Tracing Conceptual Overgeneralization in Middle School Students' Reasoning about Quantum Randomness

Jihyun Rho
University of Florida
j.rho@ufl.edu

Jiwon Kim
Purdue University
jwkim@purdue.edu

Zeynep G. Akdemir-Beveridge
University of Connecticut
zeynep.akdemir@uconn.edu

Muhsin Menekse
Southern Methodist University
menekse@smu.edu

ABSTRACT
students because it requires students' conceptual change between classical and quantum explanatory frameworks for randomness. Conceptual change in such contexts may not only involve replacing prior knowledge, but also risks overgeneralization, where newly learned concepts are extended beyond their intended scope. In this study, we investigate how students reorganize their reasoning about randomness across classical and quantum contexts following instruction. We analyzed 14,296 open-ended responses collected from pre-test and post-test assessments administered to 1,787 middle school students as part of a quantum-infused science curriculum. Student explanations were classified into classical, quantum, or non-informative reasoning using a large language model fine-tuned on a human-coded subset. A binomial mixed-effects model revealed asymmetric overgeneralization dynamics across problem contexts. Following instruction, overgeneralization decreased in quantum problem contexts, indicating improved alignment between quantum explanations and quantum phenomena. In contrast, overgeneralization increased in classical contexts, reflecting the extension of newly learned quantum reasoning into situations where classical explanations were sufficient. These findings highlight the bidirectional nature of conceptual overgeneralization and underscore the importance of instructional designs that support context-sensitive integration of new scientific concepts.

Keywords
Conceptual change, overgeneralization, quantum randomness, LLM-based classification.

1. INTRODUCTION
Research on conceptual change has emphasized that students' learning of new concepts is not a process of replacement, but one of gradual reorganization of existing knowledge structure [6,10]. Within this framework, overgeneralization can be understood as a pattern that emerges during conceptual change, reflecting students' attempts to coordinate prior knowledge and newly learned information.

Jihyun Rho, Jiwon Kim, Zeynep G. Akdemir-Beveridge, and Muhsin Menekse. Tracing Conceptual Overgeneralization in Middle School Students' Reasoning about Quantum Randomness. In Anthony Botelho, Maria Mercedes T. Rodrigo, Adish Singla, Hiroaki Ogata, Hyojeong So, and Young Hoan Cho (eds.) Proceedings of the 19th International Conference on Educational Data Mining, Seoul, Republic of Korea, June, 2026, pp. 578–584. International Educational Data Mining Society (2026).

https://doi.org/10.5281/zenodo.21039823

Prior work primarily framed overgeneralization as a phenomenon driven by prior knowledge constraining the interpretation of new information [19,20]. From this perspective, students rely on well-established prior knowledge to make sense of a new domain, sometimes leading to explanations that do not fully align with the new domain. However, this overgeneralization also occurs in the reversed direction. As students acquire new information, this newly learned information may itself be overextended, shaping or changing reasoning even in contexts where prior knowledge would otherwise be sufficient [22]. Overgeneralization, therefore, can reflect not only resistance to abandoning old frameworks but also the excessive generalization of newly acquired information beyond its intended scope.

This bidirectional view of overgeneralization is particularly relevant in domains where instruction introduces fundamentally different explanatory principles, such as quantum randomness. When learning quantum randomness, students need to reconcile probabilistic reasoning grounded in classical systems with non-classical sources of indeterminacy [4,17]. In this process, students may incorrectly apply classical randomness to a quantum context, but they may also extend quantum randomness back into classical situations, thereby reshaping their prior conceptual frameworks.

Despite its theoretical importance, overgeneralization during conceptual change has often been examined indirectly, primarily through accuracy-based measures. Such approaches can obscure whether students' responses reflect stable understanding, transitional reasoning, or systematic overextension across contexts. To address this gap, our study examines patterns of overgeneralizations across classical and quantum contexts before and after the instruction, focusing on how students' explanations shift across conceptual boundaries.

2. RELATED WORKS
2.1 Quantum Curriculum for Middle School
Recent years have seen growing interest in introducing quantum information science and engineering (QISE) concepts into K-12 education [15], driven largely by efforts to expand the future quantum workforce. National reports, including those from the National Quantum Coordination Office, have emphasized the importance of developing age-appropriate quantum learning pathways that integrate with existing science and computing curricula, rather than treating quantum topics as isolated and advanced content. In response to this need, initiatives such as The National Q-12 Education Partnership have emerged. However, their curricular recommendations primarily target the high school level. Correspondingly, existing empirical research on K-12 quantum education has

overwhelmingly focused on high school students (e.g., [5,16]), with comparatively little attention to middle school students.

Despite its imbalance, a growing body of evidence suggests that middle school students are cognitively capable of engaging with core ideas in modern physics, including foundational quantum concepts [1,7,13]. For example, prior work has demonstrated a saturation effect in post-test performance, indicating that no intrinsic cognitive barriers prevent middle school students from understanding key Einsteinian ideas such as curved spacetime or photon momentum [7]. Moreover, longitudinal findings show that students were able to recall central concepts (e.g., "space is curved," "light comes as photons") one to three years after instruction, suggesting durable conceptual understanding rather than short-term memorization [13]. Together, these findings challenge assumptions that quantum concepts are inherently inaccessible at the middle school level.

Building on this evidence, a small but growing literature on middle school quantum education has focused on two primary areas: (1) the development of effective pedagogical interventions and (2) teacher education and professional development for quantum curriculum implementation. A portion of this work has examined teacher preparation. Survey and interview studies with middle and high school teachers consistently show that the successful introduction of quantum topics depends heavily on teacher self-efficacy, conceptual confidence, and access to sustained support [2,11]. This highlights the role of professional development (PD) in co-designing their quantum curriculum [1]. Teachers often express strong motivation to engage with quantum science due to its perceived relevance and future-oriented appeal, while simultaneously reporting concerns about curricular alignment and their own content knowledge.

In parallel, research on instructional design has explored how middle school students can meaningfully engage with quantum concepts by deemphasizing mathematical formalism and instead foregrounding conceptual, experiential, and interactive learning modalities [7,13,14,21]. For instance, one study employed a gamified puzzle environment to teach quantum logic gates, finding that students required substantial scaffolding and frequently bypassed textual explanations, which motivated designs that embed learning directly into gameplay mechanics. Other work has leveraged interdisciplinary approaches, such as using music education to introduce quantum atomic models to 12-13-year-olds [21]. This interdisciplinary approach resulted in high post-test performance. Finally, the Einstein-First project exemplifies a complementary approach that replaces traditional Newtonian instruction with Einsteinian concepts through concrete physical models. Interventions in the project included using Nerf guns to represent photons (demonstrating momentum and uncertainty) and lycra sheets with marbles to visualize curved spacetime [7,13]. These studies support the feasibility of introducing modern physics concepts at the middle school level.

2.2 Teaching quantum randomness

However, demonstrating the feasibility of teaching quantum concepts at the middle school level does not imply that all quantum ideas are equally accessible. Among basic quantum concepts [9], the probabilistic nature of quantum phenomena is central for understanding the behavior of quantum entities. Prior research suggests that students often approach quantum physics with intuitions grounded in everyday experience, treating quantum entities as classical objects that follow definite trajectories through space and time [12]. This reliance on everyday experiences and classic objects creates substantial conceptual barriers when students encounter quantum phenomena that require fundamentally different explanatory assumptions.

Specifically, researchers investigated how students interpret probability and randomness in quantum contexts and categorized students' interpretations into three types, including a classic type, a mixed type, and a quantum type [17]. The classical type refers to students' interpretation where randomness is treated as epistemic uncertainty caused by incomplete information or imperfect measurement. The mixed type refers to an interpretation where students assume hidden information exists but is practically inaccessible. Lastly, in the quantum type, students treat individual outcomes as inherently unpredictable while recognizing that collective behavior can be predicted statistically. This framework highlights that the critical distinction between classical and quantum randomness lies not in whether individual outcomes are difficult to predict, but in the source of that unpredictability. In the classical view, unpredictability arises from incomplete information. Given sufficient knowledge of all relevant variables, outcomes are determinable, in principle. In the quantum view, unpredictability is intrinsic to the system. Even with complete knowledge of the system's state, individual outcomes remain fundamentally indeterminate.

2.3 Conceptual Change and Overgeneralization

Conceptual change research characterizes learning as a process of knowledge integration that unfolds gradually, during which students actively reorganize existing ideas rather than simply replacing them [6,10]. Within this framework, overgeneralization has been widely documented as a mechanism through which students extend familiar principles beyond their valid scope. Much of this work has focused on cases in which entrenched prior knowledge constrains the interpretation of new information, such as the application of whole-number reasoning to rational numbers [19], linearity assumptions in non-linear domains [20], or deterministic intuitions in probabilistic and stochastic systems [4]. In these accounts, overgeneralization is often interpreted as evidence of robust prior knowledge that resists accommodation. However, other strands of overgeneralization can also arise when newly learned concepts are overextended or inappropriately activated in contexts where previously established forms of reasoning would have been sufficient [4,8,22]. From this perspective, overgeneralization can reflect not only resistance to conceptual change, but also productive sense-making during the reorganization of knowledge structures, which occur in both directions. While prior knowledge constrains new learning, newly acquired concepts are overextended into familiar domains.

Thus, because conceptual change involves ongoing reorganization of knowledge, overgeneralization is often difficult to capture with summary performance measures alone. Accordingly, many conceptual change studies have examined students' strategies and explanations in depth. For example, prior research has analyzed students' written explanations, interview responses, and problem-solving justifications to characterize the reasoning strategies students apply across tasks [4,19,20]. However, such fine-grained analyses are typically time-intensive and therefore difficult to scale, particularly when studying large populations or multiple task contexts. As a result, relatively little work has systematically examined patterns of reasoning across contexts using large collections of student explanations.

Addressing this gap requires analytic approaches that can simultaneously preserve sensitivity to students' explanatory reasoning and operate at scale across multiple task contexts. Recent work in educational data mining has demonstrated the feasibility of using large language models to automatically assess rich, open-ended student data [3,18]. Therefore, in the present study, we adopt a Large Language Model-assisted analytic approach to examine large collections of student explanations across classical and quantum randomness contexts.

3. RESEARCH QUESTIONS

Building on prior works, the present study examines how students' reasoning about randomness reorganizes after their first formal exposure to quantum randomness in a middle school science setting. This investigation is situated within a larger evaluation of a quantum-infused curriculum that introduces quantum randomness through instruction on radioactive decay. This topic already involves classical probability within existing science standards. By comparing students' written explanations across pre- and post-test items designed to elicit classical and quantum interpretations of randomness, we analyze whether overgeneralization dynamics shift asymmetrically across contexts as students integrate new conceptual knowledge.

Accordingly, we address the following research question:

RQ) How do overgeneralization patterns differ across classical and quantum contexts before and after curriculum, and what asymmetries emerge in these dynamics?

4. DATASETS

4.1 Quantum-Infused Science Curriculum

We co-designed the 7-day-long quantum education curriculum for middle school students with middle school science teachers. We aim to align the curriculum with classroom constraints and be accessible to typical middle school students. The first half of the curriculum introduces randomness in an Earth and Space Science context, starting with instruction on radioactive decay and radiometric dating. Students begin with an archaeology-themed challenge (e.g., estimating the age of remains) and then learn how Carbon-14 decays into Nitrogen-14 over time, using half-life as a statistical model for decay rates rather than a deterministic prediction of individual events. In the following lessons, students engage with isotopes, decay types, and hands-on modeling activities (e.g., popcorn and coin-flip analogies) to build an intuitive distinction between what can be predicted in aggregate (i.e., the decay rate) versus what cannot be predicted at the level of a single atom (i.e., which nucleus decays next). At this stage, students are expected to understand this unpredictability as classical randomness, arising from practical complexity (i.e., too many variables to track). They are not yet expected to understand any fundamental indeterminacy in the system itself.

The second half of the curriculum then extends beyond existing standards-based science content by explicitly introducing the concept of quantum randomness. They now explore randomness in the context of quantum indeterminacy and learn how quantum random number generators produce unpredictable outcomes that cannot be explained by hidden variables or measurement limitations. Students apply this idea in an activity that involves generating binary numbers, converting them to hexadecimal codes using real quantum-generated data from an online quantum lab, creating random colors, and using them in their artwork. This final lesson positions quantum randomness as an extension of students' prior reasoning about probability, while discussing boundary conditions for when classical explanations are sufficient versus when quantum explanations are needed.

4.2 Participants and Dataset

Participants were middle school students enrolled in regular science classes at seven public middle schools in the Midwestern United States. Across the schools, a total of 1787 students participated in both the pre-test and post-test. Approximately 25.4 % of students were in 6th grade, 37.0% were in 7th grade, and 37.6% were in 8th grade. Female students comprised 49.3% of the sample, while male students comprised 50.7%. Students were recruited through classroom participation during regular instructional activities, and no prior instruction on quantum randomness was assumed.

Each student responded to four assessment items (i.e., two items asking the classical randomness concept and two items asking the quantum randomness concept) in the pre-test and the same four items in the post-test, resulting in a total of 8 responses per student. The example of the question is shown in Figure 1 below. Across all participants, this yielded 14,296 total responses in the full dataset used for large-scale analysis. This study was reviewed and approved by the Institutional Review Board at Purdue University (IRB-2022-1016). We randomly selected a subset of responses from 100 students for detailed human coding and model calibration. This subset comprised 800 responses (4 items × pre/post × 100 students). We used this subset to establish coding reliability, train, and evaluate automated classification models. The remaining responses formed the dataset for population-level analyses of conceptual reasoning and overgeneralization dynamics.

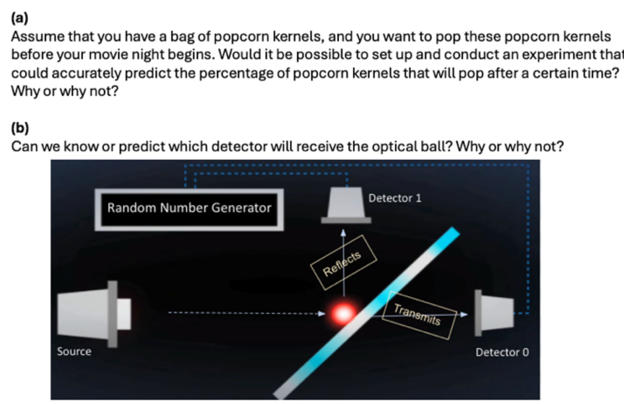

Figure 1. Example test items used in both pre-test and post-test. (a) is a classical randomness item asking whether the outcome of popcorn popping can be accurately predicted in advance and why. (b) is a quantum randomness item depicting a single-photon beam splitter setup, asking whether one can know or predict which detector will register the optical signal and why.

5. ANALYSIS

5.1 Coding Scheme

We developed the coding scheme through an iterative, theory-informed process. We first derived an initial set of codes from the theoretical distinction between classical randomness and quantum randomness, informed by prior research on middle school students' understanding of quantum randomness [17]. We then applied these preliminary codes to a subset of student responses to identify ambiguities, overlaps, and borderline cases. Through iterative discussion, we refined the code definitions to ensure conceptual clarity, mutual exclusivity, and applicability across both classical and quantum contexts. The refinement was done with the understanding that coding of student responses did not aim to assess correctness per se, but rather to characterize the dominant conceptual basis underlying each student's response. Crucially, the coding distinguished responses based on the attributed source of unpredictability, rather than on the acknowledgment of unpredictability itself. For example, in the popcorn item in Figure 1, attributing unpredictability to practical complexity (e.g., varying moisture or heat) was coded as classical, whereas treating individual outcomes as fundamentally indeterminate was coded as quantum. The latter constitutes overgeneralization because popcorn popping is governed by deterministic physical processes that are, in principle, predictable given sufficient information. This process yielded a three-category coding scheme, summarized in Table 1.

Using this coding scheme, two trained researchers independently coded a subset of students' open-ended responses. This human-coded dataset consisted of responses from 100 students, each answering four items in both pre-test and post-test, yielding a total of 800 responses. The overall inter-rater reliability (Cohen's κ) was κ = 0.8227, indicating strong agreement beyond chance. Specifically, 0.8143 for classical, 0.8450 for quantum, and 0.8227 for non-informative.

5.2 LLM-Based Response Classification

To enable large-scale analysis of student reasoning, we used a large language model (LLM) to classify the remaining student responses that were not manually coded. We fine-tuned a pretrained instruction-following LLM (*Meta-LLaMA-3.1-8B-Instruct*) using parameter-efficient fine-tuning with low-rank adaptation (LoRA). We split the human-coded student responses described above into training and validation sets. The model was trained on this training set.

We evaluated the trained model's performance using the validation set exclusively held out for model evaluation. To account for class imbalance across conceptual categories, we report both macro-averaged and weighted F1 scores, along with category-level precision and recall. After validation, we applied the trained model to fully code the remaining full dataset of 13,078 responses after excluding empty responses.

5.3 Operationalizing Overgeneralization

We operationalized overgeneralization as cross-contextual misalignment between the problem context and the category of conceptual reasoning expressed in student responses. For each response, we compared the coded conceptual category with the problem context. We coded overgeneralization as present when a student applied quantum reasoning to a classical problem context or classical reasoning to a quantum problem context.

We excluded responses coded as non-informative from overgeneralization analyses because they did not provide sufficient evidence of a coherent conceptual framework. This binary overgeneralization variable served as the primary outcome in subsequent statistical analyses examining change over test time (i.e., pre-test and post-test) and problem contexts.

5.4 Statistical Analysis

To examine changes in overgeneralization across test time (i.e., pre-test and post-test) and problem contexts, we employed a binomial mixed effects modeling approach. The dependent variable was the binary overgeneralization indicator. Fixed effects included the test time (pre vs. post), problem context (classical vs. quantum), and their interaction. Random intercepts were included for students and test items to account for repeated measurements and problem-level variability.

Table 1. Three-category coding scheme used to classify students' reasoning of randomness in classical and quantum test items.

Code	Definition	Example
Classical	Randomness is treated as epistemic rather than fundamental. Outcomes are assumed to be determined by underlying physical causes and, in principle, predictable given complete information. Uncertainty is attributed to limited knowledge, measurement constraints, or statistical aggregation.	"Yes, you would have to do a lot of retesting and starting the whole experience over again. You could look in your microwave and see how many pop in a minute."
Quantum	Randomness is treated as fundamental to the physical system. Even with complete knowledge of the system, the outcome of an individual measurement is considered inherently unpredictable, reflecting intrinsic indeterminacy rather than lack of information.	"We cannot predict which detector will receive the optical ball. It is random, so we don't know if detector 1 or detector 0 will get it."
Non-Informative	Responses that are off-topic, vague, or non-conceptual, including expression of uncertainty or yes/no answers without explanatory reasoning.	"I don't know", "idk", "No, because I don't want to wait to watch the movie."

6. RESULTS

6.1 Validating LLM-based Classification Against Human Annotation

Before analyzing the overgeneralization pattern in the full dataset, we assessed the LLM-based classification performance against the human-coded validation set. The model achieved an accuracy of 0.8841, a macro F1 score of 0.8316, and a weighted F1 score of 0.8803. The code category-specific level's F1, precision, and recall scores were shown in Table 2. Precision and recall scores indicated that the model was most reliable in distinguishing students' classical reasoning and quantum reasoning.

Table 2. Classification performance of the LLM-based model by conceptual category on the human-coded validation set.

Code	Precision	Recall	F1
Classical	0.885	0.955	0.919
Quantum	0.750	0.600	0.667
Non-Informative	0.962	0.862	0.909

6.2 Distributed Patterns of Conceptual Reasoning

Prior to instruction, students' pre-test responses in the classical problem context were predominantly classified as classical reasoning (75.32%), with only a smaller proportion reflecting quantum reasoning (4.83%). In contrast, responses to the quantum problem context showed a substantial reliance on classical reasoning (68.13%), indicating that students largely interpreted quantum phenomena using classical conceptual reasoning before the instruction. Table 3 summarizes the distribution of conceptual reasoning across problem context and test times.

Table 3. Pre-test and post-test distributions of classical, quantum, and non-informative reasoning by problem context.

Problem context	Test time	Classical	Quantum	Non-Informative
Classical	Pre	75.32%	4.83%	19.85%
	Post	64.78%	15.45%	19.77%
Quantum	Pre	68.13%	14.44%	17.43%
	Post	56.49%	29.23%	14.28%

Following the instruction on quantum randomness, the overall distribution of conceptual reasoning shifted in the post-test (Figure 2). In the quantum problem context, the proportion of quantum reasoning increased to 29.23%, accompanied by a decrease in classical reasoning to 56.49%, suggesting increased uptake of non-classical reasoning.

Notably, a parallel shift also occurred in the classical problem context. Despite the classical nature of the problems, the proportion of responses classified as quantum reasoning increased from 4.83% in the pre-test to 15.45% in the post-test.

6.3 Overgeneralization Dynamics Across Problem Context and Test Time

Our binomial mixed-effects model revealed meaningful effects of both test time and problem context on students' overgeneralization patterns. First, the main effect of time showed a positive posterior mean (β = 1.31, 95% credible interval = [1.24, 1.37]), indicating that, on average, overgeneralization became more likely following instruction. The 95% credible interval for this effect did not include zero, providing posterior evidence for an overall increase in overgeneralization across problem contexts. The main effect of context also showed a positive posterior mean (β = 4.22, 95% credible interval = [4.15, 4.29]), suggesting substantially higher baseline overgeneralization rates in the quantum problem context compared to the classical problem context.

Critically, the interaction between test time and problem context showed a negative posterior mean (β = -2.22, 95% credible interval = [-2.31, -2.13]). This pattern indicates that instructional effects on overgeneralization diverged across problem contexts. Specifically, while overgeneralization decreased in the quantum problem context following instruction, it increased in the classical problem context, reflecting a problem-context dependent reorganization of students' conceptual reasoning. These posterior estimates from the mixed-effects model are reported in Table 4.

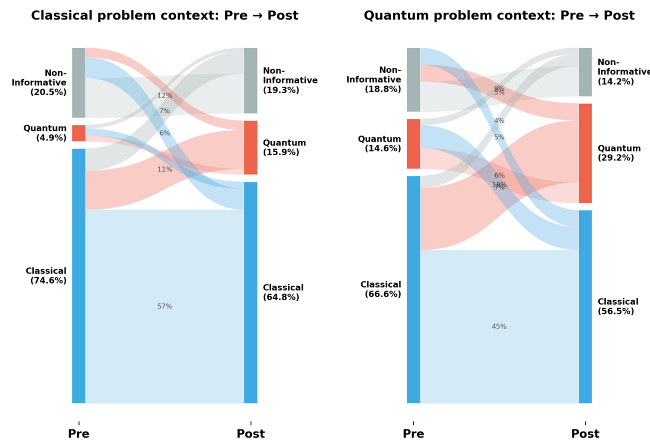

Figure 2. Sankey plot showing transitions in students' response from pre-test to post-test in classical (left) and quantum (right) problem contexts. Each band represents the flow of responses from a pre-test to a post-test.

Table 4. Posterior estimates from the binomial mixed-effects model predicting overgeneralization.

	Posterior mean (β)	Posterior standard deviation	95% Credible interval (Crl)
Intercept	-2.680	0.028	[-2.74, -2.63]
Test time	1.305	0.034	[1.24, 1.38]
Problem context	4.217	0.035	[4.15, 4.29]
Test time × Problem context	-2.219	0.045	[-2.31, -2.13]

Figure 3 summarizes model-predicted probabilities of overgeneralization across test time and problem context, derived from the binomial mixed-effects model and transformed from the logit scale. In the classical problem context, predicted overgeneralization increased from 6.42% in pre-test to 20.19% in post-test. In contrast, in the quantum problem context, predicted overgeneralization decreased from 82.31% in the pre-test to 65.11% in the post-test. These estimates reveal a qualitative asymmetry. This pattern indicates an asymmetric shift in explanatory tendencies across problem contexts: post-instruction responses showed reduced context-reasoning mismatch in the quantum problems but increased mismatch in classical problems. These changes mirror the interaction effect observed in the mixed-effects model, suggesting that shifts in overgeneralization differed systematically by problem contexts.

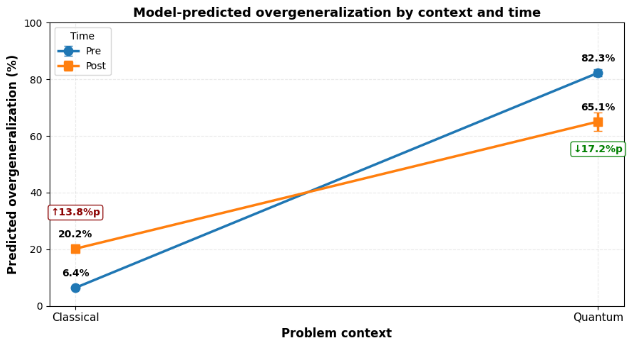

Figure 3. Model-predicted overgeneralization rates by problem context (classical vs. quantum) and time (pre- vs. post-test). Annotations indicate the change in percentage points from pre- to post-test within each context.

7. DISCUSSION

This study examined how middle school students' reasoning about randomness reorganizes across classical and quantum problem contexts following the curriculum implementation, with a particular focus on overgeneralization. From an educational data mining perspective, our work demonstrates how large-scale analyses of open-ended student explanations can uncover systematic patterns of conceptual change and overgeneralization that are challenging to observe through correctness-based metrics alone. By integrating LLM-based classification with mixed-effects modeling, we illustrate how EDM methods can be used to model fine-grained shifts in students' explanatory reasoning across contexts and time in quantum education.

Firstly, our results provide empirical support for a bidirectional account of overgeneralization during conceptual change in quantum education, revealed through large-scale analysis of open-ended responses. Overgeneralization of classical reasoning in quantum problem contexts decreased after students completed the curriculum, while overgeneralization of quantum reasoning in classical contexts increased. Prior work has largely framed overgeneralization as the transfer of entrenched prior knowledge into newly learned domains (e.g., [18, 19]). Consistent with this view, students initially relied heavily on classical reasoning when interpreting quantum randomness. Following the curriculum, such an initial reliance decreased in quantum contexts, indicating improved understanding of quantum randomness. However, newly acquired quantum reasoning was increasingly applied even to classical contexts where classical explanations were sufficient. From an EDM standpoint, this asymmetric pattern represents the value of modeling context-dependent reasoning trajectories rather than treating learning gains as uniform improvements. Large-scale classification of explanations enables detection of such directional shifts, which would likely be obscured in accuracy-only analyses.

Secondly, the observed bidirectional overgeneralization has implications for data-informed curriculum design in middle school quantum education. The increase in quantum reasoning in classical contexts suggests that students actively integrate newly learned frameworks into existing knowledge structures, even when this integration exceeds normative boundaries. Similar challenges in distinguishing explanatory frameworks have been reported in prior studies of quantum reasoning and conceptual change (e.g., [3, 16]). In our curriculum, classical and quantum randomness were intentionally taught together to align with state standards and the pre-set annual teaching plan, with instruction beginning from classical probability and concluding with quantum randomness. While such a curricular design successfully and practically promoted quantum reasoning, it may also have increased the likelihood of quantum overgeneralization in classical contexts. These findings suggest that EDM approaches can play a critical role in evaluating instructional trade-offs and thus help curriculum designers identify when instructional success in one domain introduces unintended conceptual spillover into another. Explicit scaffolds that clarify boundary conditions may therefore be necessary to stabilize students' reasoning across classical and quantum science contexts.

Last but not least, this study demonstrates the methodological potential of LLM-assisted educational data mining for studying conceptual change using open-ended student responses. Prior work has shown that LLMs can reliably assess rich student explanations at scale (e.g., [2, 17]). Building on this work, we show how validated LLM-based classifications can be integrated with statistical modeling to examine population-level patterns of reasoning reorganization. This approach is particularly valuable for underrepresented student populations, such as middle school students in quantum education, where large datasets of open-ended responses have rarely been explored and are difficult to analyze manually.

Nevertheless, limitations remain. First, the three-category coding scheme, while aligned with the theoretical distinction central to our research question, may not capture finer-grained variations such as mixed or transitional reasoning. To capture such details of students' conceptual trajectories, future studies could explore how to sophisticate coding schemes. Second, automated classification can introduce uncertainty into coding results, particularly for the quantum category, where the model achieved a recall of 0.60, suggesting that quantum reasoning may be underestimated in our analyses. This limitation likely makes our central finding conservative, as the true extent of quantum reasoning in classical contexts may exceed what is reported. Future work could address this through different strategies, such as oversampling the quantum minority class during training and applying class-weighted loss functions to improve classification performance under class imbalance. Lastly, our analysis focused on short-term instructional effects; longitudinal EDM analyses are needed to examine how overgeneralization patterns evolve over longer timescales and across instructional contexts.

8. CONCLUSION

Our study contributes to research on conceptual change by empirically demonstrating that overgeneralization during learning is not a unidirectional phenomenon but can occur as a bidirectional phenomenon. In the context of learning quantum randomness, students showed both reduced inappropriate reliance on classical reasoning in quantum problem contexts and increased application of quantum reasoning in classical problem contexts following instruction. These asymmetric shifts highlight that conceptual change involves not only replacing or constraining prior knowledge but also negotiating the boundaries of newly acquired concepts. From an educational perspective, our findings suggest that successful instruction should be evaluated not only by gains within the target domain but also by how newly learned concepts interact with students' existing knowledge across contexts. Specifically, for quantum education at the middle school level, this implies a need for instructional designs that explicitly address the scope and limits of quantum explanations, helping students differentiate when quantum reasoning is appropriate and when classical explanations remain sufficient.

9. ACKNOWLEDGEMENTS

This study is supported by the National Science Foundation (Award Number: 2422937) and the National Defense Education Program (Award Number: HQ00342110014). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the National Defense Education Program.

10. REFERENCES

1. Zeynep Gonca Akdemir. 2024. Design, Implementation, and Evaluation of a Quantum-Infused Middle-School Level Science Teaching and Learning Sequence.
2. Zeynep Gonca Akdemir, Muhsin Menekse, Erica W Carlson, Nicholas Dang, Mahdi Hosseini, and Dongyang Li. 2024. Supporting Middle School Students' Learning Outcomes and Engagement with NGSS-Aligned Quantum-Infused Science Curriculum. *ASEE Annual Conference & Exposition.* 10.18260/1-2--48039
3. Sami Baral, Eamon Worden, Wen-Chiang Lim, Zhuang Luo, Christopher Santorelli, and Ashish Gurung. 2024. Automated Assessment in Math Education: A Comparative Analysis of LLMs for Open-Ended Responses. 732–737.
4. T. Bouchée, L. de Putter - Smits, M. Thurlings, and B. Pepin. 2022. Towards a better understanding of conceptual difficulties in introductory quantum physics courses. *Studies in Science Education* 58, 2: 183–202.
5. Berit Bungum, Maria Vetleseter Bøe, and Ellen Karoline Henriksen. 2018. Quantum talk: How small-group discussions may enhance students' understanding in quantum physics. *Science Education* 102, 4: 856–877.
6. Michelene TH Chi. 2009. Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In *International handbook of research on conceptual change*. Routledge, 89–110, Routledge.
7. Rahul K Choudhary, Alexander Foppoli, Tejinder Kaur, David G Blair, Marjan Zadnik, and Richard Meagher. 2018. Can a short intervention focused on gravitational waves and quantum physics improve students' understanding and attitude? *Physics Education* 53, 6: 065020.
8. Yi Ding, Guangtian Zhu, Qiuxin Bian, and Lei Bao. 2024. Analysis of students' conceptual change in learning Newton's third law with an integrated framework of model analysis and knowledge integration. *Physical Review Physics Education Research* 20, 2: 020141.
9. Directorate-General for Communications Networks, Content and Technology (European Commission), Quantum Flagship, Greinert, Franziska, and Müller, Rainer. 2025. *European competence framework for quantum technologies (CFQT): reference framework for planning, mapping and comparing QT related educational activities, personal qualification and job requirements*. Publications Office of the European Union. Retrieved January 28, 2026 from https://data.europa.eu/doi/10.2759/8917117
10. Andrea A DiSessa. 2014. *A history of conceptual change research*.
11. Apekshya Ghimire, Jaya Shivangani Kashyap, Emily Edwards, Diana Franklin, and Chandralekha Singh. 2025. Investigating high school and pre-high school teachers' perceptions and experiences introducing quantum concepts: a survey of QuanTime and other quantum-related activities. *EPJ Quantum Technology* 12, 1: 89.
12. Fabian Hennig, Kristóf Tóth, Moritz Förster, and Philipp Bitzenbauer. 2024. A new teaching-learning sequence to promote secondary school students' learning of quantum physics using Dirac notation. *Physics Education* 59, 4: 045007. https://doi.org/10.1088/1361-6552/ad353d
13. Tejinder Kaur, David Blair, Warren Stannard, David Treagust, Grady Venville, Marjan Zadnik, Warwick Mathews, and Dana Perks. 2020. Determining the Intelligibility of Einsteinian Concepts with Middle School Students. *Research in Science Education* 50, 6: 2505–2532.
14. Tianle Liu, David Gonzalez-Maldonado, Danielle B. Harlow, Emily E. Edwards, and Diana Franklin. 2023. Qupcakery: A Puzzle Game that Introduces Quantum Gates to Young Learners. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1* (SIGCSE 2023), 1143–1149. https://doi.org/10.1145/3545945.3569837
15. Christopher Monroe, Michael G Raymer, and Jacob Taylor. 2019. The US national quantum initiative: From act to action. *Science* 364, 6439: 440–442.
16. Sara Satanassi, Paola Fantini, Roberta Spada, and Olivia Levrini. 2021. Quantum Computing for high school: an approach to interdisciplinary in STEM for teaching. *Journal of Physics: Conference Series* 1929, 1: 012053. https://doi.org/10.1088/1742-6596/1929/1/012053
17. Kristóf Tóth, Fabian Hennig, Joaquin Marc Veith, and Philipp Bitzenbauer. 2025. Secondary school learners' reasoning about quantum randomness in the context of single-photon interferometer and double-slit experiments. *Eurasia Journal of Mathematics, Science and Technology Education* 21, 8: em2686.
18. Nhat Tran, Benjamin Pierce, Diane Litman, Richard Correnti, and Lindsay Clare Matsumura. 2024. Analyzing Large Language Models for Classroom Discussion Assessment. https://doi.org/10.48550/arXiv.2406.08680
19. Xenia Vamvakoussi and Stella Vosniadou. 2004. Understanding the structure of the set of rational numbers: a conceptual change approach. *Learning and Instruction* 14, 5: 453–467. https://doi.org/10.1016/j.learninstruc.2004.06.013
20. Wim Van Dooren, Dirk De Bock, An Hessels, Dirk Janssens, and Lieven Verschaffel. 2004. Remedying secondary school students' illusion of linearity: a teaching experiment aiming at conceptual change. *Learning and Instruction* 14, 5: 485–501. https://doi.org/10.1016/j.learninstruc.2004.06.019
21. Hugo Vieira and Carla Morais. 2022. Musical analogies to teach middle school students topics of the quantum model of the atom. *Journal of Chemical Education* 99, 8: 2972–2980.
22. Joseph Jay Williams, Tania Lombrozo, and Bob Rehder. 2013. The hazards of explanation: Overgeneralization in the face of exceptions. *Journal of Experimental Psychology: General* 142, 4: 1006.