Student reasoning about sources of experimental measurement uncertainty in quantum versus classical mechanics

Emily M. Stump,[1] Courtney L. White,[2] Gina Passante,[2] and N. G. Holmes[1]

[1]*Laboratory of Atomic and Solid State Physics, Cornell University, 245 East Avenue, Ithaca, NY, 14853*
[2]*Department of Physics, California State University Fullerton, 800 N. State College Blvd., Fullerton, CA, 92831*

Measurement uncertainty and experimental error are important concepts taught in undergraduate physics laboratories. Although student ideas about error and uncertainty in introductory classical mechanics lab experiments have been studied extensively, there is relatively limited research on student thinking about experimental measurement uncertainty in quantum mechanics. In this work, we used semi-structured interviews to study advanced physics students' interpretations of fictitious data distributions from two common undergraduate laboratory experiments in quantum mechanics and one in classical mechanics. To analyze these interpretations, we developed a coding scheme that classifies student responses based on what factors they believe create uncertainty and differentiates between different types of uncertainty (e.g. imprecision, inaccuracy). We found that participants in our study expressed a variety of ideas about measurement uncertainty that varied with the context (classical/quantum) and the type of uncertainty.

2020 PERC Proceedings edited by Wolf, Bennett, and Frank; Peer-reviewed, doi.org/10.1119/perc.2020.pr.Stump
Published by the American Association of Physics Teachers under a Creative Commons Attribution 4.0 license.
Further distribution must maintain the cover page and attribution to the article's authors.

I. INTRODUCTION

One of the goals of an undergraduate physics education is that students understand the process of experimental scientific inquiry. To evaluate experimental results and claims, one must be able to assess the quality of the methods and data [1]. However, research has consistently shown that many introductory physics students hold ideas about measurement and uncertainty that differ from experts' [2].

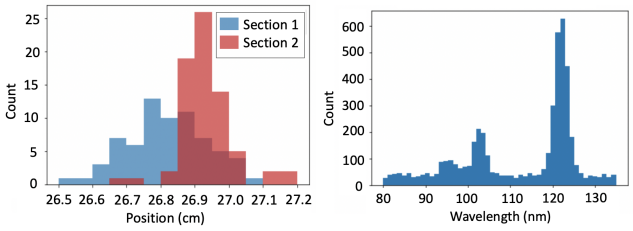
One source of confusion lies in terminology. Physicists typically use the term "error" to be synonymous with "uncertainty," which may refer to what is unknown about a measurement or the degree to which a measurement is limited. The colloquial use of error, however, is to describe a mistake. Research has found that many students interpret the physicist term "error" in this colloquial sense in classical mechanics experiments [3, 4]. This line of thinking has been observed in students' broader understanding of measurement. For example, students may expect that all errors (and thus uncertainty) can be eliminated by experts [3].

Terminology poses a similar issue for measurement and uncertainty in quantum mechanics. In quantum mechanics, uncertainty may refer to "the degree to which measurements on a microscopic scale are intrinsically non-reproducible" [5, p. 439]. This distinct definition of uncertainty, conflated with students' confusion about uncertainty in classical mechanics, can lead to further complications in student understanding.

To our knowledge, there is relatively limited research on student thinking about experimental measurement uncertainty in quantum mechanics. Several studies have shown that students draw on classical-physics reasoning to explain quantum-mechanical systems [6–8]. For example, one study found that many third-year undergraduate students interpreted quantum-mechanical uncertainty as a measure of what is unknown or limited [5]. However, a recent study found that students' epistemological views of classical and quantum mechanics are distinct [9].

In this study, we aimed to capture student ideas about sources of measurement uncertainty in both classical and quantum contexts. To classify these ideas, we drew on the Modeling Framework for Experimental Physics, which models experimental physics as generating, comparing, and refining models of physical and measurement systems [10]. Because a model is a simplified representation of the physical world [11], principles of uncertainty will be intrinsic to the modeling process. We can thus use the modeling framework to characterize student thinking in classical and quantum contexts within the same coding scheme. Research indicates that students draw on different forms of reasoning to answer different questions [12], so it is plausible that they would use different reasoning in these different contexts. However, other work found that students do not consider theoretical models when interpreting data [13], so when shown actual experimental data, students may ignore the theoretical context.

Previous work had developed an emergent coding scheme to analyze a subset of questions and scenarios from the interview data used in this study and found that students identified very different sources of uncertainty in the two contexts [14]. Here, we expand that work to evaluate whether this polarization might have resulted from the emergent coding scheme rather than underlying differences in student thinking. We also expand that work by probing student thinking in all parts of the interviews and disentangling references to precision and accuracy within the contexts. In particular, we sought to determine whether students attribute sources of uncertainty (generically, in reference to precision, or in reference to accuracy) to limitations or principles of the physical or measurement models, and how those attributions vary by context.



(a) (b)

FIG. 1. Histograms of (a) ball-drop and (b) spectrum data presented to participants during the interview. For the single-slit experiment, participants were shown the histograms in (a) with the x-axis altered.

II. METHODS

A. Participants and data collection

We used semi-structured interviews to probe student thinking about measurement uncertainty. Nineteen students were interviewed for this study, including 17 undergraduate and two graduate students. Participants were recruited from two institutions: one a research-intensive, private, and PhD-granting institution and the other a public, master's-granting, and Hispanic-serving institution. All participants had either previously taken a quantum mechanics course or were currently taking one at the time of the interview. Interviews took place in the second half of the spring semester.

The interviews focused on measurement uncertainty in the context of three experiments: one classical and two quantum-mechanical. For each experiment, the participant was presented with a histogram of fictitious data hypothetically collected by students conducting the experiment.

The classical mechanics scenario was a ball-drop experiment adapted from the Physics Measurement Questionnaire [2]. In the scenario, 63 students released a ball from a ramp on a table, measured the landing position, and produced the blue histogram in Fig. 1a. Participants were asked to explain the shape and source of the distribution. They were then asked to compare the data to the result an instructor would obtain if they measured the release height and calculated the

TABLE I. The coding scheme used to characterize student ideas about sources of measurement uncertainty, shown in relationship to the Modeling Framework for Experimental Physics [10]. The term distribution here refers to any spread or variability in the data.

	Measurement model	**Physical model**
Limitations	Distribution due to mistakes caused by either the experimenters or measuring device. (ML)	Distribution due to invalid assumptions or variables not accounted for in the model. (PL)
Principles	Distribution due to inherent limitations in the measurement apparatus, i.e. measurement devices cannot be infinitely precise. (MP)	Distribution due to inherent features of the physical model, such as the uncertainty principle, wave functions, or unknowable "hidden variables." (PP)

predicted landing position. Finally, participants were asked to compare the original data to the red histogram in Fig. 1a, a data set hypothetically collected by another group of students.

The analogous quantum-mechanical scenario was a single-photon single-slit experiment that used the same data sets as the classical scenario (Fig. 1a). In this scenario, the histograms represented the position of 63 single-photon measurements on a screen. The interview questions had the same structure as for the ball-drop scenario, with students discussing the shape of the histogram, a comparison to a calculation of the diffraction pattern from a measurement of the slit width, and a comparison between two sets of student data.

Between the experiments above, students were asked about another quantum-mechanical experiment: measurements of the wavelength of light emitted from an object (fictional data shown in Fig. 1b). Participants were asked similar questions about the shape of the histogram as above but were not asked to compare to another data set or a theoretical calculation.

In addition to the specific scenarios, participants were asked about measurement uncertainty in classical and quantum mechanics without reference to particular experiments: "*What comes to mind when you think about measurement uncertainty in classical (quantum) mechanics?*" These questions were posed twice: once between the spectrum and single-slit experiments and once at the end of the interview.

B. Coding scheme

The codes for student-identified sources of uncertainty, chosen to align with the Modeling Framework for Experimental Physics [10] and shown in Table I, were Measurement model Limitations (ML), Measurement model Principles (MP), Physical model Limitations (PL), and Physical model Principles (PP). Measurement model limitations included human error (as in Ref. [14]) or inconsistencies between trials, as exemplified by "*if the measuring stick wasn't exactly like along the same line or axis every time where, or if it was slightly shifted.*" Measurement model principles referred to measuring devices not being infinitely precise, for example, "*So it would just be a bunch of results at the same position like plus minus whatever the, what's it called, whatever precision is a yard stick goes.*" Physical model limitations included confounding variables, such as friction, air resistance, and background light, and variables missing from the physical model. Example participant statements included "*There might be like the path that the ball takes would be different and there might be like different friction coefficients along the ramp*" and "*Well that probably means that there are other variables going on. You know, that the model is not entirely accounting for everything.*" Physical model principles included sources of variability inherent to the system, particularly due to the probabilistic nature of quantum mechanics: "*Uh, always talk about the Heisenberg uncertainty principle [...] And, uh, we're taught to think of it as a, um, intrinsic property of the theory of how nature, how it works.*" In the spectrum experiment, physical model principles also included reasons for the multiple peaks in the histogram, such as, "*So like the reason why there could be multiple ones are because [...] there's I guess an increments of energy levels that you can go.*"

We coded any statement about the sources of measurement variability throughout the interview. Some statements were too vague to categorize, but otherwise all statements about sources of uncertainty were codeable based on this scheme. Rather than count unique instances of each code for each student, we identified *whether* a student described each type of uncertainty source in Table I, and then analyzed the fraction of students that drew on each type throughout the interview.

Participants' responses were also coded for the *type* of uncertainty described. Statements addressing the source of spread, imprecision, or variability were coded as *precision*. Statements about sources of systematic error or inaccuracy were coded as *accuracy*. Because statements about accuracy appeared almost exclusively in comparisons of two data sets or comparisons of a data set to the value calculated from theory, the accuracy code was further divided into two codes designating the context: *data versus data* and *data versus theory*. Statements in response to the question "*What comes to mind when you think about measurement uncertainty in classical (quantum) mechanics?*" were coded as *generic measurement uncertainty*. Participants responded similarly both times this question was posed, so we collapsed the instances together.

The first and second authors independently coded the transcripts of all nineteen interviews. We calculated inter-rater reliability through Cohen's kappa for each source code across all types of uncertainty. The Cohen's kappa values were 0.857 for measurement limitations, 0.751 for measurement principles, 0.748 for physical limitations, 0.848 for physical principles. All coding disagreements were discussed and resolved.

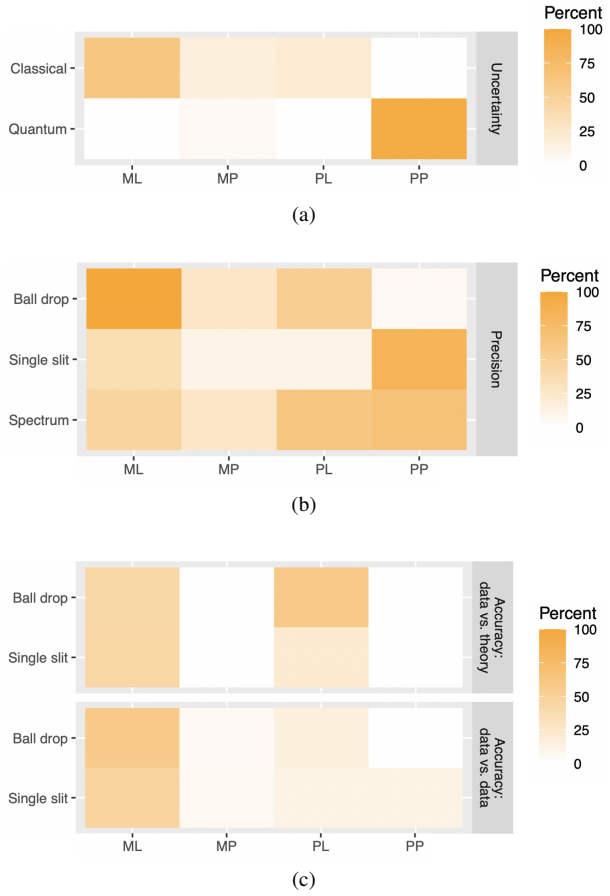


FIG. 2. Sources of (a) generic measurement uncertainty, (b) precision, and (c) accuracy mentioned by interview participants. Note that the source codes are not mutually exclusive; thus, the rows do not add up to 100%.

III. RESULTS

We present our results by each type of uncertainty and compare participants' responses between the classical and quantum contexts. We found, overall, that participants described different sources of uncertainty between the contexts.

A. Generic measurement uncertainty

Figure 2a shows the percent of participants who identified each type of source in response to the questions about generic uncertainty in classical and quantum mechanics. In *classical mechanics*, most participants focused on measurement limitations, with fewer than half identifying measurement principles or physical limitations and none mentioning physical principles. In contrast, when asked about measurement uncertainty in *quantum mechanics*, all but one of the participants attributed uncertainty to physical principles. Participants claimed that measurement uncertainty is inherent in the theory of quantum mechanics, with many explicitly referencing the Heisenberg uncertainty principle.

B. Precision

As with generic measurement uncertainty, the types of responses about precision, shown in Fig. 2b, varied between the contexts, but with new subtleties.

When asked about the ball-drop experiment (*classical*), all participants mentioned some form of measurement limitations. Compared with the generic uncertainty question, more participants also referenced physical limitations and measurement principles when considering the fictional data set. This suggest that participants considered more varied sources of uncertainty in response to the data than the generic question.

For the single-slit experiment (*quantum*), students' responses were also more varied when considering the data set than when considering measurement uncertainty in quantum mechanics generically. While sources of generic uncertainty in quantum mechanics were almost exclusively restricted to physical principles, explanations of the single-slit experiment data included physical principles (more than half the participants) and measurement limitations (fewer than half the participants), as well as a couple mentions of measurement principles and physical limitations.

For the spectrum experiment (*quantum*), participants' responses did not match their reasoning about either the ball-drop or single-slit experiments, and explanations were coded across all four categories. More than half the participants mentioned physical principles, alluding to the multiple peaks in the histogram as indicative of intrinsic properties of the material studied. However, more than half also cited physical limitations, with an emphasis on background light as a confounding variable. Several participants also referenced measurement limitations and measurement principles.

C. Accuracy

Participants' responses regarding accuracy are shown in Fig. 2c, separated by whether students were referring to the comparison between the two student data sets or between the student data set and the value calculated from theory.

Responses comparing data to theory were similar between the ball-drop (*classical*) and single-slit (*quantum*) contexts, with participants drawing only on measurement limitations and physical limitations. In both cases, measurement limitations were typically attributed to some form of human error. In the ball-drop experiment (*classical*), over half the participants mentioned physical limitations through confounding variables missing from the theory, such as friction or air resistance. In the single-slit experiment (*quantum*), only four participants mentioned physical limitations, either suggesting that the slit was too wide to produce a diffraction pattern, or that the hypothetical theory calculation would not account

for all variables present in the experiment. These similarities starkly contrast the distinct responses in the classical and quantum contexts in the previous sections.

In comparing two data sets, participants most frequently mentioned measurement limitations for both experiments, with physical limitations the next most common source. This suggests that participants attributed discrepancies between the data sets primarily to limitations, and that measurement limitations were more likely than physical limitations.

IV. DISCUSSION

In this study, we analyzed interviews with middle- and upper-division physics majors about sources of measurement uncertainty in classical and quantum-mechanical scenarios. The participants discussed measurement uncertainty differently in the two contexts, most evidently in their descriptions of generic measurement uncertainty. Consistent with the previous analysis of this data [14], we found that participants attributed generic classical uncertainty to a variety of measurement limitations, measurement principles, and physical limitations but almost exclusively discussed generic quantum uncertainty as inherent physical principles.

These views of measurement uncertainty were largely borne out in participants' discussion of precision in the ball-drop and single-slit data. However, more participants drew on additional types of sources of measurement uncertainty when discussing the single-slit experiment than when describing quantum uncertainty generically, and with higher frequency than the previous analysis found [14]. This is consistent with other work indicating that students draw on different forms of reasoning when responding to a specific scenario as compared to a decontextualized question [12].

Notably, the distinctions between classical and quantum contexts did not bear out in discussion of the spectrum experiment, where students drew almost equally on all type of uncertainty sources. Interestingly, a common explanation given for the spread in the spectrum data was background light, but no students drew on this idea in the single-slit context. One possible conclusion is that the participants did not think about the spectrum experiment as purely quantum-mechanical. This would align with previous work showing that many university students at a variety of educational levels apply mixed quantum and classical concepts in thinking about models that explain atomic spectra [7, 15, 16]. Indeed, participants drew on measurement principles and physical limitations in the spectrum experiment with similar frequency to the ball-drop experiment and on measurement limitations and physical principles similarly to the single-slit experiment.

Alternatively, the pattern in responses could be the result of cueing from the interview questions [17]. The spectrum experiment questions were posed immediately after the ball-drop experiment and before quantum mechanics was mentioned. Participants' thinking appears to have shifted discretely when quantum mechanics was explicitly mentioned in the questions about generic measurement uncertainty. That is, participants' ideas about uncertainty were more polarized when considering uncertainty generically than in response to hypothetical data sets. Future work should evaluate the role of framing and resources as students describe measurement uncertainty in different contexts. If student reasoning differs between contexts (classical vs. quantum and generically vs. in response to data), this would have implications for assessments of student thinking about measurement uncertainty.

Our analysis also sheds light on distinctions between student thinking about accuracy and precision in the classical and quantum contexts. Students frequently drew on measurement limitations in discussing both accuracy and precision across both contexts. This supports existing work that students tend to conflate sources of uncertainty and errors or mistakes [3, 4]. Furthermore, when comparing two data sets, participants tended to give similar answers about both the ball-drop and single-slit experiments. This is in stark contrast to the differences between the contexts with regard to precision and generic uncertainty. When considering a comparison between a data set and a hypothetical theory calculation, participants similarly mentioned measurement limitations with the same frequency for both experiments. However, participants drew on physical limitations, usually to question the completeness of the theory, with similar frequency as measurement limitations in the ball-drop experiment, while for the single-slit experiment, few participants drew on physical limitations. These responses are consistent with previous work indicating that upper-level students see theory both as validating and being validated by experiment [18, 19], though our analysis notably suggests that students are less likely to see theory as validated by experiment in a quantum context.

V. CONCLUSION

This study shed light on student thinking about measurement uncertainty in classical and quantum contexts, drawing on the Modeling Framework for Experimental Physics [10] to code student ideas about sources of measurement uncertainty. In addition to indicating that students think about measurement uncertainty quite differently across physics contexts and types of uncertainty, the results suggest that student thinking about uncertainty may be sensitive to cueing from questions. These results will inform the design of future assessments to probe ideas about sources of measurement uncertainty in a larger and broader sample of students.

ACKNOWLEDGMENTS

This work has been supported in part by the NSF under Grants No. DUE-1808945 and No. DUE-1809183. The authors would like to acknowledge the contribution of Martin Stein, who developed the protocol for and conducted the interviews used in this study.

[1] C. Wieman, Comparative Cognitive Task Analyses of Experimental Science and Instructional Laboratory Courses, The Physics Teacher **53**, 349 (2015).

[2] A. Buffler, S. Allie, and F. Lubben, The development of first year physics students' ideas about measurement in terms of point and set paradigms, International Journal of Science Education **23**, 1137 (2001).

[3] D. Evangelinos, D. Psillos, and O. Valassiades, An Investigation of Teaching and Learning about Measurement Data and their Treatment in the Introductory Physics Laboratory, in *Teaching and Learning in the Science Laboratory*, Contemporary Trends and Issues in Science Education, Vol. 16, edited by D. Psillos and H. Niedderer (Springer Netherlands, 2002) pp. 179–190.

[4] N. Holmes and C. E. Wieman, Assessing modeling in the lab: Uncertainty and measurement, in *2015 Conference on Laboratory Instruction Beyond the First Year of College*, edited by M. Eblen-Zayas, E. Behringer, and J. Kozminski (College Park, MD, 2015) pp. 44–47.

[5] I. D. Johnston, K. Crawford, and P. R. Fletcher, Student difficulties in learning quantum mechanics, International Journal of Science Education **20**, 427 (1998).

[6] G. Passante, P. J. Emigh, and P. S. Shaffer, Examining student ideas about energy measurements on quantum states across undergraduate and graduate levels, Physical Review Special Topics-Physics Education Research **11**, 020111 (2015).

[7] G. Kalkanis, P. Hadzidaki, and D. Stavrou, An instructional model for a radical conceptual change towards quantum mechanics concepts, Science Education **87**, 257 (2003).

[8] C. Singh, Student understanding of quantum mechanics, American Journal of Physics **69**, 885 (2001).

[9] B. W. Dreyfus, J. R. Hoehn, A. Elby, N. D. Finkelstein, and A. Gupta, Splits in students' beliefs about learning classical and quantum physics, International Journal of STEM Education **6**, 31 (2019).

[10] B. M. Zwickl, D. Hu, N. Finkelstein, and H. Lewandowski, Model-based reasoning in the physics laboratory: Framework and initial results, Physical Review Special Topics-Physics Education Research **11**, 020113 (2015).

[11] E. Brewe, Modeling theory applied: Modeling Instruction in introductory physics, American Journal of Physics **76**, 1155 (2008).

[12] J. Leach, R. Millar, J. Ryder, and M.-G. Séré, Epistemological understanding in science learning: the consistency of representations across contexts, Learning and Instruction **10**, 497 (2000).

[13] J. Ryder and J. Leach, Interpreting experimental data: the views of upper secondary school and university science students, International Journal of Science Education **22**, 1069 (2000).

[14] M. M. Stein, C. White, G. Passante, and N. G. Holmes, Student interpretations of uncertainty in classical and quantum mechanics experiments, in *2019 Physics Education Research Conference Proceedings* (American Association of Physics Teachers, 2019).

[15] R. Mueller and W. H., Students' conceptions of quantum physics research, in *Research on teaching and learning quantum mechanics: Papers presented at the annual meeting of the National Association for Research in Science Teaching* (1999).

[16] I. Testa, G. Capasso, A. Colantonio, S. Galano, I. Marzoli, U. S. di Uccio, F. Trani, and A. Zappia, Development and validation of a university students' progression in learning quantum mechanics through exploratory factor analysis and Rasch analysis, International Journal of Science Education **41**, 388 (2019).

[17] R. S. Russ, V. R. Lee, and B. L. Sherin, Framing in cognitive clinical interviews about intuitive science knowledge: Dynamic student understandings of the discourse interaction, Science Education **96**, 573 (2012).

[18] D. Hu and B. M. Zwickl, Examining students' views about validity of experiments: From introductory to Ph.D. students, Phys. Rev. Phys. Educ. Res. **14**, 010121 (2018).

[19] D. Hu and B. Zwickl, Examining students' personal epistemology: the role of physics experiments and relation with theory, in *Physics Education Research Conference 2017*, PER Conference Juried Paper (Cincinnati, OH, 2017) pp. 11–14.