—HIGHER—
EDUCATION
—POLICY—

Higher Education Policy, **2020**, **33**, (45–64)

www.palgrave.com/journals

Original Article

Enhancing Academic Quality and Collegial Control: Insights from US Policy on the Ethical Conduct of Human Subjects' Research

David D. Dill
Department of Public Policy, University of North Carolina, Chapel Hill, NC 27599-3435, USA.
E-mail: david_dill@unc.edu

Government initiatives and regulations intended to assure academic quality have been implemented in many countries over the last 25 years. Because of reservations about the effectiveness of these externally oriented policies, they have undergone continual change and adaptation. A number of countries are now experimenting with internally oriented policies focusing on the reform or "enhancement" of a university's own collegial processes for assuring academic quality in teaching and student learning. During these same years many developed countries also implemented national policies regulating human subjects' research within their universities. What might be learned from the experience with national policies on human subjects' research that could help inform the design of more effective national policies intended to improve and enhance the quality of education within universities? This question is explored through an analysis of the development of US policy on human subjects' research as well as its implementation and impacts at a major American research university.
Higher Education Policy (2020) **33,** 45–64. https://doi.org/10.1057/s41307-018-0093-9; published online 18 May 2018

Keywords: academic quality assurance; university policy; higher education policy; human subjects' research policy

Introduction

Over the last twenty-five years as access to higher education rapidly expanded in many countries national policies designed to assure academic quality spread around the world (Dill and Beerkens, 2013; Williams and Harvey, 2015). Because these national quality assurance (QA) policies have had limited success in actually assuring and improving the quality of teaching and student learning in the university sector, they have been continually changed and adapted. In the European Union QA policies are now shifting from an external or "control-oriented" approach to a "development-oriented" or "enhancement" approach focused on improving processes within each university (Hopbach, 2014). For example, Germany, which had required external accreditations of each university study field,

is now offering "system accreditation," the option of accrediting an institution's internal system for assuring the quality of its own academic programs (Grendel and Rosenbusch, 2010).

Research on public regulation (Coglianese and Mendelson, 2010) provides some insight into the observed limitations of previous national QA policies on teaching and student learning in higher education. Means-based regulations, such as policies encouraging more hierarchical administrative control of academic governance or those requiring standardized student satisfaction surveys of university courses, may be ineffective because assuring and enhancing academic quality is a complex professional activity. For example, an econometric study of US universities (Carroll *et al.*, 2012) discovered academic decisions made primarily by administrators led to an over-investment in student consumption benefits, such as amenities for student life, and to higher total costs for undergraduate students. In contrast, decisions with greater faculty participation in academic governance led to less investment in student consumption and to increased academic quality as measured by the rigor of academic program offerings as well as faculty qualifications. Correspondingly, the scores of standardized student surveys of university teaching commonly used in the USA (Stark and Freishtat, 2014), and now mandated by some QA national policies, have been discovered to be biased by discriminatory evaluations of women and minorities, positively associated with the award of inflated student grades and negatively related to direct evidence of student learning.[1]

Ends-based regulations utilizing performance-based education indicators, such as student progression and graduation rates as well as graduate employment and lifetime earnings, are similarly problematic (Johnes, 2016). These indicators of university education are indirect or proxy measures of student learning and at best reflect institutional reputation and resources. They fail to capture the efficiency of resource use — the real value added of a particular university education. Consequently, QA "transparency" policies requiring provision of this type of information to student consumers or promoting university rankings based upon such indicators often encourage institutional manipulation and "gaming" rather than actual improvement in the academic quality of education. For example, universities in a number of countries (Johnes, 2016) have been accused of increasing graduation rates by lowering academic standards (e.g., "grade inflation"), of misrepresenting their published academic performance data, and even pressuring their students to provide favorable responses to related national surveys. This reality is reflected in recent efforts by international higher education agencies (Daniel, 2016) to clarify and publicize the significant dangers to academic quality posed by corrupt university behavior.

External "command and control" types of regulations have proved particularly ineffective when applied to organizations with heterogeneous missions, which produce complex difficult to measure outputs, and whose relevant technologies are dynamic (Coglianese and Mendelson, 2010). These first two conditions are

characteristic of universities in most developed nations (Van Vught and Ziegele, 2012; Johnes, 2016). Additionally, the "technology" of teaching, student learning, and student assessment is now undergoing rapid change in higher education (Massy, 2016). In short, when organizational problems are highly complex or poorly understood and when the regulated organizations are sufficiently diverse that one-size rules do not fit all, adopting a "meta-regulatory" approach may be particularly beneficial. This type of policy seeks to induce regulated institutions to develop their own internal, self-regulatory responses to encountered performance problems. Under the noted conditions organizations probably have far greater knowledge and information about their own core processes than do regulators. Therefore, if provided appropriate incentives and guidance, the targeted organizations themselves are more likely to find the most cost-effective solution to encountered performance problems. For these reasons an independent group has proposed a reform of the well-established US college and university accreditation process based upon a meta-regulatory model (Brown *et al.*, 2017).

Effective meta-regulatory policies have been designed and implemented for the control of toxic emissions in US industries and utilities (Coglianese and Mendelson, 2010), but how would such an approach best be applied to universities with their distinctive values, outcomes, and strong traditions of collegial control? The Nobel laureate in Economics Elinor Ostrom (2010) has similarly emphasized neither market forces nor the regulatory rules of the state are the most effective institutional arrangements for managing and providing complex public goods in self-organizing organizations, which like universities have a tradition of collegial governance. Ostrom (2000) noted the socially beneficial performance of self-governing organizations is affected by government action, whether for example national or local governments publicly grant organizational members the authority to govern themselves. But the performance of these self-governing organizations is more influenced by the effectiveness of their collectively designed processes for socializing and enforcing social norms, i.e., the shared understandings — the common culture — about actions that are obligatory, permitted, or forbidden.

The influence of professional norms and a shared academic culture on academic behavior is particularly significant in the research performance of the most respected universities (Paradeise and Thoenig, 2015). Within these universities academic quality in research is primarily sustained and improved through the social interactions that occur within and between academic subunits and among academic staff. These interactions include many formal and informal internal conversations among academic staff as well as repeated self- and cross-evaluations, which strongly regulate the behaviors of faculty members in differentiated academic units. In these elite universities the communal norms generated and communicated through these collegial processes of internal regulation and socialization appear to be a primary form of organizational control over the quality of university research. Whether similar communal norms and influential collegial processes exist to

effectively control the quality of instruction and student learning within these same universities is less clear (Braxton and Bayer, 1999; Shavelson, 2010).

The French sociologist Emmanuel Lazega (2001, 2005) has further developed and empirically tested a model of the social processes indispensable for effective professional behavior in knowledge-based, collegial organizations. Lazega focuses on mechanisms (e.g., "essential values and norms," "authority to know," "lateral control mechanisms," "graduated sanctions," and "precarious professional values") which make it possible for interconnected professionals to cooperate and engage in collective actions for the efficient production of complex work. His model offers potentially valuable concepts for the design of meta-regulatory policies intended to assure and improve academic quality in collegially governed universities.

Ironically over the same time period as many countries have been experimenting with national QA policies for teaching and student learning, a number of developed nations led by the US have implemented a meta-regulatory approach to governing research in their universities (Office for Human Research Protections, 2017). This national effort to regulate university behavior in human subjects' research has received little attention in the burgeoning international literature on QA policies (Williams and Harvey, 2015). In sharp contrast to many national QA policies, the regulation of human subjects' research in the USA was designed from the outset to focus on the ethical norms essential to responsible research behavior and to increase incentives for collective actions by the faculty members within each university to strengthen institutional oversight of academic research. The US policy therefore requires rigorous peer reviews of academic research behavior at the institutional level.

How might this US experience with the meta-regulation of academic research behavior inform current national efforts to enhance faculty engagement in university academic quality assurance? Utilizing the concepts introduced above this question will be pursued through an analysis of the development and implementation of the US human subjects' research policy as well as an exploration of its impacts in a major research university.

The Development of US National Policy on Human Subjects' Research

US as well as international concern with experimental research on human subjects was initially motivated by the 1947 Nuremburg Code issued by the judges conducting the trials of Nazi War criminals (Annas and Grodin, 1992). The Code pronounced ethical principles to be observed in the conduct of research on human subjects. It clarified the right of the individual as an autonomous human being to be informed of the expected effects of the research on her or his health or person, to refuse to be a research subject, and to terminate her or his participation in a study at

any time. Because the Code was developed by a military tribunal, a more comprehensive professional code of related conduct — the Declaration of Helsinki — was subsequently developed by the World Medical Association (1964).

US national policy regarding research on human subjects was more directly influenced by revelations and legal suits in the early 1970s stemming from the Tuskegee Syphilis Study conducted by the Public Health Service.[2] Many of the low-income African-American male subjects of this study remained uninformed of their illness and needlessly died.

The Tuskegee disclosures and the increasing reliance on medical research for the development of new drugs motivated the members of the US Congress to propose a single Federal Board to approve all health-related research on human subjects. The leaders of the National Institutes of Health (NIH), fearful it would be charged with this responsibility and legal liability, instead lobbied Congress for a law formalizing local, expert group review as a basis for approving proposed research on human subjects. The NIH leaders had previously implemented expert internal group reviews of all proposed research conducted in its federally supported research hospital. In contrast to the prevalent academic tradition of relying upon the ethics of individual scientists to guide choices in research on human subjects, NIH advocated an "ethics of place" (Stark, 2012). This process reviewed proposed studies on human subjects, utilizing expert peers from the hospital, who applied their collectively defined conception of research behavior that is "obligatory, permitted, or forbidden" (Ostrom, 2000).

In 1974 the US Congress adopted the National Research Act, which set the conditions for research on human subjects eligible for federal funding. Consistent with the NIH recommendations, the Act required all relevant institutions to establish Institutional Review Boards (IRBs), composed primarily of local expert researchers, to review pertinent proposed research. The Act also required appointment of a National Commission for the Protection of Human Subjects in Biomedical and Behavioral Research (1979). This Commission was charged with defining basic ethical norms to govern the conduct of biomedical and behavioral research involving human subjects and developing guidelines for institutional IRBs. The appointed Commission consisted of 11 members, eight of whom were respected university professors. Thus, consistent with Ostrom's (2000) principles for effective collective action in self-governing organizations, the relevant social norms for research behavior were tailored by members of the academic profession itself.

The National Commission's findings (1979) specified three essential ethical principles: respect for persons, beneficence, and justice. Respect for persons requires the informed consent of research subjects. Beneficence describes the researcher's obligation to systematically assess the risks to the subjects compared to the expected benefits of the study to the subjects and to the larger society. Justice addresses the equity of subject selection. Subjects should not be chosen because of

their easy availability, their compromised position, or their potential manipulability, but for reasons directly related to the problem being studied.

In 1981, the Department of Health and Human Services (DHHS) and the Food and Drug Administration (FDA) issued regulations based on the Commission's findings. In 1991 these regulations, the so-called Common Rule, extended the IRB requirement to all researchers, research institutions, and universities engaged in federally funded research on human subjects.

The development of this US policy and its implementation offer a sharp contrast to the regulations adopted to assure academic quality in teaching and student learning in the USA as well as in many other developed countries (Dill and Beerkens, 2013). The US policy on human subjects' research did not require universities to publish new information on their research performance to assure public "transparency." Nor did the policy create a new national agency to assure institutional accountability, nor enable greater control by university administrators over academic research, nor define specific new measures to evaluate research. Rather the US policy was designed from the outset to reinforce if not strengthen the authority of each institution's collective faculty, one of Ostrom's (2000) core principles for the successful guidance of self-governing organizations. In contrast to some national QA policies the US human subjects' research policy explicitly stipulated no negative IRB decision could be overturned by a university administrator or appeal body, a ruling thus far upheld by the US courts. The overall focus of the US regulations was on clarifying and promoting the systematic communication of the ethical norms and practices deemed essential to protecting human subjects in research as well as empowering the qualitative review of relevant proposed research by respected institutional academic peers within each college or university.

How was the US human subjects' research policy implemented in universities and what have been its impacts? This question will be explored through an analysis of the policy and practices adopted at the University of North Carolina, Chapel Hill,[3] utilizing Lazega's (2005) concepts of collegial control.

Human Subjects Research at UNC

The University of North Carolina, Chapel Hill (UNC), is a comprehensive public university with over 29,000 undergraduate, graduate, and professional students. International rankings of universities generally include UNC among the top 75 institutions in the world. The University has a comprehensive collection of schools including the College of Arts and Sciences, professional schools in Business, Education, Journalism, Law, and Social Work, as well as health science schools in Dentistry, Medicine, Nursing, Pharmacy, and Public Health. In 2016 the University ranked 6th among all private and public universities in US federally funded R&D grants and contracts.

UNC initiated group review procedures for some proposed research in 1966 when the US Surgeon General required proposals submitted for federal Public Health grants first be approved by a university IRB. As the federal requirement for IRB approval broadened to include other fields, UNC's oversight of research involving human subjects also expanded, but in a highly decentralized manner. By 2000 the university had eight IRBs operating out of five separate school-based offices. The independent operations were spread among the School of Medicine (with four IRB Committees), School of Dentistry, School of Nursing, School of Public Health, and Academic Affairs (one IRB Committee each).

This decentralized structure produced redundancies and inefficiencies across offices with supposedly identical missions: five sets of operating procedures; five Web sites and databases; five application forms and processes; five documentation standards; and five channels for communicating policy, both within and outside the institution. This variability also contributed to little or no sharing of best practices across units, to inconsistency in applying a common set of federal regulations, to uneven staffing and budgeting support by the different schools, and to an erratic distribution of workload, experience, and expertise. The university was thereby attempting to fulfill university-level ethical obligations, a prerequisite for federal research funding, with school-level operations. This exposed the institution to possible compliance problems.

Due to some deaths of research subjects the DHHS began to aggressively assess the effectiveness of university IRBs in 1998 and discovered ineffective and over-taxed IRB procedures (Nelson, 2014). Consequently, federal research funding was suspended for a period of time at several respected research universities. These pressures motivated the formation in 2001 of the independent, nonprofit Association for the Accreditation of Human Research Protection Programs (AAHRPP). In 2002 UNC made a commitment to pursue AAHRPP accreditation. A University Task Force appointed to study the matter concluded the currently fragmented IRB structure at the university was unlikely to be accredited and recommended actions to better coordinate ethical oversight of human subjects' research at UNC. Subsequently, a university-wide Office of Human Research Ethics (OHRE) was created, reporting to the Vice Chancellor of Research. A faculty member in Social Medicine, who had served as head of the Office of Human Research Studies in the UNC School of Medicine, was appointed its first Director.

To support the new office a University Advisory Committee was appointed representing a breadth of research and administrative perspectives. During 2003–2004 the OHRE began to standardize best practices and coordinate processes among the existing IRBs including completion of the first set of university-wide operating procedures (Nelson, 2014). A shared Web site, as well as common application form(s), consent forms, and internal training for all IRB members and

staff were also developed. In September 2004 the office was formally announced and the new procedures and tools were implemented.

As noted in Figure 1 the total number of proposed studies submitted grew from 4079 in 1999 to 12,790 in 2013, an average annual increase of 14% a year. The OHRE continually worked to make the IRB process more efficient in terms of financial resources and faculty time, as well as more predictable and effective. However, NIH policy on human subject research also required documentation of each researcher's knowledge of ethical and regulatory obligations with regard to protection of human participants. Reflecting Lazega's (2005) model of collegial control, the US national policy focused the university's attention on means of communicating among academic researchers the ethical "values and norms essential to effective professional performance."

The OHRE accordingly sought means of strengthening the academic culture whereby academic researchers are socialized to relevant ethical principles. UNC therefore extended the ethical documentation requirement to all faculty members, staff, students, and other personnel engaged in the design, conduct, or analysis of research on human subjects carried out under the auspices of the University, regardless of the source of funding. In 2005 UNC also joined the Collaborative Institutional Training Initiative (CITI), a Web-based program offering educational modules on ethical principles regarding human subjects' research, IRB regulations, informed consent, and vulnerable populations. Each module requires completion of a graded short quiz to assess researcher understanding, and all relevant UNC researchers were required to achieve an overall passing score. Following its first year of CITI participation UNC had become the largest user out of the several hundred involved universities. In addition, OHRE began offering educational seminars and lectures in a variety of settings across campus and in the local community, which addressed ethics-related issues with students, faculty, staff, and interested members of the public.

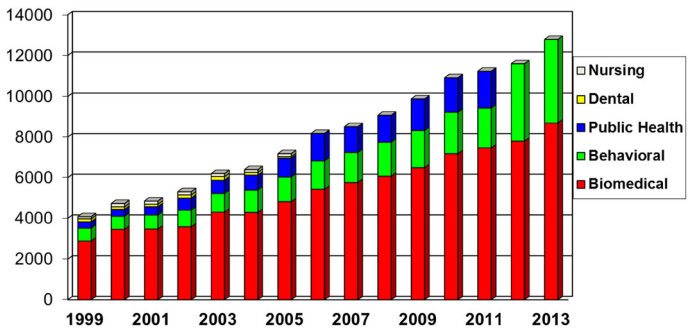

Figure 1. Growth in volume/complexity of UNC research requiring IRB review (Nelson, 2014). (Color figure online).

The US policy required each university's academic staff to collectively develop and implement IRBs, composed primarily of university academic researchers, to approve relevant proposed research. This meta-regulatory policy therefore created an incentive for UNC to design what Lazega (2005) terms a "lateral control mechanism," a primary means by which collegial organizations achieve effective quality assurance. In order to serve as an effective process for the monitoring and enforcement of social norms, such a mechanism needs to be conducted by trusted and respected institutional peers. Nominally, the appointment of UNC IRB members was made by the Vice Chancellor of Research at UNC. But over the last fifteen years the process was implemented in a highly collegial manner, with the IRB Chairs and the Director of the OHRE consulting personally with Deans, Department Chairs, and senior faculty members across the university to identify and recruit the ablest, most experienced, and best respected scholars for appointment to each IRB. As in other areas of university governance, recruiting experienced faculty members to serve became an increasing challenge, because of the growing time demands on academic staff for teaching, research, and administrative responsibilities. At UNC, similar to many US universities, release time for faculty members to participate in faculty governance activities is uncommon.

The US national policy also clarified who within the university possesses, in Lazega's (2005) terms, "the authority to know" what constitutes ethical behavior in human subjects' research. For example, university IRBs were required to include academic staff expert in medical research, research design, and ethics, as well as professionals representing the interests of vulnerable subject populations. The OHRE consequently sought to ensure any given research protocol was reviewed by an IRB with the most appropriate expertise, regardless of the researcher's academic affiliation. Therefore, over time the existing IRBs for the Schools of Dentistry, Nursing, and Public Health were phased out and six IRBS were formed across three nodes of expertise: Behavioral, Public Health/Nursing, and Biomedical (Figure 1)[4]. Every UNC IRB has 10 or more members and, following national guidelines, must have one member who is a non-scientist as well as one professional member, not otherwise affiliated with the University, who is to represent the perspective of research subjects. But the majority of each IRB are faculty members drawn from the UNC faculty. The IRBs are also truly multi-disciplinary, including the "Behavioral" IRB where proposals from faculty members in the College of Arts and Sciences and relevant Professional Schools in Academic Affairs are reviewed by researchers with experience in the sciences as well as by researchers in related social science fields.

Consistent with the organizational research on effective mechanisms of collective action (Lazega, 2005; Ostrom, 2010), UNC IRBs also developed a process of "graduated sanctions" for controlling unprofessional or opportunistic academic behavior. At UNC respected academic colleagues on the IRBs first talk

with and counsel potential research violators on means of improving flawed research proposals. Only after thoughtful and systematic efforts at personal education and socialization of a researcher have been made will a negative decision by an IRB be rendered.

One concern expressed about the development of university IRBs was "mission creep" (Gunsalus *et al.*, 2006). Regulatory oversight can encourage excessive, inefficient paperwork, and IRB reviews may be expanded to include research involving little risk to subjects. Is IRB review necessary, for example, for case study research, program evaluations, interviews with key informants about programmatic or organizational issues, or secondary use of publicly available data? The UNC OHRE invested considerable effort in carefully defining what constituted "human subjects' research" at a policy level and therefore what type of university research required IRB approval. By 2013 several hundreds of the proposed research studies submitted to the UNC IRBs, which formerly required a full review, were determined through an abbreviated process to be "exempt" from further evaluation.

US federal regulations also require IRB review of all relevant university studies. The possible inclusion of student-conducted research initially created some confusion and debate within the UNC community as well as within the IRBs, because of the US tradition of baccalaureate student honors theses as well as required student research assignments and projects in taught undergraduate and graduate classes. After extensive analysis the OHRE developed an *IRB Guidance for Student Research and Class Projects*, which clarified that many student class assignments are conducted principally for educational purposes. These guidelines carefully defined the types of student research exempt from IRB approval as well as the types of class activities which would warrant IRB review.

The OHRE regularly publishes its *Standard Operating Procedures* (UNC, 2017), a document of over 300 pages. The Procedures are constantly revised and updated to reflect changes in national guidelines as well as what the UNC IRBs have learned in the process of making decisions about submitted research proposals. The document provides detailed and up-to-date ethical guidelines on research involving vulnerable subjects. This ongoing publication thereby provides to UNC researchers more current and complete guidance on professionally responsible behavior in the design and conduct of research than can be deduced from publicly available professional standards or guidelines. In this sense the US IRB process is best understood, not as applying regulatory rules, but as "doing ethics" (Stark, 2012). This involves, as the UNC experience suggests, identifying and clarifying through an active process of "case-based decision-making" the ethical standards currently appropriate to designing and conducting research on human subjects as well as regularly communicating this information to all researchers in the UNC community.

Finally, the UNC IRBs have provided a new collegial mechanism by which the university faculty can renegotiate the shared "precarious professional values" (Lazega, 2005) essential to effective academic work. Researchers' academic freedom is an example of such a shared value, one which may become more precarious as it comes in conflict with researcher's ethical responsibility for research participants. As noted, the collegial IRB process now provides a means for each university to make, clarify, and continually communicate to the members of the academic community revised ethical standards for research based upon institutional peer-reviewed case decisions. The IRB process thereby provides to academic researchers a more immediate and respected mechanism for addressing the uncertainties and complexities in shared professional values inevitably caused by ongoing technical innovations and new developments in research.

In 2008, six years after its commitment to pursue AAHRPP accreditation, UNC completed the required self-assessment and formally submitted its application for human subjects accreditation.[5] Following a detailed review and commentary by AAHRPP, a revised version of the UNC self-assessment was accepted in 2009. AAHRPP then conducted a site visit, and UNC was accredited the following year. The AAHRPP accreditation review process differs from both US and many other national QA external reviews of teaching and learning in that its sole focus is on assessing the organization, effectiveness, and influence on academic behavior of a university's procedures for human research protection. The reputation of a university's research faculty, its research productivity, the quality of the institution's research facilities, the size of its research budget, or its means for managing and governing research activities other than the procedures for human research protection are not factors in the AAHRPP accreditation. Instead, the rigor and efficiency of a university's human research protection process in influencing faculty research behavior is investigated systemically by an external review team which follows a university's IRB approvals and rejections back to a sample of the reviewed university researchers (up to 50% of those interviewed by the external accreditation team). In their interviews with individual faculty members the external site visitors evaluate the researchers' review experiences as well as assess their personal ethical views and commitment to human subjects' research guidelines. Similar to external academic quality reviewers in the USA and other countries, the AAHRPP evaluators are trained for these reviews. But in further contrast to US academic accreditation reviews and the external QA reviews of some other countries, the major criterion for selecting the members of AAHRPP site visitation teams is relevant professional expertise. Therefore, AAHRPP university site visitors are composed solely of experienced university research administrators or university researchers with extensive experience and expertise in research on human subjects.

Implications for the Design of QA Policies

The US policy on human subjects' research and its implementation at a respected university suggest some critical issues in the effective regulation of academic behavior which could inform the design of more effective national QA policies for teaching and student learning. The preceding analysis is based upon a descriptive study of policy impacts at one university. Before exploring the issues raised for QA policy, it is appropriate to assess the generalizability of these findings. Some specific practices implemented at UNC, such as the requirement researchers successfully complete a designated external course on research ethics, may not be representative of all US research universities. But related research (Abbott and Grady, 2011; Cohen and Lynch, 2015; Stark, 2012) suggests the implementation and evolution of the human subjects' review process at UNC is similar to the experiences of other major US research universities. This is particularly true with regard to the composition and conduct of peer reviews by required University IRBs, the institutional emphasis on clarifying and communicating research ethics to all research staff, and the significant influence of collegial mechanisms of control in the university's human subjects' research practices. In addition, the discussed policies and procedures of the AAHRPP accreditation process are national in scope, thereby potentially influencing all US universities. At a minimum, therefore, the preceding analysis of US policy and its impacts at UNC offers some relevant questions for research on QA policies which may advance our knowledge regarding effective regulation in higher education.

A first issue is whether QA regulatory policies have focused sufficiently on what Lazega (2005) terms the "values and norms essential to effective professional performance." That is, the professional values and ethical obligations of academic staff regarding instruction, marking, and student assessment. As previously discussed, defining the ethical responsibility of researchers was a core component of the US national policy. Some academic critics of the US IRB review process (Cohen and Lynch, 2015) have argued the monitoring of ethical behavior in human subjects research is unnecessary, because university researchers learn ethical values and the norms of proper research design through their academic training. While this argument is debatable, a similar assertion with regard to university teachers has little empirical support. Comparative research on the academic profession (Cummings, 2010) revealed an average of 17% of academic staff from selected OECD countries reported receiving graduate training in instructional skills or learning about teaching methods during their research doctoral education. While 34% of surveyed US faculty reported such experience, an earlier national survey (Braxton and Bayer, 1999) discovered significant variability among US faculty members on proscribed behaviors regarding teaching and student assessment. US faculty concurrence with ethical norms relevant to instruction was discovered to vary significantly by subject field and type of academic institution, with the least

agreement on relevant proscribed behaviors reported by faculty members in selective research universities.

Some national QA policies and individual universities have attempted to address faculty values and norms regarding teaching and student learning. Related national guidelines have been developed in the UK by the Higher Education Academy (2017). At the institutional level "Principles of Teaching and Learning" (Eberly Center, 2017) have been developed and communicated to its faculty by Carnegie Mellon University (CMU) in the USA. These principles were derived from the rigorous and respected research on effective university course design by the CMU Open Learning Institute. A substantial amount of the international research on QA policy (Williams and Harvey, 2015) has also explored faculty attitudes toward and satisfaction with national QA policies. Their views are often reported to be critical or negative. But much less research has been conducted on the collective ethical norms and values which guide university faculty behavior in their teaching, grading, and student assessments. More such research appears warranted. But given the relative paucity of research and literature on the ethics of teaching in higher education, as compared to the burgeoning literature on the ethics of research, a "common law" approach to developing ethical guidelines on instruction and student assessment may be more appropriate for national policy. That is, rather than attempting to nationally define and enforce essential ethical norms for teaching and student assessments, as was done in US human subjects' research policy, external reviews of university quality processes might be designed to better motivate each university's collective faculty to develop, communicate, and monitor its own conception of ethical standards for teaching staff. National concurrence on relevant ethical practices thereby might evolve over time organically rather than by administrative or government fiat.

A second critical issue is the challenge posed to both national and institutional QA policies by rapidly changing technology. As the UNC experience suggests, technological advances in academic research contribute greater complexity to human subjects' research proposals and IRB reviews. Rapidly changing technology now poses similar challenges to the effectiveness of customary methods of performing university instruction as well as to the validity of traditional means of student marking and assessment. A carefully designed study of undergraduate courses in US research universities utilizing a "hybrid learning" form of instruction (Bowen *et al.*, 2014), which combined online instructional software with traditional forms of teaching, discovered the students achieved the same student learning outcomes as traditionally taught courses. But they did so with 25% less student time investment and lower overall university costs. The respected senior author subsequently called upon universities to implement systematic collegial reviews of the instructional methods employed by academic staff in all subjects or programs:

> Decisions … have to be made as to how to shape the export and import of new pedagogies across institutions as well as across fields of study. Advances in technology make it imperative to move away from historical notions that departments must drive all decisions of this kind. Moving away from a vertical, departmental "silo," approach to resolving important questions will not be easy, but it is essential. We have to organize ourselves to think more horizontally (Bowen, 2016, 14–15).

Bowen is essentially advocating the development within universities of a "lateral control mechanism" (Lazega, 2005), similar to the US IRBs, which could review all instructional programs and courses as a means of assuring academic quality and efficiency. A comparable collegial control mechanism, termed "academic quality work," has been designed and implemented by Massy (2016) in Hong Kong and the US university systems of Missouri and Tennessee. However, such university-wide academic quality assurance processes have often been resisted by academics who advocate a "federal" conception of collegial governance (Tapper and Palfreyman, 2010). This type of collegial control provides academic autonomy for schools and academic departments regarding means of instruction and student marking. But as Lazega's (2005) model further suggests, contemporary collegial organizations grappling with changing technology also need collective means for renegotiating "precarious professional values" such as the idea of academic freedom. Reflecting this view Bowen argued the independence of thought required by university academic staff to advance knowledge and properly educate students is linked to professional responsibilities, "which include the obligation to adhere to professional norms and to discipline those who fail to do so" (Bowen and Tobin, 2015, 201). The US experience with the IRB process for the peer review of human subjects research suggests a possible means for both collectively enforcing essential professional norms and negotiating the delicate balance between academic freedom and collegial accountability for academic quality.

A third critical issue for QA policy is whether it provides sufficient incentives for the regulated universities to develop principled, factually informed deliberation about the relevant terms of professional accountability. A weakness of many universities revealed in external quality assurance reviews is their failure to develop a common "culture of evidence" for assuring academic quality in teaching and student assessment (Shavelson, 2010). That is, does a university possess a governance and information structure which captures valid and reliable evidence on student learning, feeds it back to all academic levels, and rigorously monitors program progress on academic improvement?

Who should serve on the groups overseeing the quality of teaching and student learning, both external QA review teams and collegial committees within universities? Research on national QA policies emphasizes their influence on

university behavior is a function of their perceived political independence, their scientific knowledge, as well as the compelling authority of their expertise (King, 2009). Within collegial organizations the monitoring and enforcement of shared professional norms is most effectively conducted by trusted colleagues who possess relevant experience and expertise and are thereby awarded what Lazega (2005) terms "the authority to know." US research universities have traditionally made distinctions in authority based on knowledge (Dill, 2014). For example, full-time instructional staff in US academic departments are generally accorded the right to participate in decisions on the curriculum, course assignments, and junior appointments, but only tenured full professors are awarded the "authority to know" who among their departmental colleagues should be granted academic tenure and a professorship. Similarly, US national policy on human subjects' research requires all university IRBs include academic peers who are expert in ethics and research design, as well as professionals knowledgeable about human subjects, to thereby assure the proposed studies are ethically appropriate, are scientifically valid, and have been subjected to truly independent review by knowledgeable professionals (Emanuel *et al.*, 2000). As previously indicated, all IRBs at UNC include academic peers who, while not necessarily representative of the specific area of research under review, possess the knowledge necessary to determine if the proposed research is ethically responsible and uses accepted scientific principles and methods. In addition, the mentioned AAHRPP accreditation review teams are composed exclusively of knowledgeable academic professionals with extensive experience and relevant scientific expertise in human subjects' research. With this designation of the "authority to know," the US human subjects research policy appears to have motivated regulated universities to implement a collectively designed process for socializing and enforcing relevant professional norms, i.e., the shared understandings about research actions that are obligatory, permitted, or forbidden (Ostrom, 2000).

Accumulating research in the field of "learning science" (Massy, 2016) is now making significant contributions to our understanding of how effective learning at university level takes place and the means by which instruction and student assessment can be improved for maximum effectiveness. Arguably, effective assessment of academic quality assurance should therefore include a rigorous evaluation of whether an entity — a university, an academic program, or a course within a university — possesses teaching and student assessment processes reflecting the principles emerging from research on learning science. But unlike US policy on IRBs, and the required composition of AAHRPP review committees, there appears to be no similar expectation in US or other national QA policies that external review committees or university committees engaged in QA be staffed by academics with the scientific expertise to rigorously evaluate the validity, reliability, and efficiency of methods of instruction, student marking, and assessment.

The relevance of scientific expertise to QA policy is illustrated by the experience of a German university (Ganseuer and Pistor, 2016), which is developing a more evidence-based approach to assuring the quality of teaching and student learning in pursuit of the country's new form of "system" accreditation. The university was already subject to means-based QA regulations by its state, which mandated university administration of student satisfaction surveys of academic instruction and required use of these data in the evaluation of academic staff. But as previously noted standardized student surveys have been discovered to exhibit significant bias and are poorly related to effective instruction and student learning (Stark and Freishtat, 2014). To better monitor and improve instruction, direct assessments of teaching behavior appear more valid, such as systematic appraisals of instructional materials and classroom observations by academic peers, evaluation methods much less commonly employed within universities than student satisfaction surveys. Student evaluations of teaching do make valuable contributions to improving instruction and should be encouraged, but as with other evaluation measures, their benefit depends upon their design.

The German university similarly discovered the standardized student course surveys mandated by the state were of little value to academic staff seeking to improve their instruction. Instead, the university has designed and encouraged faculty adoption of new qualitative tools for obtaining student comments on their instructional experiences. These included mid-course student polls, which provided instructors with detailed, activity-oriented student feedback, and the election of student course representatives, who personally meet with the instructor to discuss potential problems. There is little evidence the problematic nature of standardized student satisfaction surveys discovered by this university has been acknowledged by other universities or publicly questioned by external QA review teams (Dill and Beerkens, 2013). This raises the question as to why available and relevant scientific expertise is not required for membership on university internal QA committees or external QA review teams, a practice clearly required by US human subjects' research policy.

In sum, an analysis of US human subjects research policy and its influence at a respected research university raises a number of provocative and potentially valuable issues for the design of more effective policies on academic quality assurance. These include the need for policies to place a greater focus on the professional values and norms essential to effective teaching and student learning, the need to motivate universities to develop or strengthen their collegial mechanisms for negotiating, monitoring, and enforcing these professional norms, and the need to utilize relevant scientific expertise in the conduct and practice of these policies. While the analysis suggests the potential for a meta-regulatory approach to academic quality assurance policy, the UNC experience also indicates more effective collective faculty action at the university level will likely involve an

increased investment of faculty time, at least for those peers actively engaged in educational oversight. At the same time, the analysis suggests an appropriately designed "lateral control system" can be dynamic, providing the opportunity for universities over time to become more genuine "learning organizations" in the assurance and improvement of academic quality.

Conclusion

As the policy experience of the US suggests, the primary means for protecting human subjects in academic research has not been through the competitive pressure of market forces, increased "transparency" for university research, greater authority for university administrators, or more specified indicators of university research performance. Rather it has been pursued by clarifying relevant ethical beliefs as well as strengthening the collegial processes within universities by which these academic norms are communicated, monitored, and enforced. This suggests the potential value of a similarly designed micro-regulatory approach to national QA policies, which focuses on reforming and strengthening each university's collegial processes for assuring the quality of teaching and student learning.

In this spirit more systematic research is needed to explore the challenges posed in designing effective regulation for self-governing, knowledge-intensive, collegial organizations (Lazega, 2005). With regard to universities, what are the shared professional norms and values essential to effective university teaching and student assessment? How do universities successfully organize and conduct teaching and student assessment? How do they assure and improve instructional quality? How do they preserve professional unity among teaching staff? How do they control academic deviance among those responsible for teaching, grading, and assessing students? How do they balance academic continuity among instructional staff with the need for continual technical change?

One can best address these questions by rigorous studies utilizing relevant theoretical models of the internal collegial processes of universities, identifying means of rationalization through effective collective action. In the newly changing environment of higher education, knowledge of the social mechanisms for achieving durable cooperation among professionally rival academic peers remains the best means for both improving academic quality and lowering the costs of universities in all countries.

Acknowledgements

I am indebted to Professors Maarja Beerkens, Larry Churchill, Mary Henkel, James Kehoe, Douglas MacKay, Daniel Nelson, John Nelson, and Dr. Elizabeth Kipp

Campbell as well as two anonymous journal reviewers for their advice on the development of this paper. I remain solely responsible for the arguments presented.

Notes

1 Recent research in France has discovered similar issues of reliability and validity in the standardized student satisfaction surveys now used in many other countries (Boring *et al.*, 2016).
2 This discussion of the evolution of US policy and the activities of the National Institutes of Health is based primarily on the historical review and analysis in Stark (2012).
3 The UNC analysis is based upon a review of relevant UNC faculty governance documents (https://facultygov.unc.edu/faculty-council/resolutions/), the Annual Reports and Standard Operating Procedures of UNC's Office of Human Research Ethics (http://research.unc.edu/human-research-ethics/), and in-depth interviews with the former and current Directors of the UNC Office.
4 As of 2016 these three nodes have been merged into "Behavioral" and "Biomedical" IRBs.
5 The discussion of AAHRPP is based upon materials available on their Web site (http://www.aahrpp.org/) as well as an interview with a senior staff member of the Association.

References

Abbott, L. and Grady, C. (2011) 'A systematic review of the empirical literature evaluating IRBs: What we know and what we still need to learn', *Journal Empirical Research on Human Research Ethics* 6(1): 3–19.

Annas, G.J. and Grodin, M.A. (1992) *The Nazi doctors and the Nuremberg code: human rights in human experimentation*, New York: Oxford.

Boring, A., Ottoboni, K. and Stark, P.B. (2016) 'Teaching evaluations (mostly) do not measure teaching effectiveness', *Science Open* https://doi.org/10.14293/s2199-1006.1.sor-edu.aetbzc.v1, accessed 23 October 2017.

Bowen, W.G. (2016) 'Issues facing major research universities at a time of stress and opportunity', https://doi.org/10.18665/sr.278053, accessed 23 October 2017.

Bowen, W.G., Chingos, M.M., Lack, K.A. and Nygren, T.I. (2014) 'Interactive learning online at public universities: evidence from a six-campus randomized trial', *Journal of Policy Analysis and Management* 33(1): 94–111.

Bowen, W.G. and Tobin, E.M. (2015) *Locus of authority: the evolution of faculty roles in the governance of higher education*, Princeton: Princeton University Press.

Braxton, J.M. and Bayer, A.E. (1999) *Faculty misconduct in collegiate teaching*, Baltimore: Johns Hopkins University Press.

Brown, J., Kurzweil, M. and Pritchett, W. (2017) 'Quality assurance in U.S. higher education: the current landscape and principles for reform', *Ithaka S + R* https://doi.org/10.18665/sr.303620, accessed 23 October 2017.

Carroll, K., Dickson, L. and Ruseski, J. (2012) 'Do faculty matter? effects of faculty participation in university decisions', http://economics.umbc.edu/working-papers/#2013, accessed 23 October 2017.

Coglianese, C. and Mendelson, E. (2010) 'Meta-regulation and self-regulation' in M. Cave, R. Baldwin, and M. Lodge (eds.) *The Oxford handbook on regulation*, Oxford: Oxford University Press, pp. 146–168.

Cohen, I.G. and Lynch, H.F. (2015) *Human subjects research regulation: perspectives on the future,* Cambridge: MIT Press.

Cummings, W.K. (2010) 'Teaching versus research in the contemporary academy', *Research Institute for Higher Education (RIHE) International Seminar Reports No.15,* http://rihejoho.hiroshima-u.ac.jp/pub_en/seminar_en.html, accessed 23 October 2017.

Daniel, J.S. (2016) *Advisory Statement for Effective International Practice. Combatting Corruption and Enhancing Integrity: A Contemporary Challenge for the Quality and Credibility of Higher Education,* https://www.chea.org/userfiles/PDFs/advisory-statement-unesco-iiep.pdf, accessed 10/23/2017.

Dill, D.D. (2014) 'Academic governance in the US: implications of a 'commons' perspective', in M. Shattock (ed.) *International trends in university governance: autonomy, self-government and the distribution of authority*, Oxford, UK: Routledge, pp. 165–183.

Dill, D.D. and Beerkens, M. (2013) 'Designing the framework conditions for assuring academic standards: lessons learned about professional, market, and government regulation of academic quality', *Higher Education* 65(3): 341–357. Academies Press.

Eberly Center for Teaching Excellence and Educational Innovation (2017) *Principles of teaching and learning,* https://www.cmu.edu/teaching/principles/, accessed 23 October 2017.

Emanuel, E.J., Wendler, D. and Grady, C. (2000) 'What makes clinical research ethical?', *Journal of the American Medical Association* 283(20): 2701–2711.

Ganseuer, C. and Pistor, P. (2016) *From tools to a system: the effects of internal quality assurance at the university of Duisburg-Essen,* UNESCO International Institute for Educational Planning, http://www.iiep.unesco.org/en/our-expertise/governance-quality-assurance, accessed 23 October 2017.

Grendel, T. and Rosenbusch, C. (2010) 'System accreditation', *Higher Education Management and Policy* 22(1), pp. 1–12.

Gunsalus, C.K., Bruner, E.M., Burbules, N.C, Dash, L., Finkin, M., Goldberg, J.P., Greenough, W.T., Miller, G.A. and Pratt, M.G. (2006) 'Mission creep in the IRB world', *Science* 312(5779): 1441.

Higher Education Academy (2017) *Supporting your teaching practice and career development,* https://www.heacademy.ac.uk/individuals, accessed 23 October 2017.

Hopbach, A. (2014) 'Recent trends in quality assurance? Observations from the agencies' perspective', in M. J. Rosa and A. Amaral (eds.) *Quality assurance in higher education: contemporary debates*, Basingstoke: Palgrave Macmillan, pp. 216–230.

Johnes, J. (2016) 'Performance indicators and rankings in higher education', in R. Barnett, P. Temple, and P. Scott (eds.) *Valuing higher education*, London: UCL IOE Press, pp. 77–105.

King, R. (2009) *Governing universities globally: organizations, regulation and rankings,* Cheltenham: Edward Elgar.

Lazega E. (2001) *The collegial phenomenon: the social mechanisms of cooperation among peers in a corporate law partnership*, Oxford: Oxford University Press.

Lazega E. (2005) 'The theory of collegiality and its relevance for understanding professions and knowledge intensive organizations', in T. Klatetzki und V. Tacke (eds.) *Organisation und profession*, Wiesbaden: Verlag für Sozialwissenschaften, pp. 221–251.

Massy, W.F. (2016) *Reengineering the university*, Baltimore: Johns Hopkins University Press.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral research (1979) *Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (Belmont Report), http://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/, accessed 23 October 2017.

Nelson, D. (2014) 'Looking back and looking ahead: reflections on 20 years in the IRB business', Paper presented at Research Ethics Grand Rounds, April 17, 2014, University of North Carolina, Chapel Hill, NC.

Office for Human Research Protections (2017) *International compilation of human research standards,* U.S. Department of Health and Human Services, https://www.hhs.gov/ohrp/international/compilation-human-research-standards/index.html#, accessed 23 October 2017.

Ostrom, E. (2000) 'Collective action and the evolution of social norms', *Journal of Economic Perspectives* 14(3): 137–158.

Ostrom, E. (2010) 'Beyond markets and states: polycentric governance of complex economic systems', *American Economic Review* 100(3): 641–672.

Paradeise, C. and Thoenig, J.-C. (2015) *In search of academic quality*, New York: Palgrave Macmillan.

Shavelson, R.J. (2010) *Measuring college learning responsibly: accountability in a new era*, Stanford: Stanford University Press.

Stark, L. (2012) *Behind closed doors: IRBs and the making of ethical research,* Chicago: University of Chicago Press.

Stark, P.B., and Freishtat, R. (2014) 'An evaluation of course valuations', *Science Open* https://doi.org/10.14293/s2199-1006; 1.SOR-EDU.AOFRQA.v1, accessed 23 October 2017.

Tapper, T. and Palfreyman, D. (2010) *The collegial tradition in the age of mass higher education*, Dordrecht: Springer.

The University of North Carolina at Chapel Hill (2017) *Human research protection program: standard operating procedures,* https://research.unc.edu/files/2017/05/SOP-June-2-2017-bookmarked-and-TOC-links.pdf, accessed 23 October 2017.

Van Vught, F.A. and Ziegele, F. (2012) *Multidimensional ranking: the design and development of U-Multirank*, Dordrecht: Springer.

Williams, J. and Harvey, L. (2015) 'Quality assurance in higher education', in J. Huisman, H. de Boer, D. D. Dill, and M. Souto-Otero (eds.) *The Palgrave international handbook of higher education policy and governance*, London: Palgrave, pp. 506–525.

World Medical Association (1964) *Declaration of Helsinki—ethical principles for medical research involving human subjects*, https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/, accessed 23 October 2017.