Examining Physicists' Ethical Reasoning: A New Methodology

Tyler Garcia,[1] Caitlin Solis,[1] Caleb Linville,[2] Bill Bridges,[1] Wyatt Jones,[1] Jonathan Herington,[3] Scott Tanona,[2] and James T. Laverty[1]

[1]*Physics, Kansas State University, Manhattan, Kansas, 66506*
[2]*Philosophy, Kansas State University, Manhattan, Kansas, 66506*
[3]*Philosophy, University of Rochester, Rochester, New York, 14627*

Researchers across all scientific disciplines routinely face ethical decisions in their work, from addressing conflicts of interest to deciding whether and how to make data available for reproducibility. To help strengthen their ethical reasoning skills, they are encouraged to take online training programs like the CITI program. Ethics training is insufficient for improving ethical behavior. Better understanding of how scientists make decisions and reason about ethics is needed. To develop that understanding, we need expanded, asset-based measures of ethical reasoning that can be applied to open-ended responses and discussions. As part of a year-long intervention on a group of fifteen scientists' value-based reasoning, we conducted pre/post interviews that included open-ended questions about ethical scenarios. For this paper, we explore an application of three theories of ethical and stakeholder reasoning to those answers, and determine that we can use them to examine quality, principles, and subjects of their reasoning in open responses.

2022 PERC Proceedings edited by Frank, Jones, and Ryan; Peer-reviewed, doi.org/10.1119/perc.2022.pr.Garcia

I. INTRODUCTION

Scientists across all disciplines undergo Responsible Conduct of Research (RCR) training. This training, often offered in self-paced online modules, tends to focus on topics like authorship, conflicts of interest, research misconduct, and the protection of human subjects [1]. Despite its intentions, research suggests that RCR training does not improve scientists' ethical knowledge, skills, attitudes, or behavior [2]. This is partly due to the fact that usual RCR training does not show applications of ethical principles to real-world problems [3].

Ethics training is still needed in physics as well as other disciplines. In 2004, Kirby and Houle found that around 39% of junior members of APS who responded to the survey had observed or had personal knowledge of at least one ethical violation, such as data falsification, plagiarism, or authorship [4]. In the absence of effective training, many physicists learn from their peers. However, there has been less written about ethical issues in physics than in other fields in science [5].

Whereas most ethics training focuses on applications of ethical theories or codes of ethics, we argue that discussing the *values embedded in ordinary scientific practice* is a key pathway to improving moral reasoning among scientists [3]. While scientists receive training in methodologies for distinguishing factual or theoretical questions, they receive little training in how to adjudicate questions about values. In this research we rely on the "values in science" literature that notes the many ways decision-making about scientific methods are laden with non-epistemic (e.g. ethical, social, legal, or economic) values [6, 7].

For this paper, we will use the Four Component Model to define ethical decision making. This model contains four parts: moral sensitivity, moral reasoning, moral motivation, and moral implementation. These components will be explained more in the theory section. With the aim of learning how to improve scientists' ethical decision making, we had scientists from different fields meet together in a fellowship to talk about scientific topics that are rooted in value judgements. For this paper, we investigate the moral reasoning from the physicists from the fellowship by: (i) looking at the *quality* of their reasoning according to the Neo-Kohlberg Schema, (ii) analyzing the *content* of their reasoning using a typology developed from Beauchamp and Childress' Principles of Biomedical Ethics, and (iii) identify the *moral subjects* of their reasoning using Stakeholder Theory.

Currently, many researchers measure the quality of ethical decision making by using tests like the Defining Issues Test (DIT or DIT2). [8]. These types of tests consist of asking someone what they would do in an ethically questionable scenario. The choices for the answers are in a multiple choice style and ranked based on their choice of action. We are interested in finding different ways to investigate moral reasoning that does not assume we know the ethically correct answers, includes multiple facets of reasoning, and can be applied to open-ended responses.

Rest et al argue that in order someone to become more ethical, they must improve on each of four moral components [9]. Findings from the developmental moral psychology literature show that changes in moral reasoning in particular are correlated with improvements in ethical conduct in the professions [8].

In this paper, we describe the theories of the Four Component Model, Neo-Kohlberg Schema, Principles of Biomedical Ethics, and Stakeholder Theory, and describe how we used them to characterize moral aspects of open responses. We describe the fellowship and provide examples of how we coded the data. We end with implications for scientists' moral reasoning and future work. Through all of this, we will be answering this research question: **How can we characterize scientists' ethical reasoning to investigate how their reasoning changed?**

II. THEORY

Rest's theory of moral development: Rest et al's four component model (FCM) of the psychology of moral behavior states that in order for someone to behave morally in a situation, they must demonstrate moral sensitivity, moral reasoning, moral motivation, and moral implementation [9]. Moral sensitivity is when someone is able to identify the morally salient features of the scenario, moral reasoning is when someone identifies the morally right action for the right reasons, moral motivation is where someone intends to do what is morally right, and moral implementation is where the person follows through on their motivation to behave morally. In this project, we focus on moral reasoning.

Neo-Kohlberg Schema: For this paper, we will be using the Neo-Kohlberg moral theory to categorize the participants' quality of reasoning by classifying their value-laden statements into categories(schema). This theory comes from the discipline of philosophy since physics has not looked into ethical reasoning. The Neo-Kohlberg theory includes three different schemas that are included in the DIT2: Personal interest schema, maintaining norms schema, and post-conventional schema. The personal interest schema focuses on the personal gains and losses that result from moral action. This schema focuses on reasoning that benefits oneself instead of using personal experience. The maintaining norms schema involves reasoning based on compliance with generally accepted social norms that maintain order and social cooperation. The post-conventional schema involves reasoning about moral obligations based on shared ideals that are fully reciprocal and open to ethical scrutiny [8, 10].

Principles: Beauchamp and Childress identified four principles to adhere to when it comes to ethical decisions in clinical medicine [11]. These principles are designed for making ethical decisions in biomedical ethics, we suggest that we can use them to categorize someone's context of reasoning in the STEM field. Definitions and examples can be found in Table I:

Stakeholders: For our definition of stakeholder, we will

TABLE I. Definitions and examples of autonomy, non-maleficence, beneficence, and justice from the Four Principles.

Principle	Definition	Example
Autonomy	The ability to make one's own decisions. There are two conditions that are essential for autonomy: liberty (independence from controlling influences) and agency (capacity for intentional action).	Participant gives informed consent
Non-maleficence	Non-maleficence is the obligation to not harm anyone intentionally. This can be both an action or a non-action that is made to prevent any harm.	Avoiding unnecessary harm towards patients
Beneficence	Beneficence is the obligation to contribute to someone's overall welfare. This is different from non-maleficence since beneficence is about taking positive acts to help others instead of refraining from harmful acts.	Improving someone's health through surgery
Justice	Justice is the fair, equitable, and appropriate treatment of other people. This principle states all people must be treated equally and must provide benefits to all people regardless of features like gender or race.	Everyone receives life saving supplies regardless of social class

be using Freeman's definition of a stakeholder: a stakeholder is any group or individual who can affect or is affected by the achievement of the organization's objectives [12]. This theory comes from the philosophy of business management. Freeman then identified two groups that are classified as stakeholders: "can affect" and "affected". In this paper we will refer to these as "involved stakeholder" and "affected stakeholder" respectively [13].

To further break down the classifications of stakeholders, Ulrich identified the different roles stakeholders can occupy [13, 14]:

- **Source of motivation/client:** The stakeholder whose values are being served
- **Source of control/designer:** The stakeholder who has the power to decide
- **Source of expertise/planner:** The stakeholder who contributes to the overall expertise of the project
- **Sources of legitimation:** The stakeholder that contributes the necessary responsibility of those involved

III. METHODOLOGY

A. Context

This study is part of a larger design-based research project to improve scientists' ethical decision-making. This project involved a year-long fellowship for faculty in the sciences. The fellowship consisted of 10 sessions over the course of an academic year. The fellowship started out with 15 participants with one of the participants dropping out. The participants come from the fields of biology, chemistry, physics, geology, and biochemistry. The fellowship was designed to foster directed exploration and discussion of the role that values play in decisions made within science. Throughout the fellowship, the participants were made aware that they were explicitly addressing value questions. They knew the research project was about ethics in science, but the research goals were not discussed.

We define an individual values something if the individual *prefers* that thing over some other thing [15]. In science, values can be "goals", "aims", "principles" or anything else agents want to be realized. Things can be valued intrinsically, because they are preferred on their own, or instrumentally, because they help one achieve other goals.

B. Data Collection

Our data for this project comes from individual, semi-structured interviews of fellowship participants conducted both before and after the fellowship series, respectively called the pre-interview and the post-interview. All interviews were video and audio recorded. Due to one of the fellowship participants dropping out, we only have data for fourteen of the participants. The participants were asked to respond to questions involving ethics in their research and also to a series of vignettes. This paper will focus on their responses to the vignettes since the Four Component Model is specifically concerned with how should a person reacts to specific situations.

The vignettes involved fictional scenarios where scientists were faced with an ethical choice regarding the conduct of their research. These were adapted from the Ethical Decision Making Measure (EDM) [16]. The interviewee was asked what would they do in that situation, and to describe their reasoning. Note that while these vignettes are not specifically related to physics, these concerns are relevant to physicists [4]. The RCR vignettes included six different scenarios:

- A grad student bringing up the concern that a scientist did not get informed consent to a project
- A project was conducted between two groups and one group got faulty data so the groups only wanted to use the "good" data
- A faculty revoked his promise that his grad student that they would be first author on the paper
- A scientist received a paper to review on a project that was similar to the one he was working on
- A grad student used unpublished data in their own project
- A scientist is in charge of hiring a new faculty member and could only bring in three male candidates but not the female candidate

C. Data Analysis

We made transcripts for all 28 (14 pre-interview plus 14 post-interview) of the video/audio recordings using otter.ai as the initial transcriber. We corrected the transcripts for all of the sub questions and responses starting from the first question of the interview and ended as soon as the questions about vignettes were done.

We then selected two of the physicists post-interviews to test if our proposed theories were viable for application to all participants. We put these transcripts into Excel with the interviewee questions in one column and the interviewee responses in the column next to the questions. We did this for every question and response in the vignette section of the transcripts. Once we put all the questions and responses in the Excel document we separated everything by the vignette scenario.

To code the responses, we highlighted where they were reasoning. For their Neo-Kohlberg schema, we looked to see if they were reasoning about themselves (self-interest), reasoning about how things should be based on social, legal, or scientific norms (maintaining norms), or reasoning for the betterment of others (post-conventional). For their principles invoked, we looked to see if their reasoning involved respecting other's choices (autonomy), avoiding intentional harms (non-maleficence), maximizing benefits (beneficence), or wanting what's fair for everyone (justice).

Using Ulrich's groups to help us classify stakeholders, we created a list of specific groups of stakeholders for ourselves which are: **self, students, scientists, administration, and public**. Note that stakeholders could be categorized in more than one group depending on the context of response. We did this for every vignette and for both physicists. A more detailed breakdown on how we coded the responses is in the "Vignette Responses" section below.

For inter-rater reliability, four of the authors of this paper met together and discussed each response and what we should code for each response. We came to an agreement on all of the responses.

IV. VIGNETTE RESPONSES

A. First Vignette

The following is a question-response from one of the physicists for the vignette regarding a grad student bringing up the concern that a scientist did not get informed consent to a project. In this, the italicized words indicate the question from the interviewer; the response is in normal text:

> (*How should [the scientist] respond?*) If he can really help improve the reliability of drugs, then maybe it's worth inflicting some pain on people...

For this response, the physicist's justification for inflicting pain is "If he can really help improve the reliability of drugs". While the action in this quote is about causing pain, the reason for inflicting pain is to increase benefits for other people (reliability of drugs), so this would be classified in the post-conventional schema.

As mentioned in the previous paragraph, the reasoning here is to increase benefits for other people. The principle here would be beneficence since the interviewee wants to increase benefits of the drugs.

In this response, the "he" would be the scientist from the vignette in charge of the project and the "reliability of drugs" affects the populace. The groups for stakeholders in this response would be "scientists" for the involved stakeholder and the "public" for the affected stakeholders since the interviewee is reasoning for the scientist in charge of the project to run the study to improve drugs for the people.

This next quote is from the same physicist from the same vignette:

> (*How do you think about causing pain to people?*) Well, I mean, if people are dying from the bad drugs, then I guess it's okay. Maybe to give some people shocks...

For this response, the action might seem unethical since they are saying that the scientist should perhaps shock people, which would cause harm to them. However, their reasoning that "if people are dying from the bad drugs, then I guess it's okay" implies that the shocks are to improve the safety of the drugs, which would ultimately prevent less deaths. This puts this quote into a post-conventional schema.

For the principle of this response, the interviewee is implying that they want to reduce the deaths from the bad drugs, which would be minimizing the harms from the drugs. The avoidance of harm would make this an application of non-maleficence. However, they also are suggesting shocking people would be justified by this consequence, which, since it is intentionally causing harm to individuals not offset by benefit to them, would go against the principle of non-maleficence.

The groups for stakeholders in this response would be "scientists" for the involved stakeholder, since they are still talking about the scientist from the vignette running the experiment, and the "public" for the affected stakeholder, since the scientist is running the project improve drugs for the people.

B. Second Vignette

The following is a question-response from a different physicist for the vignette regarding a scientist is in charge of hiring a new faculty member and could only bring in three male candidates but not the female candidate:

> *(And so why what's the obligation [for bringing in the female candidate]?)* Where's it come from? It's better for science, you're getting different, you're getting different perspectives, the perspective of her life, everything her life has told her everything about her environment, as a woman could be significantly different from that

> of a male.... And so you need this diversity in personnel in order to better guarantee diversity in how we solve problems, how we see the data, how we treat the data, etc. I think that's, that's essential.

In this vignette responses, the physicist is appealing to post-conventional reasoning since in this response the physicist is appealing to the fact that promoting diversity through hiring practices would be better for science, as is seen from their reasoning quotes of "It's better for science" and "you need this diversity in personnel in order to better guarantee diversity in how we solve problems".

This is the one of the responses we saw in the vignettes where there were more than one principle in each response and there was a supporting principle in this response. The interviewee responds with "It's better for science" where the "It's" in this response is diversity. This response would be a beneficence response since diversity is improving science (maximizing the benefits of science). Later on in this response, they say that diversity is needed for all of these problems and do not explicitly state that diversity would maximize benefit or reduce any harms from these problems. This might then be a justice statement. In this whole response, the beneficence in the first part with diversity would then go on to support the justice of diversity in solving those problems.

The stakeholders here are more complicated than the other vignette responses. The issue lies where the interviewee states "It's better for science". The "It's" would be the diverse group of scientists, so the involved stakeholder would be "scientists", but the affected stakeholder in this response would be "science". Since "science" is not an individual or group, we would then categorize this as "beneficiaries of science", which does not fit in any of the groups as mentioned above since all of the groups of stakeholders we found would benefit from science.

V. DISCUSSION

In order to fully analyze how these physicists are reasoning in ethical scenarios, we attempt to use the theories of Neo-Kohlberg Schema, principles, and stakeholders. Rather than saying that there is a correct way to ethically reason like some methods of measuring ethical reasoning, these theories allow us to view ethical reasoning in a more asset-based way since we are looking at how the physicists are reasoning with schema, principles, and stakeholders in their own way. However, there are troubles with our stakeholder definition, which are described more below. We would also like to note that the application of this combination of theories to open-ended responses is novel. Since we have shown that we can apply these theories to categorize the physicists' reasoning, we can attempt to use these theories for the rest of the scientists and eventually look to see any changes in their reasoning.

Neo-Kohlberg Schema: From the responses, we can identify the invoked Neo-Kohlberg schema with relative ease. The physicists invoked a mixture of post-conventional responses and maintaining norms responses with no self-interest responses. This could possibly mean that the level of reasoning is at a high level already for these physicists, or they could also be trying to make themselves look better in front of the interviewer.

Four Principles: The principles we do see in the above responses are beneficence, justice, and non-maleficence with none of the responses in this paper containing the principle of autonomy. This is not to say that the physicists don't invoke autonomy at elsewhere in the interviews. Beauchamp and Childress's principles of biomedical ethics can be applied to our setting; however, there were responses where it was not entirely clear what principle would align with the interviewee's reasoning. This may be due to lack of detail in the interviewee's responses, especially for non-maleficence, which is more clear in a clinical setting where one person has substantial responsibility and the potential to harm others directly.

Stakeholders: Identifying who the physicists are reasoning towards is not as simple as the traditional stakeholder theory would make it. The issue with our current theory of stakeholders is what happens when the the stakeholder is an abstract topic or object like in the last vignette response where the stakeholder is identified as "science" since our definition of stakeholder is an individual or group. This shows that our current use of stakeholder theory needs to be refined and our groups of stakeholders needs to be adjusted since the physicists reasoned with an abstract idea rather than an individual.

VI. CONCLUSION

In the end, we can conclude that through the Neo-Kohlberg schema, the Principles of Biomedical Ethics, and Stakeholder Theory, we can analyze scientists' ethical reasoning by looking at their quality of reasoning (Neo-Kohlberg schema), context of reasoning (principles), and subjects of reasoning (stakeholders). For future work, we plan on using these theories in combination on all of the participants. This includes looking at the changes in their reasoning with the theories from pre-interview to post-interview.

ACKNOWLEDGMENTS

We would like to thank the Kansas State University Physics Department for their support. We would also like to thank KSUPER for their valuable insights in our study. This material is based upon work supported by the National Science Foundation under Grant No. [1835366]

[1] Responsible Conduct of Research (RCR) Basic | CITI Program.
[2] S. Powell, M. Allison, and M. Kalichman, Effectiveness of Responsible Conduct of Research Course: A Preliminary Study, Science and engineering ethics **13**, 249 (2007).
[3] M. Mumford, *Read "Fostering Integrity in Research" at NAP.edu*.
[4] K. Kirby and F. A. Houle, Ethics and the Welfare of the Physics Profession, Physics Today **57**, 42 (2004), publisher: American Institute of Physics.
[5] M. Thomsen, Ethics Inside and Outside the Physics Lab, in *Handbook of Research Ethics and Scientific Integrity*, edited by R. Iphofen (Springer International Publishing, Cham, 2020) pp. 937–954.
[6] M. J. Brown, Values in Science beyond Underdetermination and Inductive Risk, Philosophy of Science **80**, 829 (2013).
[7] H. Douglas, Inductive Risk and Values in Science, Philosophy of Science **67**, 559 (2000), publisher: [The University of Chicago Press, Philosophy of Science Association].
[8] J. Rest, D. Narvaez, S. Thoma, and M. Bebeau, DIT2: Devising and Testing A Revised Instrument of Moral Judgment, Journal of Educational Psychology **91**, 644 (1999).
[9] J. R. Rest and D. Narvaez, eds., *Moral development in the professions: Psychology and applied ethics*, Moral development in the professions: Psychology and applied ethics (Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, 1994) pages: xii, 233.
[10] S. J. Thoma, Measuring moral thinking from a neo-Kohlbergian perspective, Theory and Research in Education **12**, 347 (2014).
[11] T. L. Beauchamp, J. F. Childress, and Oxford University Press, *Principles of biomedical ethics* (Oxford University Press, New York; Oxford, 2013) oCLC: 827736605.
[12] R. E. Freeman, *The Stakeholder Concept and Stragegic Management*, 2 (Pitman Publishing Inc., 1984).
[13] M. C. Achterkamp and J. F. J. Vos, Critically identifying stakeholders, Systems Research and Behavioral Science **24**, 3 (2007), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sres.760.
[14] U. Werner, Critical heuristics of social systems design, European Journal of Operational Research **31**, 276 (1987).
[15] D. M. Hausman, *Preference, Value, Choice, and Welfare* (Cambridge University Press, 2011).
[16] M. D. Mumford, L. D. Devenport, R. P. Brown, S. Connelly, S. T. Murphy, J. H. Hill, and A. L. Antes, ARTICLES: Validation of Ethical Decision Making Measures: Evidence for a New Set of Measures, Ethics & Behavior **16**, 319 (2006), publisher: Routledge _eprint: https://doi.org/10.1207/s15327019eb1604_4.