class activities (learning experiences), content, and assessments are specifically coherent with those outcomes.

Tractenberg (2020-B) describes an Assessment Evaluation Rubric (AER). The AER summarizes can be utilized to guide the development and evaluation/revision of assessments that are already used. The AER is structured around four features of an assessment (test, quiz, or assignment): its general alignment with learning goal(s); whether the assessment is intended to/effective as formative or summative; whether some systematic approach to cognitive complexity is reflected; and whether the assessment (instructions as well as results) itself is clearly interpretable. Each of the four AER dimensions (alignment; utility; complexity; clarity) has four questions that can be answered as present/absent (or yes/no) – identifying 'gaps' (nos) that need to be re-considered by the instructor. An alternative way to use the AER dimensions and questions is to evaluate another instructor's work with something like a three-level ordinal scale describing "present-useable", "possibly present - needs clarification", and "absent". Having someone other than the developer of the assessment to provide these ratings gives important information about the likelihood that students are able to perceive these features, and having that rater be experienced with/knowledgeable about cognitive and educational sciences increases the likelihood that actionable reviews will be made available. *Actionable* evaluation means that a decision or action will be supported by the results of the evaluation (National Institute on Learning Outcomes Assessment, 2016). Any instructor can use the AER to evaluate their own assessments and ensure that existing, or to-be-developed assessments, will promote learning and learner centered teaching.

If existing materials are evaluated with the AER, and opportunities to strengthen the four features in the AER are identified, then this would enable the 'strengthening' to incorporate ethical reasoning, professional practice standards, and others of the Rules suggested in this paper.

7. Prioritize observability and actionability of the target behaviors to be taught and learned (and demonstrated).

Samuel Messick (1994) outlined three questions that can help us develop valid assessments, where "validity" is defined as enabling decision-making that is based on the results of the assessment (Messick, 1989):

1. What are the knowledge, skills, and abilities (KSAs) the instruction or curriculum should lead to? (Messick #1)
2. What actions/behaviours by the students will reveal these KSAs? (Messick #2) And
3. What tasks will elicit these specific actions or behaviours (that reveal KSAs) (Messick #3)?

These three questions are important because they focus attention on both what our learning outcomes are (the KSAs the instruction should lead to), and on the fact that only observable behaviors that we specifically ask students to do in our assessments can indicate whether or not the learning objectives were achieved (by any given student). "Whatever tasks students are asked to carry out in the assessment (i.e., Messick #3) should have some relationship with observable verbs - because to accomplish Messick #2 the actions that students do to demonstrate they have learned what was intended must be observable (and, although it is by no means the only source, verbs from Bloom's Taxonomy (Bloom et al. 1956) work for this; see also Moseley et al. (2005) for other cognitive frameworks besides Bloom's). If instructors keep the three Messick questions in mind as they develop any assessment, then it will be more likely to achieve/the assessment will satisfy all of the features in the alignment dimension." (Tractenberg, 2020-B; p. 7).

There are six levels of complexity (from Bloom's 1 "B1" to Bloom's 6 "B6") in the original Taxonomy: **B1** Remember/Reiterate (performance is based on recognition of previously seen example); **B2** Understand/Summarize (performance summarizes info already in question (and/or answers)); **B3** Apply/Illustrate (performance extrapolates from seen examples to (really) new examples); **B4** Analyze/Predict (performance requires application/following of a rule; organize/sort using concrete criteria; some use of judgement in conjunction with the rules is needed); **B5** Create/ Synthesize (find patterns, innovate – to fill gaps in knowledge); and **B6** Evaluate/Compare/Judge (make judgments in the absence of "truth" or concrete criteria). The *minimum* level of performance to strive for, for ethical practice, is **B4**, as shown in Figure 1.

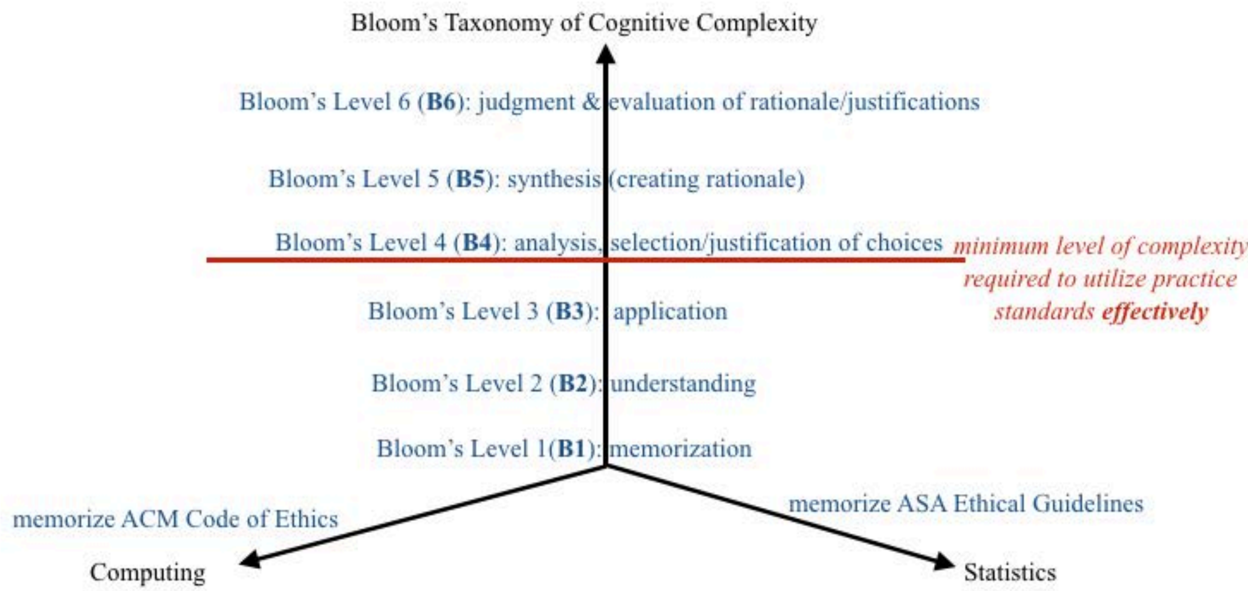

Figure 1: Bloom's cognitive complexity level 4 is the *minimum* required to apply the judgment needed to utilize ethical guidance. Memorization is not sufficient.

The ability to explain/justify why one or another ethical guideline or principle pertains, or should be prioritized over other considerations, is the minimum required (Briggle & Mitcham, 2012; see Gunaratna & Tractenberg 2016 for examples) for engagement in "ethical data science practice"; similar complexity arises from assessing the potential harms and benefits for stakeholders -especially when these conflict (Tractenberg, 2019). Students can be asked whether they agree that both ASA and ACM offer guidance (a recognition, B1-level, task) or to compare the decisions that the guidance documents offer (an evaluation, B6 level task), as appropriate for their level in any training or university course setting. In recognition that *judgement*, rather than memorization, is essential to ethical practice and ethical reasoning in data science, Bloom's taxonomy of cognitive behaviors[46] can and should be leveraged: students need to work with both the KSAs of ethical reasoning, and with the ethical practice standards, at low Bloom's levels so that some familiarity with these key elements of ethical practice are remembered, understood, and their applicability is attempted (B1-B3).

Increasing the Bloom's complexity requirement can promote both deeper processing (increasing the likelihood of long term remembering) and the potential for constructing a schema that will accommodate increasing sophistication in later training and after work experience. A schema is defined as "… the process of interpreting sensory data … in retrieving information from memory, in organizing actions, in determining goals and sub goals, in allocating resources, and, generally, in guiding the flow of processing in the system" (Rummelhart, 1980). Expecting students to see/memorize Guidelines and then be able to "guide the flow of processing" that leads to defensible decisions about ethical practice is unreasonable. Leveraging the fact that Bloom's taxonomy is a hierarchy (Tractenberg, 2017) and organizing instruction to exploit this, will better support the development of a schema for ethical practice in statistics and data science that will remain relevant for, and useful by, students as they move forward in their training and professional lives.

8. Assess progression towards fluency in terms of communication and self-assessment.

Quantitative courses typically do not utilize narrative, essay, or other fluency-based assessments. They also do not feature (teach/practice) self-assessment and the development of *metacognition*, defined as "the process of reflecting on and directing one's own thinking." (National Research Council, 2001: 78). A focus on metacognitive development will help to achieve instructional objectives at higher Bloom's levels while also strengthening the field (in which statistics and data science is being taught) by training creative and critical thinkers. Importantly, **"critical thinking" and the mechanics of metacognition require the application of Bloom's levels 4-6 to one's own thinking.** Thus, students need to learn the higher-level Bloom's thinking *and also* learn to apply those complex thinking skills to their own learning. Metacognition needs to be taught and practiced. The National Research Council (1999) summarized the literature supporting their recommendation that:

> A "metacognitive" approach to instruction can help students learn to take control of their own learning by defining learning goals and monitoring their progress in achieving them. (p. 14)

They also noted,

> Integration of metacognitive instruction with discipline-based learning can enhance student achievement and develop in students the ability to learn independently. It should be consciously incorporated into curricula across disciplines and age levels. (p. 17)

Attention to fluency in communication and self-assessment can engage students in documenting their own achievements - and charting new learning goals for themselves. These can be important within a course, as students identify where their study or learning habits do and do not work (see, e.g., Ambrose et al. 2010). Ethical reasoning KSAs, or analyses that employ these KSAs, will be narrative and training quantitative students to both communicate statistics and data science ideas, analyses, and results fluently will benefit from the integration of opportunities to self-assess, and communicate about their ethical practice.

Case analysis is a well-known approach to teaching ethics, however, most ethics education for non-ethics majors has also been shown repeatedly to not change unethical behaviors in research contexts. Dow et al. (2015) evaluated the theory and empirical work supporting case analyses for instruction in (professional) ethics; case analyses involve *active learning* (Handelsman et al. 2007), known to be an effective pedagogy, and require reasoning, which should be the target for ethics education (Briggle & Mitcham, 2012; National Academy of Engineering, 2013). Integrating authentic cases (Dow et al. 2015; Gunaratna & Tractenberg, 2016; Tractenberg, in review-B) together with ethical reasoning with professional practice guidelines unifies the pedagogic strengths of the case study method with authenticity that andradogic principles (Knowles et al. 2006) proscribe for effective training in "ethical data science". However, case analysis requires instruction and practice with feedback for students who are not familiar with this activity. Instructors may need some training in how to utilize this approach in class as well as in out-of-class assignments. Leveraging case analysis as a teaching approach – but with clear learning outcomes that feature ethical reasoning (following a learnable, improvable set of KSAs) and concrete ethical practice standards – could lead to improvements in focal areas like communication, self-assessment and metacognition, *and* in the abilities to reason through ethical challenges using the professional practice standards maintained by organizations like the ASA and ACM.

9. Plan to integrate, then evaluate. Be prepared to make additional, evidence-informed, changes that will optimize the chances of the learning outcomes being achieved by most, if not all, students.

Evaluation can focus on whether or not learning outcomes were achieved. Actionable outcomes mean that specific decisions about what to change can be supported depending on the results of the evaluation (NILOA, 2016; Hutchings et al. 2015). Formal evaluation – not just the 5-level Likert satisfaction and "compared to other instructors" ratings that educational institutions typically utilize, which usually offer no information whatsoever about what the instructor can or should change to improve the likelihood of achievement of learning outcomes – is an essential feature of a formal instructional design paradigm/effort. This is yet another reason why the Curriculum Development Guideline A is to choose the paradigm you will follow: the paradigm will include *actionable evaluation*. If you fail to choose/follow a paradigm, evaluation is unlikely - and actionable evaluation is *highly* unlikely. The Curriculum Development Guidelines (Tractenberg et al. 2020) include an emphasis on evaluation (Guideline C) as both an end in itself when designing new instruction (or revising existing courses/syllabi or programs of study) and as a way to ensure that learning objectives are, in fact, achievable. Planning to evaluate changes that one makes to instruction drives decision making towards what is observable and reproducible – while also allowing instructors to model what the education and learning sciences have empirically demonstrated over decades about teaching for learning. Planning an evaluation ensures that documentation of changes, justification of new content, teaching approaches, or learning objectives, and other aspects of instructional development are in place before the new learning opportunities are offered. This also enables sharing decision making with stakeholders (including the institution and program, other instructors, and the learners). This level of documentation can make it easier to follow **Rule 10**, and engage with colleagues at your institution and beyond.

10. Engage with colleagues - instructors in the same course or in later courses – who are also interested in integrating ethical quantitative practice instruction, so that students will have the opportunities to continue learning, practicing, and becoming more proficient at ethical quantitative practice.

Determination that the learning in a course was sustained/sustainable can be accomplished with a follow-up survey of the students. This determination of sustainability can be part of the formal course or instructional evaluation. A sustainability survey was published in 2017 (Tractenberg et al.) that can be adapted for users of these 10 rues. The survey used in Tractenberg et al. (2017), an adapted version of which appears in the Appendix, features ethical reasoning because the survey was developed for, and used in, a graduate level course specifically *on* ethical reasoning, and not a course where ethical reasoning was a featured part of <other instructional objectives>. Thus, this survey would need some modification for more general use, but preliminary evidence of validity for decisions about learning objectives was generated in the 2017 study.

Leveraging your colleagues' interests in continuing -and contributing to- the development of ethical reasoning will enable your program (or team) to engage in truly developmental, and integrated, ethics instruction. Options include, but are not limited to, incorporating student responses on questions about their perceptions of the sustainability of their learning about ethical practice within a capstone experience. If a capstone component will feature ethical reasoning, then both engagement by other instructors, and the structural features recommended throughout these ten rules, will support a meaningful integration of ethics instruction that students can reflect on appropriately in the capstone. Getting buy-in from colleagues for instructional change is an important component of effective, long term curriculum improvement (see Diamond, 2008), and the true integration of ethics instruction throughout a program of study – culminating in at least one feature of the capstone experience – will encourage student buy-in as well.

3. Discussion

"Teaching ethics" is an oversimplified instructional objective. The integration of "ethics" into every quantitative course is possible – and desirable (Tractenberg 2016-a, Tractenberg, 2016-b), primarily through purposeful articulation of learning objectives specific to aspects of ethical practice that are realistic and achievable. This integration can be accomplished in an evaluable way by following the ten rules outlined here. The approach advocated is to *teach ethical reasoning*, specifically, how to make decisions throughout statistical and data science practice *ethically*. This requires much more than just an interesting case or important issue like "privacy" or "algorithmic fairness"; the knowledge, skills and abilities of ethical reasoning are useful in situations where these – and any other – ethical issues arise. It should be noted that, while "teaching" an interesting case is certainly simpler, faster, and easier to implement widely, the likelihood of a lasting impact – for example, promoting ethical practice or ethical reasoning, or even raising awareness of the practitioner's ethical obligations as they practice – is very low without learning objectives that involve reproducible and observable behaviors by students.

If the integration of "ethics" into quantitative courses focuses on teaching ethical reasoning steps with the ASA Ethical Guidelines for Statistical Practice (with or without the ancillary contributions from the ACM Code of Ethics), then instructors can initiate discussions in class as well as assign, and evaluate, objectively-assessed student work. The important feature of ethical reasoning is that the emphasis is not on "the right answer", it is instead on how the practitioner (and/or student) arrives at the answer/decision, and that the decision-making is evaluable and reproducible; assumptions and background knowledge are recognized so that, if these change, then the decisions can be revisited -in equally evaluable ways- if needed. Ideally, instruction in ethical reasoning with professional practice standards will teach *ethical reasoning*, so that if other practice standards become relevant (e.g., in the work place), the same approach can be used with the new material, or to identify and justifiably make decisions about new issues. Also ideally, this instruction in ethical reasoning will last beyond the end of the course (Knapper, 2006; see also Tractenberg et al., 2017)), and would even be leveraged and built upon in later quantitative courses.

A great deal of work is required to make an authentic case that represents an effective learning experience (but see Tractenberg, 2019); considerations in planning the instruction (per the CDGLs, Tractenberg et al. 2020) should include that student learning will be assessable by the instructor -and that what the students will/should learn is a realistic learning objective (or set of these). All formal instructional and curriculum development guidelines will specify that effective learning is promoted with instruction that keeps learning goals in mind at all times.

The potential impact of improving the development of ethical reasoning with practice standards is great. For example, the National Academies (NASEM 2017) called attention to “disruptive research practices” in their 2017 report, “Fostering Integrity”:

> “Actions such as failing to retain or share data and code supporting published work *in accordance with disciplinary standards*, practices such as honorary or ghost authorship, and using inappropriate statistical or other methods of measurement and data presentation to enhance the significance of research findings are clearly detrimental to the research process and may impose comparable or even greater costs on the research enterprise than those arising from research misconduct.” (p. 206; emphasis added).

The NASEM report urges greater attention to be placed on *training new scientists* to avoid these practices, underscoring the focus of this article on the integration of ethics training – to promote sustainable learning of that important material and the abilities to reason with it – for all who use statistics and data science, and not just the statisticians and data scientists. The ASA and ACM could argue that their disciplinary standards are the ones to be followed, by all who utilize their methods and technologies. Since some non-statistics/non-data science disciplinary norms, *particularly those in biomedical research*, as reported by Wang et al. 2018 (see also Martinson et al. 2005), might actually reinforce the disruptive practices the NASEM report discusses, the full and intentional integration, not concatenation, of training in ethical practices of statistics and data science are especially important for all quantitative courses (Tractenberg, 2016-A, 2016-B). Both the ASA, representing roughly 18,000 practitioners worldwide and the ACM, representing roughly 100,000 computing professionals worldwide, assert that their ethical practice guidance should pertain to members and non-members alike who utilize their methods and techniques. Thus, promoting the ethical use of statistical and computing practices using their practice standards, i.e., ethical statistics and data science, is justified (Martinson et al. 2005; Institute of Medicine, 2009; NASEM, 2017; Stark & Saltelli, 2018; Wang et al. 2018) and timely (National Academies of Sciences, Engineering, and Medicine, 2018).

Acknowledgements:

While this paper is intended to be *factual* - and the facts reported are correct - any opinions expressed are my own, and do not represent the perspectives of the American Statistical Association, Association for Computing Machinery, or Georgetown University.

 This manuscript was previously published as: Tractenberg RE. (2020, August 15). Ten simple rules for integrating ethics into statistics and data science instruction. SocArXiv. https://doi.org/10.31235/osf.io/z9uej. Apart from formatting, these documents are identical.

References

Ambrose SA, Bridges MW, DiPietro M, Lovett MC, Norman MK. (2010). *How learning works: Seven research-based principles for smart teaching*. San Francisco, CA: John Wiley & Sons.

American Statistical Association (ASA) *ASA Ethical Guidelines for Statistical Practice-revised* (2018) downloaded from https://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx on 30 April 2018.

Association for Computing Machinery (ACM). *Code of Ethics* (2018) downloaded from https://www.acm.org/about-acm/code-of-ethics on 12 October 2018.

Barnett SM & Ceci SJ. (2002). When and where do we apply what we learn? *Psychological Bulletin* 128(4): 612-37.

Bloom BS (Ed), Englehart MD, Furst EJ, Hill WH & Krathwohl DR. (1956). *Taxonomy of educational objectives: the classification of educational goals, by a committee of college and university examiners. Handbook I: Cognitive Domain.* New York: David McKay.

Briggle A, & Mitcham C. (2012). *Ethics and science: An introduction.* Cambridge, UK: Cambridge University Press.

Diamond RM (2008). *Designing and Assessing Courses and Curricula (3E).* San Francisco CA: Jossey Bass.

Dow MJ, Boettcher CA, Diego JF, Karch ME, Todd-Diaz A, Woods KM. (2015). Case-based learning as pedagogy for teaching information ethics based on the Dervin sense-making methodology. *Journal of Education for Library and Information Science* 56(2-Spring): 141-157.

Fink LD (2013). *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses, 2E.* San Francisco, CA: Jossey Bass.

Grace S & Trede F. (2011). Developing professionalism in physiotherapy and dietetics students in professional entry courses. *Studies in Higher Education.* DOI: 10.1080/03075079.2011.603410.

Gunaratna N, Tractenberg RE. (2016). Ethical Reasoning for consulting with the 2016 Revised ASA Ethical Guidelines for Statistical Practice. *Proceedings of the 2016 Joint Statistical Meetings*, Chicago, IL. Pp. 3763-3787.

Handelsman, J., Miller, S., and Pfund, C. (2007). *Scientific teaching*. New York: W.H. Freeman.

Hutchings P, Kinzie J & Kuh GD (2015) Evidence of student learning: What counts and what matters for improvement. In GD Kuh, SO Ikenberry, & NA Jankowski (Eds.). *Using Evidence of Student Learning to Improve Higher Education*. Somerset, NJ: Jossey-Bass. Pp 27-50.

Institute of Medicine. (2009). *On Being a Scientist: A Guide to Responsible Conduct in Research: Third Edition.* Washington, DC: The National Academies Press. https://doi.org/10.17226/12192

Knapper C. (August 2006). *"Lifelong learning means effective and sustainable learning: Reasons, ideas, concrete measures."* Seminar presented at 25th International course on Vocational Training and Education in Agriculture, Ontario, Canada. Downloaded from h6p://www.ciea.ch/documents/s06_ref_knapper_e.pdf on 2 Oct 2013.

Knowles MS, Holton III EF, Swanson RA. (2005). *The Adult Learner, 6E*. Burlington, MA: Elsevier

Martinson, B.C., Anderson, M.S., and de Vries, R. "Scientists Behaving Badly." *Nature* 435(2005):737-738.

Messick S (1989). Validity. In: *Educational Measurement. 3rd ed*. New York, NY: Macmillian. Pp. 13-103

Messick S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher* 23(2): 13-23.

Moseley D, Baumfield V, Elliott J, Gregson M, Higgins S, Miller J, Newton DP. (2005). *Frameworks for Thinking: A handbook for teaching and learning*. Cambridge, UK: Cambridge University Press.

National Academy of Engineering (2013). *Practical Guidance on Science and Engineering Ethics Education for Instructors and Administrators: Papers and Summary from a Workshop December 12, 2012*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18519.

National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering Integrity in Research*. Washington, DC: The National Academies Press. https://doi.org/10.17226/21896.

National Academies of Sciences, Engineering, and Medicine (NAS). (2018). *Data Science for Undergraduates: Opportunities and Options*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25104.

National Institute for Learning Outcomes Assessment (2016, May). *Creating sustainable assessment through collaboration: A national program reveals effective practices.* http://learningoutcomesassessment.org/documents/Occasional_Paper31.pdf

National Research Council. (2001) *Knowing What Students Know*. National Academy Press: Washington, DC, USA.

National Research Council. (1999). *How People Learn*. National Academy Press: Washington, DC, USA.

Nicholls G (2002). *Developing teaching and learning in higher education*. London, UK: Routledge.

Nilson L (2016) *Teaching at its best: A research-based resource for college instructors, 4th E.* San Francisco, CA: Jossey Bass.

Paterson, M., J. Higgs, S. Wilcox, and M. Villenuve. 2002. Clinical reasoning and self-directed1earning: Key dimensions in professional education and professional socialisation. *Focus on Health Professional Education* 4, no. 2: 5–21.

Rios CR, Golde CM, Tractenberg RE. (2019). The preparation of stewards with the Mastery Rubric for Stewardship: Re-envisioning the formation of scholars and practitioners. *Education Sciences* 9(4), 292; https://doi.org/10.3390/educsci9040292

Rumelhart DE. (1980). Schemata: The building blocks of cognition. In RJ Spiro, BC Bruce, & WF Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Erlbaum. pp. 33-58.

Santa Clara University. (no date). *Ethical Reasoning*. Downloaded from http://www.scu.edu/ethics/ 29 November 2009.

Schwänke U. (2009). "*Sustainable learning - how storyline can support it.*" Paper presented at the Nordic Storyline Conference, Gothenburg, NE. Downloaded at http://www.storyline-methode.de/mediapool/43/436167/data/Sustainable_learning_-_nachhaltiges_Lernen.pdf 10 October 2013. pp 1-2. (**this URL has not been active since 2017**)

Stark PB & Saltelli A. (2018). Cargo-cult statistics and scientific crisis. Significance 15(4): 40-43. Downloaded from https://rss.onlinelibrary.wiley.com/doi/10.1111/j.1740-9713.2018.01174.x on 10 September 2018.

Tractenberg RE. (2016-A). Why and How the ASA Ethical Guidelines should be integrated into every quantitative course. *Proceedings of the 2016 Joint Statistical Meetings,* Chicago, IL. Pp. 517-535.

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. Pp 115-139.

Tractenberg RE. (2017, October 6). *Preprint*. The Mastery Rubric: A tool for curriculum development and evaluation in higher, graduate/post-graduate, and professional education. Published in the *Open Archive of the Social Sciences* (SocArXiv), osf.io/preprints/socarxiv/qd2ae

Tractenberg RE. (2019, April 23). *Preprint.* Teaching and Learning about ethical practice: The case analysis. Published in the *Open Archive of the Social Sciences* (SocArXiv), https://doi.org/10.31235/osf.io/58umw

Tractenberg RE. (2020-A, February 20). Concordance of professional ethical practice standards for the domain of Data Science: A white paper. published in the *Open Archive of the Social (SocArXiv*), DOI 10.31235/osf.io/p7rj2

Tractenberg RE. (2020-B, April 21). *Preprint.* The Assessment Evaluation Rubric: Promoting learn*ing* and learn*er* -centered teaching in higher education. Published in the *Open Archive of the Social Sciences* (SocArXiv), 10.31235/osf.io/bvwhn

Tractenberg RE. (in review-A October 2020). *Ethical Reasoning for the Quantitative World*. (Book~280 pages)

Tractenberg RE. (in review -B October 2020). *Ethical Practice in Statistics and Data Science*. (Book~450 pages).

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, FitzGerald KT, & Collmann J. (2017). Evidence of sustainable learning with the Mastery Rubric for Ethical Reasoning. *Education Sciences*. *Special Issue: Consequential Assessment of Student Learning. Educ. Sci. 2017, 7(1), 2; doi:10.3390/educsci7010002 http://www.mdpi.com/2227-7102/7/1/2*

Tractenberg RE, Lindvall JM, Attwood TK, Via A. (2020, April 2) *Preprint.* Guidelines for curriculum and course development: a whitepaper for higher education and training. Published in the *Open Archive of the Social Sciences (SocArXiv*), doi 10.31235/osf.io/7qeht.

Tractenberg RE, Russell AL, 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. DOI 10.1007/s11948-014-9613-1

Trede F. (2012). Role of work-integrated learning in developing professionalism and professional identity. *Asia-Pacific Journal of Cooperative Education* 13(3): 159-67.

Wang MQ, Yan AF, Katz RV. (2018). Researcher requests for inappropriate analysis and reporting: A us survey of consulting biostatisticians. Annals of Internal Medicine 169(8), 554–558.

APPENDIX: Sustainability survey, adapted from Tractenberg et al. (2017)

The following are instructions and survey items (yes/no plus essay prompts) that can be utilized to engage students in reflecting on, and documenting, their perceptions of the extent to which learning about ethical practice was sustained, following the four-dimensional definition of "sustainable learning" of Schwänke **(2009). This survey was originally published in 2017; the text below has been modified in the highlighted text to help others adapt these items to their own courses.**

Please read the following definitions of "lifelong learning", "changing your behavior as a result of acquiring new knowledge", "a process of personal development continuing beyond the course", and "deconstruction/reconstruction". Based on these definitions and thinking about your own **reasoning** in your daily life or work *since **you completed the <course topic> course***, answer the questions below. Each question asks for ONE example, but if you can estimate how many times each type of experience has occurred, please do so. Describe as many experiences as you choose (no word limit!)—we only NEED one experience on each of the four dimensions, but will use everything you can provide to improve our understanding of your learning and the effectiveness of the ethics and ethical reasoning instruction in <***the program of study***>.

1. **Lifelong learning**: an additional level of depth, or dimension, that the course's emphasis on ethical reasoning skills and reflecting on your own thinking enabled you to bring to a course or experience *unrelated to <**the course topic**>.*

Since the end of the **<*course topic*>** course, have you had at least one experience that fits the "lifelong learning" definition?

YES (ESTIMATED NUMBER:) NO

Briefly describe how the **course's emphasis on ethical reasoning <*what did your course emphasize that should be sustained??*>** and your own reflection on your thinking led to an additional level of depth, thought, or reflection in another course or other experience. IF this has happened to augment your awareness of *ethical challenges* <*what was your course about!?*>, please note that! However, this question is asking about your application of what you learned in our course to something other than ethical challenges.

2. Changing your learning behavior as a result of *acquiring ethical reasoning KSAs (achieving the learning outcomes of the course)*: describe how your learning (fact-finding, thinking, understanding of something, or approach to learning something new) changed as a result of having completed the ***<course topic>*** course.

Since the end of the **<*course topic*>** course, have you had at least one experience that fits the "changing your learning behavior" definition?

YES (ESTIMATED NUMBER:) NO

Briefly describe how the course led to changes in your learning. *This question is asking about your application of the specific KSAs you learned in our course* to other types of learning—including *ethics (e.g., if you participated in a seminar or course on ethics*) but not limited to that. For example, if your participation in/benefit from journal club meetings changed, use that experience.

3. **A process of personal development continuing beyond the course**: Something you did, or initiated, for your own sense of learning (i.e., not taking a course as part of your program, but a learning or training experience that you sought, created, or identified that was not already planned). Something you feel you will/would/would have benefitted from, *that was not already in progress*.

Since the end of ***the <course topic>*** course, have you had at least one experience that fits the "process of personal development beyond the course" definition?

YES (ESTIMATED NUMBER:) NO

Briefly describe how the course led to your identification or creation of this experience (including, if appropriate, your recognition of an opportunity as useful that you hadn't considered to be "useful" before—maybe you only thought it would be "interesting", but not specifically contributory to your own personal development). This question is asking about whether and how your level of awareness of your *own reasoning*, and how to try and strengthen it, was changed after or beyond the course itself—including ***the <course topic>*** (e.g., if you participated in another seminar or course on ***the <course topic>***) but not limited to that. For example, if your perception of an opportunity's potential to contribute to your own development changed, use that! This can also include decisions NOT to participate in one opportunity because you perceived that the potential to strengthen or increase your own skills to be less than you desired.

4. **Deconstruction/reconstruction**: An idea or concept that, *prior to our course*, you thought you understood, but that you recognized you did not truly understand (*deconstruction)* and so sought to understand more deeply, and discovered an error in your original understanding that you remedied or sought to remedy (*reconstruction*).

Since the end of ***the <course topic>*** course, have you had at least one experience that fits the "deconstruction/reconstruction" definition?

YES (ESTIMATED NUMBER:) NO

Thinking about the time after the course ended, briefly describe an idea or concept that you thought you understood before taking our course, but that you later recognized you did not truly understand. Describe how the course *or reasoning KSAs* ***the <course topic>*** led you to recognize that you did not truly understand it, and whether and how these *same KSAs* helped you to reconstruct the concept without the original error.

This question is asking about your ability to identify ideas in your head that might require deconstruction (in order to determine if you truly understand them)—and also about your ability to reconstruct them—by seeking new information or experiences, and rebuilding the concept without the original error or gap. For example, you may have chosen NOT to participate in work or learning opportunities because you originally perceived that there was no potential to benefit from them; however, developing the ability to isolate and reflect on *your reasoning skills* may have led you to perceive new/unknown potential for learning in experiences was greater than you were capable of recognizing before the course.