

## **Exploratory factor analysis of the QMCA**

Adam Quaal and Gina Passante

*Department of Physics, California State University Fullerton, 800 N. State College Blvd., Fullerton, CA, 92831*

Steven J. Pollock

*Department of Physics, University of Colorado Boulder, 390 UCB, Boulder, CO, 80309*

Homeyra R. Sadaghiani

*Department of Physics and Astronomy, California State Polytechnic University Pomona, 3801 W. Temple Ave., Pomona, CA, 91768*

This investigation is situated within an ongoing project which seeks to understand student thinking in upper-division introductory quantum mechanics courses. Recently, the Quantum Mechanics Concept Assessment (QMCA) was revised to include additional items in spin-basis contexts to reflect the rising prevalence of the “spins-first” instructional paradigm. In this work, we utilize exploratory factor analysis to group items on the QMCA based on common variance. Student responses were collected from several large institutions over the 2018-2019 academic year, with the three largest institutions following a “spins-first” curriculum. In interpreting our factor structure, we focus on the placement of isomorphic questions and the original concept framework of the QMCA, as well as a tentative interpretation of factor groupings. We conclude by discussing how these groupings may be further investigated, as well as implications for subsequent iterations of the QMCA and research on student thinking in these two contexts.

## I. INTRODUCTION

This investigation is situated within a larger, ongoing project which seeks to understand student thinking in upper-division introductory quantum mechanics (QM) courses. Specifically, we aim to differentiate between student performance on multiple-choice questions in *position* and *spin contexts* to provide insight into how students approach comparable problems in each context. Position-context questions refer to those that assess student understanding of QM concepts through position-basis scenarios and notation, while spin-context questions assess the same concepts in the spin-basis. We believe this difference in context is important for two reasons. First, these question contexts reflect two common introductory QM instructional paradigms, known as *position-first* and *spins-first*. The former refers to the instructional paradigm that begins with the Schrödinger equation and its solution in potential wells [1]; the latter paradigm begins with sequential Stern-Gerlach experiments in spin-1/2 systems [2]. Second, while the concepts assessed by these questions and the mathematical procedures to solve them are similar, it is not clear if undergraduates at this level tend to view them as such. Research on students' conceptual difficulties in position contexts is well-established [3–7], but literature on difficulties in spin contexts is relatively sparse.

The instrument we use to explore the distinction between students' answers to items in the position-basis and the spin-basis is the Quantum Mechanics Concept Assessment (QMCA). The QMCA was initially developed and validated with items written primarily in position contexts [8]. In recent years, the QMCA has been modified with the goal of moving toward a concept inventory that is appropriate for instructors across curricular contexts. One major modification was the addition of several questions in the spin context that are analogous to questions in the position contexts. We refer to these questions as *isomorphic* items. Based on their conceptual and structural similarities, we would expect an expert to view two isomorphic questions as measuring the same broad concept.

Results from prior administrations of this modified QMCA show comparable averages on seven of eight isomorphic pairs [9]. However, similar aggregate scores on these items do not necessarily indicate a pattern in the paired responses of individual students. Though these questions are authored as isomorphic, it is not clear that students answer them in the same manner. Furthermore, the way students answer these isomorphic items may provide insight into important conceptual difficulties that may be more pronounced in either the position or spin context. These distinctions could impact the use of the QMCA across instructional contexts, as well as its continued development.

In this paper, we utilize exploratory factor analysis (EFA) to group items on the QMCA and use those groupings to postulate how students answer isomorphic questions. EFA is a data reduction technique that measures the common variance among a set of items and groups them accordingly into *factors*. This method has been used previously in physics

education research to determine the extent to which student performance on an instrument aligns with the content framework the authors intended to assess [10, 11]. We are particularly interested in two implications of the QMCA's factor structure. First, whether isomorphic pairs appear in the same factor groupings, which would suggest that students answer these items in a similar manner. Second, whether the factor structure matches the five concept domains outlined by the authors of the QMCA [8]. We conclude with a discussion of the implications of our factor structure to subsequent iterations of the QMCA and the study of student thinking in these two QM contexts.

## II. REVIEW OF PRIOR WORK

In Sadaghiani's and Pollock's QMCA development and validation study, items on the concept inventory were classified into five domains [8]:

- Measurement (M)
- Time-independent Schrödinger equation (S)
- Time evolution (T)
- Wave functions or boundary conditions (W)
- Probability or probability density (P)

These reflect a faculty consensus of essential content that should be covered in an introductory-level QM course. A single item may fall into multiple domains, so these domains are not mutually exclusive. In addition, these domains were determined by faculty consensus when the QMCA contained primarily position-context questions. Isomorphic spin-context items that were added to the QMCA were largely classified as "measurement" questions.

Previous work on the modified QMCA shows that students exhibit comparable performance on eight of the nine isomorphic pairs of questions [9]. However, these descriptive statistics do not show whether students approach these questions *consistently*. Isomorphic items on the QMCA have isomorphic responses to the greatest extent possible (see one such example in Fig. 1). A student who is responding consistently to both would choose the same isomorphic response for both items irrespective of whether they are correct. Comparing every set of isomorphic response for every pair of isomorphic questions would require substantially larger sample sizes than we have collected. The results of our EFA allow us to see simply whether students are consistently correct or incorrect, since we expect consistently answered questions to fall into the same factor grouping.

Our decision to utilize EFA was motivated by a desire to understand the construct validity of the QMCA in its current state of development. Classical test theory distinguishes between several types of test validity, two of which are *content* validity and *construct* validity [12]. Content validity measures the extent to which an assessment tool covers the content domains it intends to measure. For the QMCA, the five concept domains determined by faculty consensus give us a rough idea of the assessment's content validity. Construct

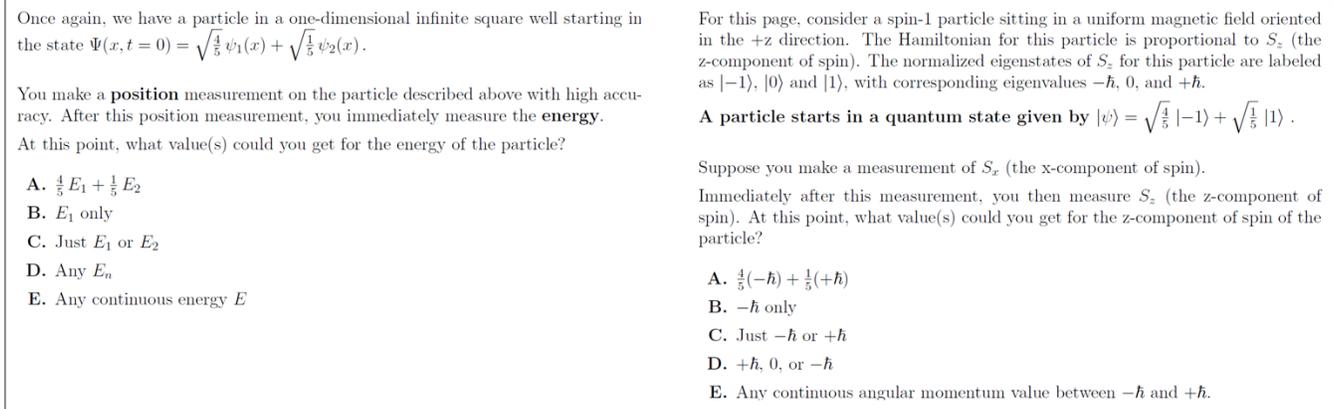


FIG. 1. A pair of isomorphic questions in position (left) and spin (right) contexts, with scenarios and responses constructed using comparable notation and values.

validity is a measure of what constructs are truly measured by student responses—for example, students’ understanding of time evolution or mathematical reasoning abilities. While content validity is important in choosing a concept inventory to administer, construct validity determines how student responses can be interpreted and used to inform instructional practices.

**III. METHODOLOGY**

**A. Research context**

The QMCA scores in this analysis were collected in the fall 2018 and spring 2019 academic terms, primarily at three institutions in the US. One of these institutions is a large, doctoral-granting institution with very high research activity; the other two are large, primarily nonresidential public universities with high undergraduate enrollment and some post-baccalaureate programs. Instructors at these three sites taught their introductory QM courses with the spins-first instructional paradigm via McIntyre’s textbook [2]. In addition, a small number of scores were collected from administrations at several pilot sites. Collectively, these administrations provided our sample size ( $N = 281$ ).

**B. Exploratory factor analysis**

Exploratory factor analysis is a statistical method often employed on multiple-choice assessment instruments to reduce data into groups, or factors, based on their common variance [13]. These factors represent latent variables which cannot be directly measured, such as student understanding of a specific concept (e.g., quantum measurement). Our analyses were run in the FACTOR program [14]; this program was chosen due to its ability to compute tetrachoric correlation matrices, which

are necessary when conducting an EFA with binary data. We used parallel analysis determine the optimal number of dimensions, and our factor rotation was carried out using the normalized direct oblimin method. We chose an oblique rotation method because we do not expect factors corresponding to student understanding of content domains to be independent in this context. Our minimum factor loading was set to 0.30. An item’s factor loading reflects the correlation between an item and the factor into which it is loaded.

Since our correlation matrix was *not positive definite* due to a small sample size, the FACTOR program allowed us to apply a smoothing algorithm [15]. However, this algorithm destroyed a substantial amount of covariance in the process, which required us to remove several items that were highly correlated to other items in order to generate a correlation matrix acceptable for EFA. To accomplish this, we examined several items on the QMCA which occur in sequential pairs. These pairs are not the same as isomorphic questions since these sequential pairs occur within the same position or spin context. Typically, the first item in a sequential pair poses a “Yes/No” question and the subsequent item asks for a follow-up rationale. Items in these pairs are highly correlated for reasons we can explain without EFA, so we removed the “Yes/No” question from each of these sets prior to re-running the EFA. This decision ultimately excluded five questions from the 38-item assessment, with two of those five questions comprising an isomorphic pair. With these items removed, the correlation matrix returned a Kaiser-Meyer-Olkin (KMO) value of 0.88. The KMO measure of sampling adequacy measures the amount of variance that might be attributable to latent factors; values above 0.80 are considered “meritorious” for factor analysis [16].

**IV. RESULTS**

The factor structure we obtained is shown in Table I, with four factors explaining about 42% of the variance in scores

on the QMCA. For reasons explained above, this structure does not include five of the “Yes/No” questions (4, 14, 20, 25, 37) on the QMCA. In addition, seven items (3, 12, 13, 26, 32, 33, 38) simply did not load into any factor, suggesting that they do not explain a significant amount of variance in student responses to the QMCA.

TABLE I. Factor structure of the QMCA <sup>a</sup>.

Item	Factor Loading	Context	Domain(s)	Isomorph
<b>FACTOR 1 (22.8% of variance)</b>				
1	0.67	Position	M	22
2	0.58	Position	M, P	23
8	0.54	Position	M	29
17	0.41	Position	T, W, P	
18	0.53	Position	T, W, P	
19	0.52	Position	T, W, P	
22	0.55	Spin	M	1
23	0.83	Spin	M	2
29	0.50	Spin	M	8
31	0.47	Position	S, W	
34	0.50	Spin	M	
35	0.33	Spin	M, W	
36	0.49	Spin	M	
<b>FACTOR 2 (6.7% of variance)</b>				
10	0.46	Position	S, W	
11	0.34	Position	S, W	
16	0.70	Position	T	
17	0.37	Position	T, W, P	
19	0.38	Position	T, W, P	
22	-0.35	Spin	M	1
<b>FACTOR 3 (6.3% of variance)</b>				
5	0.33	Position	M, T	26
9	1.02	Position	M	30
15	0.40	Position	T	
30	0.72	Spin	M	9
<b>FACTOR 4 (6.0% of variance)</b>				
2	0.31	Position	M, P	23
6	0.48	Position	M	27
7	0.59	Position	M	28
21	0.30	Position	M, T	
22	0.40	Spin	M	1
24	0.40	Spin	M	3
27	0.55	Spin	M	6
28	0.60	Spin	M	7

<sup>a</sup> A row in italics indicates that the item’s isomorph does not appear in the same factor grouping; there are five such rows.

Once we generated this factor structure, we returned to the QMCA items to interpret the item groupings. As shown in Table I, several items (2, 17, 19, 22) appear in multiple fac-

tors. In fact, item 22 falls into three of the four factors, though it appears with a negative factor loading in Factor 2 (F2). Given the ratio of position and spin contexts questions on the QMCA, there does not appear to be a distinct “position context” or “spin context” factor. In addition, these factors do not seem to reflect item difficulty, as the most and least difficult items are distributed throughout the four factors.

Regarding isomorphic questions, six of the eight isomorphic pairs we included in the EFA had both items load into the same factor at least once, which suggests that these isomorphs measure a common latent variable across position and spin contexts. Three of those pairs loaded into F1, which is the predominant grouping in terms of explained variance. Two pairs of isomorphs (items 3 and 24; items 5 and 26) had one item each (item 3 and 26, respectively) that did not load into any factor. Item 3 proved to be the most difficult item in the student responses we examined, with only 23% of students answering it correctly. (The average across all questions on the QMCA was 57%.) In addition, it is the only item that did not perform comparably with its isomorph in a prior study of these questions [9]. This may explain its absence from the factor structure. Items 5 and 26 are both questions that prompt students for rationales to a “Yes/No” question. While item 26 poses the same question as its isomorph, its choice selections are not isomorphic. This is because the choice selections of both items were written based on common student responses to prior, open-ended administrations of the QMCA.

Some isomorphic pairs were grouped in one factor but ungrouped in another. For example, the isomorphic pair of items 2 and 22 both load into F1, but only item 2 loads into F4. This pattern occurred with four of the six isomorphic pairs that loaded into the factor structure, and it complicates attempts to measure the same construct across position and spin contexts because it suggests there is an additional latent variable measured by some items that is not measured by its isomorph.

Regarding the original concept framework of the QMCA, there does not appear to be a neat separation of items along any of the five domains. However, a few patterns appear in the distribution of domain types between each factor. For example, F2 contains entirely S, T, and W items alongside one M item that is negatively correlated with this factor. Similarly, F3 contains solely M and T items. However, plenty of these domain types intermingle in F1 and F4.

A closer examination of each item grouping indicates that these factors might represent students’ attempt at mathematical sense-making or applying symbolic forms [17]. Most of the items in F1 ask students to interpret a probability amplitude or involve the probability density of a particle, both written as a superposition. As such, students must answer questions by recognizing and attending to the particular structure of the equation given for the state of the particle. All the items in F2 with positive factor loadings are in the position context and involve one-dimensional infinite square potential wells, though there are few commonalities beyond their physical setup. In addition, F2 has the distinction of containing a negative factor loading with item 22 (a spin question about

maximum value), which implies that question measures this latent variable in the opposite direction as the other items. The factor loadings of items in F3 suggest that isomorphic items 9 and 30 explain substantially more variance than items 5 and 15. Items 9 and 30 are both about the interpretation of relative phase in a superposition state. Finally, many of the items in F4 task students with interpreting or determining the state of a system, usually after a measurement.

These are not definitive interpretations of these factors, as there are some questions of each type that do not appear in these factors. For example, item 23 directly asks students for the state of a system after a spin measurement, but it is notably absent from F4. Rather, they collectively suggest that the constructs measured by the QMCA may be related to students' command of mathematical formalism or use of symbolic forms in problem solving. This interpretation agrees with prior work suggesting that students struggle with both mathematical formalism and categorization of problems in upper-division introductory QM courses [3].

## V. CONCLUSIONS

The factor structure generated from our EFA of the QMCA shows that there exists some measurable relationship between isomorphic questions in the concept inventory. However, the isomorphic questions are not entirely comparable, since four of the six pairs in the factor structure do not group together consistently. Further study should be pursued in the comparison of these isomorphs, particularly in the form of qualitative methods that might tease out the nuances in student thinking across contexts. One suggestion would be structured interviews focused on students' attempts to solve isomorphic

problems.

Qualitative analysis of these problems may also give insight into the latent variables measured by each of these factors, since items on the QMCA did not group in a predictable manner. The inconsistency between the QMCA's content domains and the constructs measured according to our factor structure reflects a common theme in the factor analysis of concept inventories [10, 11, 13]. These constructs are critical to determine, as they inform how students responses can be interpreted and used to inform instructional practice. In the same vein, the concept framework domains of the QMCA may be altered, slightly, to accommodate the addition of spin-context items.

Finally, these results indicate that some modifications could be made to the QMCA to maximize its utility to instructors across instructional contexts. Data from test administrations should continue to be collected, as a larger sample size will provide for more powerful statistical techniques. As all of the instructors in our sample taught with the spin-first instructional paradigm, researchers should also consider gathering student results from position-first courses. At the same time, modifications should be considered to items that failed to load into factors, particularly isomorphic questions that did not load with their respective pair. These alterations will ensure the QMCA provides useful feedback for instructors regardless of their curricular paradigm.

## ACKNOWLEDGMENTS

We would like to thank our departmental PER groups and colleagues for their support and feedback, particularly James Wells and Benjamin Schermerhorn. This work was funded by NSF DUE-1626280, DUE-1626594, and DUE-1626482.

- 
- [1] D.J. Griffiths and D.F. Schroeter, *Introduction to Quantum Mechanics* (Cambridge University Press, Cambridge, 2018).
  - [2] D. McIntyre, C.A. Manogue, and J. Tate, *Quantum Mechanics: A Paradigms Approach* (Pearson, Upper Saddle River, NJ, 2012).
  - [3] C. Singh, *Student understanding of quantum mechanics at the beginning of graduate instruction*, Am. J. Phys. **76**, 277 (2008).
  - [4] G. Passante, P.J. Emigh, and P.S. Shaffer, *Examining student ideas about energy measurements on quantum states across undergraduate and graduate levels*, Phys. Rev. ST - PER **11**, 020111 (2015).
  - [5] C. Singh and E. Marshman, *Review of student difficulties in upper division quantum mechanics*, Phys. Rev. ST - PER **11**, 020117 (2015).
  - [6] G. Passante, P.J. Emigh, and P.S. Shaffer, *Student understanding of time dependence in quantum mechanics*, Phys. Rev. ST - PER **11**, 020112 (2015).
  - [7] G. Passante, P.J. Emigh, and P.S. Shaffer, *Student ability to distinguish between superposition states and mixed states in quantum mechanics*, Phys. Rev. ST - PER **11**, 020135 (2015).
  - [8] H.R. Sadaghiani and S.J. Pollock, *Quantum mechanics concept assessment: Development and validation study*, Phys. Rev. ST - PER **11**, 010110 (2015).
  - [9] S.J. Pollock, et al., *Designing, validating, and contrasting quantum mechanics questions for spin states and spatial wave functions*, in *Physics Education Research Conference Proceedings, 2018*, edited by A.L. Traxler, Y. Cao, and S.F. Wolf (Washington, DC, 2018).
  - [10] D. Huffman and P. Heller, *What does the force concept inventory actually measure?*, Phys. Teach. **33**, 138 (1995).
  - [11] T.F. Scott, D. Schumayer, and A.R. Gray, *Exploratory factor analysis of a Force Concept Inventory data set*, Phys. Rev. ST - PER **8**, 020105 (2012).
  - [12] P.V. Engelhardt, *An introduction to classical test theory as applied to conceptual multiple-choice tests*, in *Getting Started in PER*, edited by C. Henderson and K.A. Harper (AAPT, College Park, MD), Review in PER Vol. 2.
  - [13] L. Ding and R. Beichner, *Approaches to data analysis of multiple-choice questions*, Phys. Rev. ST - PER **5**, 020103 (2009).

- [14] U. Lorenzo-Seva and P.J. Ferrando, *FACTOR 9.2: A comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models*, Applied Psychological Measurement **37**, 6 (2013).
- [15] U. Lorenzo-Seva and P.J. Ferrando, *Not positive definite correlation matrices in exploratory item factor analysis: Causes, consequences and a proposed solution*, Structural Equation Modeling: A Multidisciplinary Journal (2020).
- [16] H.F. Kaiser, *An index of factorial simplicity*, Psychometrika **39** (1974), p.35.
- [17] B.W. Dreyfus, et al., *Mathematical sense-making in quantum mechanics: An initial peek*, Phys. Rev. ST - PER **13**, 020141 (2017).