# Methodological Reflections on Designing Surveys to Explore Faculty and Administrator Perceptions of AI Ethics Education

Kylee Shiekh
*Department of Engineering Education*
*Virginia Tech*
Blacksburg, United States
kshiekh@vt.edu

Caitlyn Geraci
*Department of Psychology*
*University of Mississippi*
Oxford, United States
cgeraci@go.olemiss.edu

Deborah Mower
*Center for Practical Ethics*
*University of Mississippi*
Oxford, United States
dsmower@olemiss.edu

Glen Miller
*Department of Philosophy*
*Texas A&M University*
College Station, United States
glenmiller@tamu.edu

Qin Zhu
*Department of Engineering Education*
*Virginia Tech*
Blacksburg, United States
qinzhu@vt.edu

*Abstract— Despite the rapid expansion of AI ethics education interventions across various institutions, there is a notable absence of empirical research systematically mapping or comparing these interventions. Furthermore, the extent and mechanisms of their effectiveness or ineffectiveness remain largely unexplored. To address these gaps, we aim to design and conduct the first-of-its-kind national survey on the state of AI ethics education. This survey will examine how faculty and administrators across the spectrum of American colleges and universities approach AI ethics education. This paper presents our reflections on the methodology behind designing two surveys intended to explore faculty and administrator perceptions of AI ethics education. It also briefly introduces the structures of the two surveys. We hope the findings from this paper will be of particular interest to scholars and educators focused on evaluating institutional-level ethics education in technology and engineering disciplines.*

*Keywords—AI ethics, methodological reflections, survey design, faculty, administrators*

## I. Introduction

Every day brings new reports on how artificial intelligence (AI) will reshape social and economic relations. AI evangelists promise a new utopia as tasks (first menial, and then complex) are automated, new solutions discovered, greater efficiency found, human shortcomings overcome [1]. Research organizations are spending billions of dollars to develop new and better technologies to achieve these aims. This tremendous commitment to technological development challenges the ability of social structures to respond adequately, running the risk of "cultural lag" where norms, values, and regulation change more slowly than developing technologies [2].

In the case of AI, academics are aware of the possible social and technical transformations. Nearly each day, academics are bombarded by emails, announcements, and articles on AI in education from their department chairs, institutional leadership, professional societies' heads and listservs, the Chronicle of Higher Education, and popular press writers. While there are many important technological questions regarding AI development (such as its capacity, scale, and practical applications), educators also have begun to grapple with a second set of questions of what and how to teach students about AI and its ethical design and use [3, 4].

In response to such technological, cultural, and academic pressures, colleges and universities have developed and are developing a multitude of interventions to prepare their students for the ethical, social, and practical impact of AI. These interventions are being developed by faculty in various disciplines embedded in a wide range of institutions (research universities, private and state colleges and universities, community colleges, and minority-serving institutions) whose capacities and needs vary significantly. Despite the rapid expansion of AI ethics education interventions across various institutions, there is a notable absence of empirical research systematically mapping or comparing these interventions [5]. Furthermore, the extent and mechanisms of their effectiveness or ineffectiveness remain largely unexplored. To address these gaps, we received funding from the National Science Foundation's Ethics and Responsible Research (ER2) program to conduct the first-of-its-kind national survey on the state of AI ethics education. This survey will examine how faculty and administrators across the spectrum of American colleges and universities approach AI ethics education. The research project is constructed to build meaningful collaborations between research-intensive universities and their regional partners, including minority-serving institutions (MSIs), community colleges, and research-intensive universities throughout the grant, from instrument design to dissemination.

## II. Methods

The project team developed two surveys, one for faculty members and another for administrators. The survey for faculty members was designed for those who have developed or are

interested in developing courses or modules related to AI ethics. It collects information on their experiences on curricular development, pedagogy, assessment, and institutional support. The second survey is designed, for administrative personnel, is designed to develop an assessment of the current state of AI ethics interventions and initiatives at their institutions.

To begin this process, two graduate students were given the description and aims of the project and tasked with initial development of the surveys. Both surveys were developed in an iterative process, with twelve versions of the faculty survey and four versions of the administrative survey developed over the course of approximately three months. Drafts of the surveys were updated based on feedback from the project PIs, and the graduate students presented updated versions to be reviewed collaboratively during scheduled development meetings. The PIs were integral to developing the structure and wording of the surveys. The development of both surveys is detailed below.

### A. Faculty Survey

Our faculty survey was initially inspired by two existing frameworks in engineering ethics education: Hess and Fore's [6] engineering ethics interventions framework and Martin et al.'s [7] multi-level framework for the ethics education system.

More specifically, Hess and Fore's engineering ethics intervention framework emphasizes that the effective design of ethics education interventions relies on a harmonious alignment of content (or curriculum), assessment, and delivery (or pedagogy or instructional strategy). Hess and Fore posit that this framework can function as an analytical tool for systematically assessing the design of a specific ethics intervention [6]. Furthermore, it allows for a rigorous comparison of ethics interventions across diverse disciplinary and institutional settings. Hence, we opt for this framework to direct our research in systematically gathering and analyzing information concerning AI ethics education interventions across various institutions.

While Hess and Fore's ethics intervention framework centers on the course level of engineering ethics education, Martin et al.'s multi-level framework delves into how faculty, serving as decision-makers in engineering education, perceive and establish connections across different levels of the engineering ethics education system. This includes not only the individual (e.g., instructor's practices and beliefs) level but also institutional (e.g., ethics education policies in programs, departments, or institutions), policy (e.g., policies adopted by governmental and accrediting bodies or funding agencies), and cultural (e.g., values and norms in engineering education and practice) levels. We adopt this framework to capture faculty perceptions of how various factors at different levels manifest on campus and influence their decision-making processes in AI ethics education.

### B. Administrator Survey

Our administrator survey was built upon Beever et al.'s [8] institutional epidemiology framework. Such a framework delves into ethics education on a community or institutional scale, investigating the specific locations within the institution where ethical learning takes place. It can function as a tool for a critical examination of: (1) the values inherent in administrators' decision-making; (2) the impact of institutional characteristics and cultures on the planning of the ethics curriculum; and (3) the interplay between the ethics curriculum and the broader STEM curriculum. We employ this theory to describe the institutional contexts and factors that influence AI ethics education.

### III. STRUCTURE OF THE SURVEYS

### A. Faculty Survey

The faculty survey consists of four sections that aim to gather information about faculty, their expertise, and efforts in developing courses and modules, and the support and challenges that they have encountered in their institution. The "Background" section gathers demographic information about both the individual and their institution. It aims to give us insights into who is developing, or is interested in developing, AI ethics interventions, and how the institutional context (institution types, size, and location, as well as the department or school in which the faculty work) may shape their development. Participants have the option to upload their CV, or to answer questions about their educational history and current roles. This is important to understand the position in which faculty are engaging AI ethics, as the widespread interest and concern about AI have drawn faculty from a range of backgrounds working in a variety of roles. It is worth noting that we plan to capture information from all faculty respondents, regardless of whether they are in the position to teach courses or modules on AI ethics.

The second section of the survey collects information on "Standalone Courses," , i.e., a course in which more than half of the content is related to AI ethics. It asks respondents who have taught a standalone course to provide information about the course, where it fits in different degree programs, and about the curriculum they develop, their pedagogical approach, and the assessments they use. Respondents may upload a syllabus. A maximum of 3 courses are allowed for entry, so that participants weren't overwhelmed by questions to answer.

The third section of the survey, "Course Components," collects information on AI ethics modules that are used in courses that have a different primary focus. We aim to learn about the course in which the module is used, the percentage of coursework focused on AI ethics, and about the objectives of the module. In addition to entering basic information similar to what is collected for standalone courses, respondents are also given the opportunity to upload documents provide more information about the assessments they use and the case studies they find most trenchant.

We also seek to collect information from survey participants who have not taught a standalone course or a course with a component on AI ethics. The survey aims to determine why these participants have not done so, to learn why this is the case. It will help us determine if a lack of opportunity, resources, skills, pedagogical constraints, or other barriers are impeding the development of AI ethics interventions, and to see this at the national scale.

The last part of the survey, "Curricular Integration and Institutional Support," aims to collect faculty perceptions about the institutional contexts in which they are working. The section is designed to complement the administrator survey; this section provides the faculty view of the institutional measures and context around them. We inquire about current AI ethics content, various stakeholders interested in the development of

this content, opportunities for faculty development, and resources provided to support faculty.

*B. Administrator Survey*

Similarly to the faculty survey, the administrator survey begins with a "Background" section which parallels questions with the faculty survey about institutional size and type. We also query the participants' educational background and current administrative role to allow for sorting and deeper analysis of awareness of curricular and initiatives at their institutions. Participants have the option to upload their CV.

To understand the current institutional context for AI ethics interventions and courses, we use the "Current Initiatives and Development Plans" section to ask about the scope of initiatives for students, initiatives for faculty, and funding for AI ethics education. We also ask about cross-institutional collaboration and give the option to give a website link or upload documents related to AI ethics initiatives.

We also want to understand what the future may look like for AI ethics intervention, and how to better support and encourage such initiatives. The "Perceived Barriers and Opportunities" section aims to understand what factors lead to increased or decreased interventions, including implementation from an institutional perspective and the student perspective. We also ask how participants see change in the next 2-3 years to get an understanding of how much change is expected to happen. This, combined with questions about the importance of AI ethics education, help us track how administrators see the future of AI ethics education unfolding.

To better understand the factors that administrators consider at the institutional level, we use the "Stakeholders and Institutional Integration" section. This section asks about stakeholders both within and external to the institution, such as students, faculty, legislators. To gain understanding of the importance and role of these stakeholders, we inquire about the influence of each stakeholder category on AI ethics education decisions. We also inquire about whether, and which, specific institutional characteristics influence decisions about AI ethics education. We ask about current approaches to interdisciplinary work to understand how different institutions are implementing collaborative approaches to AI ethics education across fields of study. Lastly, we ask about the impact of planned initiatives, and how institutional history and culture has influenced strategies.

Participants are given an opportunity to provide feedback on questions they felt we should have asked but did not. This allows participants to provide as much information as they like and deem most relevant, and enables us to see areas of importance we have missed.

## IV. Reflections on the Design of the Two Surveys

*A. Faculty Survey*

The faculty survey was designed with two research questions in mind:

RQ#1: What are the curriculum priorities, assessment strategies, and pedagogical approaches commonly employed in (existing and planned) AI ethics interventions at the undergraduate and the graduate levels?

RQ#2: What are the different ways that faculty approach AI ethics interventions?

To address RQ#1, the survey was initially split to encompass the topics of "curriculum," "pedagogy" and "assessment." Curriculum was understood to detail course content, pedagogy to include course delivery, and assessment to include how students were examined and tested in the course.

We initially approached the faculty survey through a quantitative lens; as such, the questions were crafted with structured response options to ensure clarity and precision. For example, when asking about course content, as many potential course topics as possible were specified as a "choose all that apply" item. One advantage of this approach would be in facilitating the ease of data analysis following survey distribution. However, we were also aware that this runs the risk of biasing the survey participants in their responses through the presentation of topics. Participants may potentially answer "yes" to course content they do not include because it appears similar, while also omitting relevant course content because they felt constrained by the presented options or those we neglected. Following feedback from the PIs, the survey was adjusted to include more flexibility in item responses. Specifically, flexibility meant that the fixed menu of survey answers were expanded to include open-ended responses (See Figure 1 as an example). Although this may mean more time devoted to data analysis following survey distribution, there are options to assist in this process which have been described in the project description, including using tools such as natural language processing.
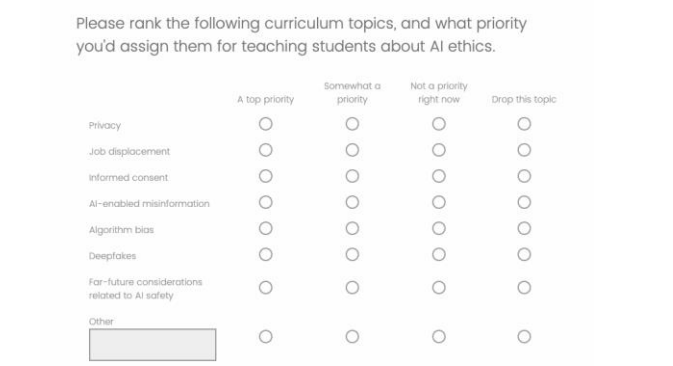


Please rank the following curriculum topics, and what priority you'd assign them for teaching students about AI ethics.

Figure 1. Example of including open-ended options (such as "Other") to increase flexibility in completing survey questions

Additional flexibility through adding open-ended responses also meant adjustments in how blocks of questions were ordered. This was motivated due to awareness that delineating between course topics, course delivery, and assessment ran the risk that questions would become redundant and lead to participant attrition. This also meant that we were asking broadly about any content, pedagogy, or assessment an instructor had ever used, which would render us unable to draw connections between these groups in a classroom setting. Ultimately, this led to a switch to ask about specific courses or course modules. The survey format was reoriented to instead ask if faculty had ever taught a stand-alone course related to AI ethics or a section of a course related to AI ethics. If someone

answered yes to either of these questions, they would be led to more specific questions in a branched-pattern about the design of that specific course or course component (see Figure 2).

Another unexpected challenge was the possibility that faculty may teach specific courses in high frequency. For example, a faculty member might teach multiple sections of the same AI ethics course each semester for the last two years, totaling six instances. Answering the same questions multiple times for each instance of the course would not only be repetitive and time consuming, but also increase the likelihood that memory loss and fatigue would degrade answers on the third or fourth round of questions. Solving this challenge required a two-fold solution. First, using the built-in functionality of Qualtrics, we were able to add a section at the beginning of each series of questions detailing which courses for which a person had already provided answers. Second, we included a question which allowed participants to upload a syllabus for repeated instances of courses instead of answering repetitive questions. Questions were made optional to allow for flexibility to upload a syllabus. Although tracking uploaded files imposes an additional time cost for researchers, the ease of responding by uploading a syllabus should lessen the cognitive and time burdens for participants. We hope this change will increase question response rates and quality and completion rates for the survey as a whole.

> You are currently responding about the following course: Responsible AI for Computer Scientists
>
> The following questions will be about pedagogy and assessment
>
> Course 1: What three topic areas are most important for students to understand from your AI ethics course?
>
> [                    ]

Figure 2. Example of questions asked for a particular course

Addressing RQ#2, "What are the different ways that faculty approach AI ethics interventions?" required that we dive further into faculty perceptions. In their multi-level review of engineering ethics education, Martin et al. [7] identify four aspects of faculty perceptions which we address in our survey. First, the individual level: we ask faculty how they perceive their own efforts to teach and assess ethics. Second, the institutional level: how do faculty perceive institutional efforts to implement and measure ethics in programs, departments, and institutions? To address this, we ask faculty about their awareness of institutional interventions and activities. Likewise, we ask about their perception of these interventions of activities, and whether they are positively received.

Martin and colleagues [7] also suggest examining policy and cultural aspects of faculty perceptions. To examine these, we have a section in the faculty survey to understand the context of their institution including the kind of support they experience and their awareness of AI ethics education policies.

While developing the faculty survey, we discussed the population we were targeting and whether there was a specific population of faculty that most typically teach AI ethics courses. After extensive discussion, we aimed to understand the practices and experiences of any faculty members involved in AI ethics courses and course practices. This could potentially include faculty who are housed in computer science, philosophy, medicine, and beyond. To understand participant differences by department and potential research areas, we included specific questions about this in the demographic and background section of the survey.

In this section, we also included questions about where these faculty would be housed within academia and within the national landscape of the United States. We expect that location and type of institution will impact how faculty enact and perceive AI ethics interventions, particularly where there are differences in state legislature that might impact how faculty teach, the size of the institution, and the makeup of the academic departments within the institution. To avoid an arbitrary classification of institutional sizes, we drew from the National Center of Education Statistics for guidance.

### B. Administrator Survey

The administrator survey was guided by the following research question and sub questions:

RQ #3: How do different types of institutions understand their needs and capacity for AI ethics education?

> RQ #3a. What are the current initiatives and developmental plans that institutions have undertaken regarding AI ethics education?
>
> RQ #3b. What factors do administrators perceive as barriers or opportunities for AI ethics education at their institutions?
>
> RQ #3c. How do administrators integrate their institutional characteristics including institutional size, institutional type, institutional cultures and histories, degree offerings, student demographics, and student interest into the support for and institutional plan of AI ethics education?
>
> RQ #3d. What influences have shaped and which stakeholders do institutional administrators consult with when making plans for AI ethics education (students, alumni, business and industry leaders, legislators)?
>
> RQ #3e. How do administrators address the relationship between current or planned AI ethics interventions and their larger university-wide curricular development plans?

The development of the administrator survey took place after the completion of the faculty survey. At this point, important decisions had been made regarding the faculty survey structure, which we found could be extended to an administrative lens. When designing the administrator survey, a less rigid survey structure was maintained, with additional focus on open-ended questions. This meant there were fewer versions of this survey, and it was written more quickly.

The graduate students who designed the initial drafts of the surveys were not experienced as administrators, and had limited knowledge of what these roles would entail and therefore the kinds of questions to ask. For this survey, the other authors were important for developing the questions and for the ordering of those questions.

Similarly to the faculty survey, we discussed the population of administrators, considering what level of information and access they would have to curricular details or initiatives and whether it would be limited by their own field or area of expertise. After extensive discussion, we aimed to understand institutional-level curricular initiatives, particularly interdisciplinary or cross-institution initiatives, the range of stakeholders, and opportunities and barriers particularly as they related to institution type and size.

The sections of this survey were designed to follow the sub questions of RQ#3. During the demographic information, we asked questions related to institutional size, location, student demographics and more to help us in answering RQ#3c. To answer RQ#3a, 3b, 3d, and 3e we used open-ended questions asking directly about these topics.

## C. Accessibility in Survey Design

Multiple accessibility concerns were raised during the development of both surveys. For all participants, we wanted to ensure that the survey structure would be easy to understand and that specific questions were clear with response options that minimized time and effort to select. This design goal imposed extensive thought and discussion on the wording of each item, ensuring that the broad populations that we are anticipating responses from can clearly understand the questions and their intention. This also meant that we were purposeful with branching, ensuring that only relevant questions were asked of each participant. For example, those who indicated having taught no stand-alone courses would not be asked to describe those courses or their details.

We also considered the timing of the survey. Both faculty and administrators are extremely busy people, and sending a lengthy survey would be disrespectful to their time and facilitate incomplete responses. As often is the case, there is a difficult balance between gathering desired level of data and creating a survey that does not impose a significant time demand. While our goal is to have each survey streamlined for completion in 20 minutes or under, we are still completing this review and revision.

Finally, when launching the survey for the review process, Qualtrics flagged some of our design decisions. First, because there are a variety of ways for someone to open the survey link and complete the survey, we had to adjust some questions to be legible on a phone screen. There were also accessibility concerns with certain question types, as they would be difficult to understand with a screen reader or with low vision. The changes recommended through Qualtrics were added before the initial drafts of both surveys were published for final review by the PI team. Current versions of both the faculty and the administrator survey are optimized according to all accessibility standards.

## D. Big Takeaways

As graduate students who are learning on the job, our big takeaways throughout the process of survey design have been intentionally moving away from rigidity, prioritizing ease of access in opening and completing the survey, referencing the core research questions to guide decision making, and identifying clear qualifiers for who will have access to completing the survey. We also want to remain mindful of the larger picture of the project and maintain alignment with funding goals. Finally, since additional data will be collected in the second year of the project through interviews, , we focused on thinking critically about what these two surveys were being designed to accomplish as compared to other components of the project. And while we are not focused on developing questions to drill deeper into specific patterns as part of the survey development, we have selected some promising lines of query for the future interviews as explorations and expansions of the surveys.

We realize that the development of a national baseline survey of AI ethics education is ambitious. Partly it is ambitious in its scope across the total number of individuals and institutions, but it is also ambitious in the range of information it seeks to capture with a single survey. The process of developing the surveys involved a team of researchers using models from the literature on survey design and ethics education, yet increasingly confronting the fact that the great variety of roles, institutional structures, institutional sizes, and initiatives poses two additional challenges. The first is that the survey may be too liberal: questions which are too general are likely to merely replicate this diversity without meaningful patterns. Second, and conversely, is that the survey may be too conservative: very specific questions may insufficiently capture important differences in approaches to AI ethics education in our efforts to design such a large survey. We think that we have found a Goldilocks approach with our design that successfully balances specific options and presented arrays of choices to participants along with free text responses to capture meaningful specificity and difference.

We have also come to appreciate how important it is to capitalize on the differences between institutions, even in the collaborative development of a survey. The many discussions we had over the course of months to develop these two surveys revealed assumptions each of us had about the curricular options, academic structure, policies, and initiatives that faculty and administrators are involved in or have sufficient knowledge about. It is a great strength of the survey that researchers from three different institutions were involved in its design. The overall approach and many of the questions would have differed greatly if the survey had been designed without such collaboration.

Additionally, this survey was developed as a pilot study to share with our partners at each of the three regional hubs for feedback as part of the developmental process. While the survey has not yet been shared with our partners they were ever present in our minds. Each of us would query or challenge others about how a particular question or batch of questions would be understood by the faculty and administrators at partner schools within our own hubs. Because we have each partnered with very different institutional networks, those differences also had a strong effect on the overall shape of the survey design and the

questions we developed therein. The survey also would have differed greatly if it been designed without keeping our partner networks in mind as our future collaborators.

## V. Summary and Future Work

Based on the reflections above, it is clear that survey design is a complex process. Such a process involves not only designing items based on research questions or theoretical frameworks but also including practical considerations, such as participant characteristics, cultures of organizations they are affiliated with, and the resource and expertise constraints of the research team.

For future work, we will employ various methods to validate the two surveys. The two most important methods include: (1) sharing the surveys with faculty and administrators affiliated with the three research hubs—University of Mississippi, Texas A&M University, and Virginia Tech; and (2) conducting interviews with selected faculty and administrators who complete the surveys, followed by triangulating the quantitative and qualitative data.

## VI. Acknowledgements

## References

[1] D. Dorr, L. Adams, and P. Embi, "Harnessing the promise of artificial intelligence responsibly," JAMA, vol. 329, pp. 1347–1348, 2023, doi: 10.1001/jama.2023.2771.

[2] W. F. Ogburn, Social Change with Respect to Culture and Original Nature, New York, NY: Huebsch, 1922.

[3] M. Coeckelbergh, AI Ethics, Cambridge, MA: MIT Press, 2020.

[4] S. M. Liao, Ethics of Artificial Intelligence, New York, NY: Oxford University Press, 2020.

[5] L. Weidener and M. Fischer, "Teaching AI ethics in medical education: A scoping review of current literature and practices," Perspectives in Medical Education, vol. 12, pp. 399–410, 2023, doi: 10.5334/pme.954.

[6] J. L. Hess and G. A. Fore, "A systematic literature review of US engineering ethics interventions," Science and Engineering Ethics, vol. 24, pp. 551–583, 2018, doi: 10.1007/s11948-017-9910-6.

[7] D. A. Martin, E. Conlon, and B. Bowe, "A multi-level review of engineering ethics education: Towards a socio-technical orientation of engineering education for ethics," Science and Engineering Ethics, vol. 27, p. 60, 2021, doi: 10.1007/s11948-021-00333-6.

[8] J. Beever, S. M. Kuebler, and J. Collins, "Where ethics is taught: An institutional epidemiology," International Journal of Ethics Education, vol. 6, pp. 215–238, 2021, doi: 10.1007/s40889-021-00121-7.