# Group learning for increasing sense of belonging in undergraduate quantum mechanics.

*E.V. Hansen* *Drexel University, Dept of Physics* *Term project for GRAD 513, Spring 2019*

## 1 IDENTIFICATION OF INSTRUCTIONAL TECHNIQUE

### 1.1 *Topic Identification and Course Details*

For this project I selected "Introduction to Quantum Mechanics" (i.e. Drexel 326), specifically focusing on units for Schrödinger's equation and wave functions. Learning objectives might include:

- (LO1) Understanding the classical foundations of quantum mechanics.
- (LO1) Identifying terms of the Schrödinger equation and connection between them
- (LO2) Deriving and adapting solutions to problems using the Schrödinger equation.
  - (LO2.1) Deriving and adapting solutions of one-dimensional problems.
  - (LO2.2) Deriving and adapting solutions of two-dimensional square wells (finite and infinite).
  - (LO2.3) Deriving and adapting solutions of the one-dimensional harmonic oscillator.

Such courses are taken by senior undergraduate physics majors, typically in their last year. Depending on the institution, class sizes could range from 10-100. The demographics of these courses are predominantly male and predominantly white or Asian; the most recent statistics from the American Institute of Physics show that only 21% of bachelor's degrees in physics were awarded to women, and only 1.7% were awarded to women belonging to underrepresented racial/ethnic minorities [1][1, 2]. In general, these demographics are a good approximation for the makeup of a quantum physics / mechanics, as students generally take this course in their junior or senior year. We know that women, *especially* women of color, suffer from stereotype threat and self-isolation in science, technology, engineering and mathematics [3, 4, 5, 6, 7, 8], which only increases with increased mathematical difficulty[9].

I chose this course because quantum mechanics (specifically this unit) requires advanced mathematics, up to and including two-variable second order linear differential equations. It also requires retention of material from previous physics and mathematics courses (e.g. mechanics, E&M, linear algebra, advanced calculus). Traditionally quantum mechanics has been thought of as a difficult course due to its abstract concepts - this unit in particular has nearly no tangible / hands-on experiment to accompany it, so has almost always been taught using lecture. An instructor, traditionally, will derive equations and solutions in class and show off potential applications without directly solving a homework/exam-like problem in front of the class. Most problem-solving skills and strategies are learned during homework or through reading a textbook.

### 1.2 *Selection of evidence-based instructional technique*

The instructional technique I would like to focus on is group work. This would specifically involve single-problem assignments to the whole class (within a lecture setting) where students would work on a set problem in small (2-4 people) groups. I believe that this unit of quantum mechanics lends itself well to group work in class - there are many potential initial conditions for problems using the Schrödinger equation, and many potential misconceptions that students can work through in groups during class time. For example, Singh & Marshman (2015)[10] report that students have difficulty correctly sketching the shape of a wavefunction (solution to the equation), even

[1] defined by AIP as African American, Hispanic, and Native American

if their derived mathematical function is correct. Instead of requiring students to discover this individually, an instructor could ask students to work together to derive energy levels and sketch a function, and then return to the class to compare and contrast to other groups. Students could then break back into small groups with a new set of initial conditions (maybe even different sets for different groups) and examine how the problem changes.

I selected this technique in particular, as an analysis of social-psychological factors by Lewis et.al. (physics specific) [7] suggested that direct intervention which increased positive social connections in the classroom was necessary to overcome the main contributing factors to lack of belonging in women in STEM: lack of peers and role models, and the existence of negative ability stereotypes. The authors suggest that these positive social connections, where women are not predominantly in a support role, would allow female students to observe other students struggling with the same material and to observe the variability in their colleagues traits and abilities. This hypothesis lends itself easily to the implementation of group work, the format of which is discussed further in Section 3.1.

## 2 LITERATURE REVIEW

In their meta analysis, Springer et.al. [11] found that small group work was positively and significantly correlated to achievement, persistence in introductory STEM[2] courses, and attitudes toward STEM. They specifically claim that students "learn best through active, collaborative, small-group work inside and outside the classroom." Springer et.al. describes three theoretical perspectives in which group work may improve learning: motivational, affective, and cognitive perspectives. The former describes group work as building an incentive structure where all students can succeed as a collaborative unit, as opposed to competing for top grades in a class where one students success decreases the chances that others will succeed. The affective philosophy of education emphasizes group work as a non-threatening environment where students have a higher chance of being heard by their colleagues (compared to an instructor serving as an "unquestioned authority"). Cognitive perspective theorizes that the opportunity for students to debate and discuss their own works is more conducive to learning, specifically when the assignments are open-ended problems with multiple paths and multiple acceptable solutions. This is supported by the National Research Council 2000 report titled "How People Learn" [12]; the report supports 'community centered' learning, defined by active participation of students who learn from one another. It is in the cognitive perspective where one might discuss differences between 'cooperative' (students work together on a task with assigned roles and procedures) and 'collaborative' (students are given a task with less structure and navigate on their own) learning.[3] However, this meta-analysis was limited to only undergraduate STEM courses and did not distinguish between the types or extent of group work.

The benefits of peer instruction and group work in introductory physics courses has been well established. In 2001, Crouch and Mazur [13] published a ten-year project at Harvard University which reported increased conceptual and quantitative skills upon implementation of peer instruction. Students were assigned selected problems from the homework and worked in small groups. Instructors circulated the room to help with arising difficulties. At the end of the week, students were graded on individual reporting of their homeworks. Crouch and Mazur report that while students were skeptical and resistant to the approach, learning assessments (specifically the Force Concept Inventory[14] and the Mechanics Baseline Tests[15]) demonstrated dramatic improvements in student learning after the introduction of peer instruction. This study was limited to Harvard University but covered both algebra- and calculus-based courses; a confounding factor may be the introduction of additional required readings during the same study period. Additionally, the study did not include a control group but rather presented observational data.

Additional studies at the University of Pittsburgh (Singh 2005[16]) report a significant gain in a conceptual in-

[2] Springer et.al. uses the term 'SMET' to describe what we now consider 'STEM' courses. I use 'STEM' to avoid confusion.
[3] This proposal will not distinguish between the two at present.

ventory of electromagnetism in a calculus-based introductory physics course for students who worked in pairs. The study was done experimentally, with two courses hosting group intervention and two as a control; students were allowed to choose their own partners and were familiar with group work from previous classes. Unlike Crouch and Mazur, instructors did not intervene or facilitate while peer discussions were taking place. The author reports that students in the peer instruction groups not only outperformed those in the control classes, but also demonstrated "co-construction" of knowledge in 29% of cases. From the cognitive perspective, co-construction of knowledge is defined in Singh's paper as the effect of two students producing the correct answer when neither student would have answered correctly independently as determined by pre- and post-tests. When the peer instruction premise was applied to quantum mechanics, Singh and Marshman report co-construction in 25% of cases ([17][4]).

Once again, these studies are more formally based in introductory undergraduate physics courses, but there is evidence to suggest that the same methods can be transferred to an upper division course. Multiple metacognition studies [10, 18, 19] dove into misunderstandings of undergraduate and graduate students in quantum mechanics and have shown at a fundamental level that upper-division students face the same basic difficulties as students in introductory courses. More specifically, students struggle with (1) scaffolding knowledge on an existing knowledge base and (2) categorizing similar problems into groups. For this reason, the studies cited above suggest that group problem solving tools which have been shown to help students build on prior knowledge in introductory courses may also be effective in upper division courses.

There is some observational evidence for this: to look more closely at the need for active learning and group work in quantum mechanics, we turn to Carr and McKagan (2009) [20] who looked at reforming quantum mechanics courses at the graduate student level. While this is above the (undergraduate) level of this proposal, the complex mathematics and course size/structure are closer between graduate and undergraduate quantum than they are between introductory and quantum courses at the undergraduate level. Carr and McKagan reported observational results - they had students break up into groups of two to four to work on problems, and write their results on whiteboards around the room. Teams were only assigned if the instructors observed particular students working together often. As further evidence for the affective perspective, instructors also sat down while students worked to decrease the level of professional threat in the room. They report that student response was positive. While this was not a laboratory study, and the authors do state that it was difficult to disentangle this additional structure from the difference in textbooks or additional resources, they do report - from student surveys - that students felt more in control of their courses and reported increased sense of belonging (which has been directly correlated with academic success in female physics students [7]).

There is limited evidence to suggest that small group work has been beneficial to students in upper division courses in a laboratory setting as well. Deslauriers and Weiman [21] compared two cohorts in a modern/quantum physics course - one with a 'very highly rated traditional lecturer' and another with small group activities. Both classes had optional recitations where students could work on problems in groups with a teaching assistant present, but the active-learning cohort experienced interactive lecture time including group worksheets and simulation in-class activities. Similarly to Crouch and Mazur, the authors found initial skepticism in students in the non-traditional lecture which dissipated as the course progressed. However, assessment with the Quantum Mechanics Conceptual Survey showed that students with the traditional lecture scored 19% lower. Interestingly, the authors also followed up with students 6 and 18 months after completion of the course (with no interim material) and demonstrated that there was only a few percent decrease (not significant) for both cohorts. The authors interpret this as suggesting that students will retain the same percentage of knowledge regardless of teaching method, but students exposed to group work start out with a higher threshold.

[4] The citation here describes the mechanisms by which students struggle with quantum mechanics - exact methods and results leading to the '25%' are technically unpublished.

There are limits to this research. With the exception of Deslauriers and Wieman, I was unable to find any laboratory studies in which group learning took place in upper division undergraduate courses, although there are plenty of tutorials and lesson plans (i.e. [22]) to suggest that there are researchers actively working on this problem. Additionally, as was discussed earlier in this proposal, any study that wants to quantitatively determine the effect of group work on female students in a research setting is limited by statistics due to the small number of women in physics courses; all of the studies I have listed above do not take gender identity into account, and do not describe in any detail the homo- or heterogeneity of their groups.

## 3 ASSESSMENT PLAN

### 3.1 *Course description*

This plan would be implemented in a course at a large, primarily undergraduate institution. A quantum mechanics course at this location would have 60-80 students. Lectures will be supplemented with group work; after a period of equation /concept review, the instructor will assign a homework-like problem to the students to work on in groups; ideally this problem would include some mathematics or advanced concepts. Students will self-select into groups of 4-5, resulting in 12-20 groups. These groups can change class to class. Additionally, self-selection is important to this study - there is evidence to suggest that assigning groups reduce sense of community in minoritized groups (e.g. LGBT and female students [23, 24]).

The instructor and TA's will float between sets of groups to answer questions but will encourage groups to work problems out together before requesting instructor's help. If this is the first course students have done group work in, instructors should be more encouraging of inter-student collaboration. A future study could evaluate the effects of professional threat by limiting instructor interaction, like [20]. Near the end of the session, students will decide on a reporter from their group, and come together to report answers to the class as a whole. If there are too many groups or limited time, groups would randomly selected to report - a variant on this could be to provide groups with slightly different 'tweaks' on the same problem. A single group from each 'tweak' would be selected to explain why their problem is different and how the solution changes. Students would then turn in a summary of their work to the instructor at the end of class, with all student names attached; this summary would be graded for participation.

Ideally, students engaging in this sort of group work would report a high sense of community as well as conceptual understanding. In order to gauge how well this worked, there will be three assessment tools: a pre-/post-course concept inventory, a mid-semester reflection, and a post-course reflection.

### 3.2 *Pre and post concept inventory:*

The Quantum Mechanics Conceptual Assessment (QMCA), a concept inventory proposed and validated in [25], would be given in the first class; it would be graded on participation and would be worth 2% of the course grade. The QMCA has specifically been validated for instructional and research purposes, and has been shown to be effective at pinpointing common difficulties among student populations. This assessment would primarily be used for comparison to a post-course inventory, but could also be used to modify existing curriculum based on prior knowledge of the cohort. Of special interest to this project would be questions coded for wave function, probability density, and Schrödinger equation solutions as outlined in the learning objectives for this project (these questions make up approximately half of the QMCA).

A post-course concept inventory would be given in class at the end of the term. This test would be graded for correctness and worth 2% of the course grade. Individual scores will be compared to pre-course QMCA for improvement. The cohort average and cohort improvement will be compared to literature to ensure that the changes in the course did not negatively affect student learning. These scores will also be analyzed in parallel with the reflections, as discussed below.

### 3.3 *Mid-semester reflection*

At the midpoint of the semester, the instructor will ask the class to complete a reflection / evaluation which will specifically target 'sense of belonging' within the space of the course. It would be administered over Qualtrics (typing preferable to handwriting), with completion of the survey resulting in 2% of the course grade. Students would be required to respond to each question, although "Do not wish to respond" would be an option for each.

The survey would be composed primarily of numerical scale questions, as modeled after [9] which surveyed women and people of color for sense of belonging in mathematics. As established previously, quantum mechanics courses are mathematically intensive - it is not an unreasonable assumption that this survey, although designed for mathematics, would be useful in this case. The survey will also include the opportunity for long-form responses for students to reflect on their sense of community or share particularly illustrative events from class. This method of evaluation clearly has the ability to trigger an emotional response, so ideally students should take this kind of survey in a safe space for them - this means that students should not take the survey in class (perhaps the weekly homework load could be lightened in order to make space for this). The instructor should take some time in class to indicate that they want to evaluate how welcome the class environment is, and feedback from students is important for that goal. Additionally, the survey would *not* be anonymous but it should be heavily stressed that selecting "Do not wish to respond" for any/all questions would not impact the course grade in any way.

Questions may include the following (questions marked with ⋆ are reverse coded):

*With respect to this quantum physics course I...*

1. *...feel different from others in this course [26]* ⋆
2. *... feel alone or isolated [26]* ⋆
3. *... feel like I belong in the classroom community [9]*
4. *... feel like an outsider [9]* ⋆
5. *... enjoy being an active participant [9]*
6. *... trust my instructors to be committed to helping me learn [9]*
7. *... feel like I fit in [9]*
8. *... feel at ease / comfortable [9]*
9. *... feel anxious / nervous [9]* ⋆

Ideally the survey would also ask students to evaluate or describe their impression of the course as a whole, with special attention to ways in which the students believe the instructors have succeeded or failed in producing an equitable and inclusive learning environment. There should also be a long form question soliciting possible solutions to any problems the respondent raised.

Results from the numerical scale survey will be aggregated per individual and compared to demographic information, looking for subcohort trends. Additionally, each question will be compared across cohort (average and distribution) looking for any question that was particularly polarizing. Long-form responses will be coded for themes such as belonging, identity, anxiety, and isolation. These themes will be compared to numerical scale results to ensure responses are consistent on an individual level. The frequency of these coded themes should also be compared to demographic information, looking for sub-cohort trends.

At this point, the instructor should use these as a mechanism for classroom change, should any particularly alarming trends arise. The instructor has the opportunity to make corrections to the course structure or community:

- Skepticism about the format of group work can be largely ignored at this point [13, 20] without further evidence from the post-course conceptual inventory.
- Compare responses to student demographics or self-identification to ensure that there are no particular groups that are reporting isolation or anxiety. If there are, the instructor should reach out to on-campus resources for inclusive classroom best practices (for Drexel University, this would be the Office of Equality and Diversity).
- Any responses that indicate individual students are not feeling included should be followed up on depending on the particular circumstances. For example, if a student is feeling alone or isolated and the instructor (in reviewing the student's in-class assignments) notices that they are constantly grouped with the same people, it might be worth attempting to split that group and splice them with others, or to ask that all students attempt to create new groups in the second half of the semester.

Ultimately this survey should not dissuade an instructor from continuing to use group work as an instructional technique, or deviating too far from the original course structure (at least based on the mid-course survey alone), but it should encourage them to ensure that their behavior (and the behavior of students) in the classroom is equitable and inclusive. If there are any particular suggested solutions that are reasonable to implement, instructors should consider them.

### 3.4 *Post-course reflection:*

At the end of each course, students would be asked to complete a similar survey to the mid-semester reflection, with additional questions related to their opinions of the group work framework. This might include asking students if they tended to work with the same people the entire term, or if they felt that that group work was a reasonable use of their in-class time. Any changes made from the mid-course survey should also be evaluated for effectiveness, and the survey should solicit new ideas / reflections from students. The survey would be collected via Qualtrics and completion of the survey (answering each question regardless of response) would be worth 2% of the course grade.

This reflection would be analyzed and coded identically to the mid-course reflection but would have additional analysis steps. Individual responses should be compared (pre- vs post-) for any changes in sense of belonging, especially through the lens of individual demographics. The instructor should also compare the frequency of themes in long form responses, specifically focused on any changes made based on responses to the mid-course survey. Finally, individual scores in the numerical scale questions would be compared (pre- vs post-) for trends between the sense of belonging scale and the concept inventory, with the expectation that students who reported a high (or increased) sense of belonging would show additional understanding.

### 3.5 *Analysis and Further Steps*

Specifically the surveys aim to compare male and female responses with the addition of small group work; I expect all students who participated in small group work to respond with an increased sense of community, but that female responses will be disproportionately higher (compared to literature like [9]). With the addition of the scores from the post-course conceptual inventory, the instructor could compare a student's reported sense of belonging to their overall conceptual understanding of the course material, and examine these comparisons for trends across the class as a whole as well as by demographic information.

The instructor should specifically look at this data through a lens of modifying the course structure or in-class behavior; this might include re-evaluating group structures (self-selected vs assigned, heterogeneity), how long/difficult in-class problems are, and how much they aided in conceptual understanding specifically. It also might include ways in which an instructor could change the language they used in the course, or if special accommodations should be made for sets of students for whom group work was, in fact, detrimental. Instructors should pay special attention to the changes implemented from the mid-course survey and evaluate if they were effective (via increased 'sense of belonging' scores) and if they should be implemented/modified in the future.

A confounding effect of these surveys could be the group formation in class. While this work disfavors assigning students to groups in case it disproportionately increases threat and risk to minoritized groups, it would be important to report overall results through a lens of the heterogeneity of self-selected groups. Future work could address this through a couple methods: students could self-report through the survey the type of group they tended to work with (all male, mostly male, roughly equal, mostly female, all female) or a more nuanced analysis could analyze the ties between students and correlation with their responses both to the conceptual inventory and the survey (group data is available through students in-class work, with names on each submission).

Additionally, the instructor should be wary of assigning any particular belief to a specific demographic in their classroom (i.e. 'Sarah will feel less included because she is a black woman'). First, this sort of analysis unfairly associates students with a limited axis of their personal identity and ignores the individuality or additional intersecting axes of any one student. Second, as discussed previously, the demographics of a quantum physics classroom are extremely limited - in this sort of classroom only $\sim$20% of students are women, only $\sim$12% of students belong to underrepresented racial/ethnic minorities, and only $\sim$2% belong to both groups [1, 2]. These surveys therefore should be used to carefully evaluate teaching methods for any overarching trends, supported by individual student experiences, but any conclusions drawn from these surveys should be limited to the course, instructor behavior, and instructional method specifically.

## 4 OPEN QUESTIONS

There are many routes of inquiry remaining to measure the effect of group learning in undergraduate quantum mechanics. A common question in group learning relates to the difficulty of the problem set given to the class in comparison to the methods used to assess learning - in other words, there may be confounding factors in using concept tests to study group learning where the problem sets are homework-like (“back of the book") problems. Additionally, it is unclear to what extent instructor intervention during problem solving time would have an impact on learning - studies like [20] were unclear as to the impact of professional threat on their results.

However, with the specific focus on ‘sense of belonging’, another problem comes to mind. A lecture based classroom is completely dominated by a two-point interaction: the instructor and the cohort. With the introduction of group learning, the instructor opens the classroom to additional complexity as interactions between students become dominant. There are several open questions in the field of cognitive psychology with relation to group dynamics in the classroom (stemming primarily out of [27]); of particular interest to this work is the specific demographic makeup of groups. Studies of group dynamics in problem solving are in disagreement: it appears that diverse groups solve problems better in a professional setting [28], but in the large undergraduate classroom there is evidence that students may benefit when groups are assigned based on similar problem solving styles [29] or that neither heterogeneity nor homogeneity is particularly beneficial [30]. Due to lack of understanding as to which type of groups (if any) are more effective, many studies on the effectiveness of group learning either assign groups or allow students to self-select based on instructor preference or classroom efficiency [13, 16, 20].

As discussed previously, this work suggests that students self-assign into groups instead of being assigned by an instructor; this is primarily due to studies reporting that minoritized groups are put at risk in the classroom

when assigned to groups (e.g. LGBT and female students [23, 24]). However, in protecting minoritized students, it is possible that any or all students will not assign themselves in groups most conducive to their own learning, thereby negating benefits from the introduction of group learning [31]. One question naturally arises:
*Knowing that 1) assigning groups presents cognitive threat to minoritized students, but 2) allowing students to self-select groups could create groupthink that negates the benefits of group learning, which method is more worthwhile?*

Research into this question would allow studies to disentangle the effect of group dynamics from questions plaguing group learning as a topic of inquiry - for example, knowing that students are in the most beneficial group structure, one could look more closely at the effect of problem difficulty or similarity to assessment tools. Having a confident solution to this question would also allow for studies into the impact of other classroom features on sense of belonging; for example, one could examine the use of culturally-diverse instructors or textbooks, or multiple methods of engagement/assessment within the group learning structure (suggestions from *Guide for Inclusive Teaching at Columbia* [32])

## REFERENCES


[1] Laura Merner and John Tyler. African American, Hispanic, and Native American Women among Bachelors in Physical Sciences and Engineering. Technical report, American Institute of Physics, 2017.

[2] Anne Marie Porter and Rachel Ivie. Women in Physics and Astronomy, 2019. Technical report, American Institute of Physics, 2019.

[3] Steven J Spencer, Claude M Steele, and Diane M Quinn. Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, 35:4–28, 1999.

[4] Li Shen, Hany Farid, and Mark A. Mcpeek. *Talking about leaving: Why undergraduate leave the sciences Boulder*. 2008.

[5] Sarah L Eddy and Sara E Brownell. Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. *Physical Review Physics Education Research*, 12, 2016.

[6] Marie-NoÃńlle Delisle, FrÃľdÃľric Guay, Caroline Senécal, and Simon Larose. Predicting stereotype endorsement and academic motivation in women in science programs: A longitudinal model âŸĘ. *Learning and Individual Differences*, 19:468–475, 2009.

[7] Karyn L Lewis, Jane G Stout, Steven J Pollock, Noah D Finkelstein, and Tiffany A Ito. Fitting in or opting out: A review of key social-psychological factors influencing a sense of belonging for women in physics. *Physical Review Physics Education Research*, 12(020110), 2016.

[8] Karyn L Lewis, Jane G Stout, Noah D Finkelstein, Steven J Pollock, Akira Miyake, Geoff L Cohen, and Tiffany A Ito. Fitting in to Move Forward: Belonging, Gender, and Persistence in the Physical Sciences, Technology, Engineering, and Mathematics (pSTEM). *Psychology of Women Quarterly*, 41(4):420–436, 2017.

[9] Catherine Good, Aneeta Rattan, and Carol S. Dweck. Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, 102(4):700–717, 2012.

[10] Chandralekha Singh and Emily Marshman. Review of student difficulties in upper-level quantum mechanics. *Physical Review Special Topics - Physics Education Research*, 11, 2015.

[11] Leonard Springer, Mary Elizabeth Stanne, and Samuel S. Donovan. Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis. *Review of Educational Research*, 69(1):21–51, 1999.

[12] Commission on Behavioral Committee on Developments in the Science of Learning, Social Sciences, and National Research Council Education. *How People Learn: Brain, Mind, Experience, and School*. National Academy Press, Washington D.C., 2000.

[13] Catherine H Crouch and Eric Mazur. Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 2001.

[14] D Hestenes, M Wells, and G Swackhamer. Force Concept Inventory. *The Physics Teacher*, 30:141–166, 1992.

[15] D Hestenes and M Wells. A Mechanics Baseline Test. *The Physics Teacher*, 30:159–166, 1992.

[16] Chandralekha Singh. Impact of peer interaction on conceptual test performance. *American Journal of Physics*, 73(5):446–451, 5 2005.

[17] Emily Marshman and Chandralekha Singh. Framework for understanding the patterns of student difficulties in quantum mechanics. *Physical Review Special Topics - Physics Education Research*, 11, 2015.

[18] Guangtian Zhu and Chandralekha Singh. Improving students' understanding of quantum measurement. I. Investigation of difficulties. *Physical Review Special Topics - Physics Education Research*, 8, 2012.

[19] Guangtian Zhu and Chandralekha Singh. Surveying students' understanding of quantum mechanics in one spatial dimension. *Citation: American Journal of Physics*, 80:277, 2012.

[20] L D Carr and S B Mckagan. Graduate quantum mechanics reform. *American Journal of Physics*, 77:308, 2009.

[21] Louis Deslauriers and Carl Wieman. Learning and retention of quantum concepts with different teaching methods. *Physical Review Special Topics - Physics Education Research*, 7, 2011.

[22] Chandralekha Singh. Interactive learning tutorials on quantum mechanics. *American Journal of Physics*, 76:400, 2008.

[23] Katelyn M. Cooper and Sara E. Brownell. Coming Out in Class: Challenges and Benefits of Active Learning in a Biology Classroom for LGBTQIA Students. *CBE Life Sciences Education*, 15(37):1–19, 2016.

[24] N Ding and E Harskamp. How Partner Gender Influences Female Students' Problem Solving in Physics Education. *Journal of Science Education and Technology*, 15(6), 2006.

[25] S.B. McKagan, K.K. Perkins, and C.E. Wieman. Design and validation of the Quantum Mechanics Conceptual Survey. *Physical Review Special Topics - Physics Education Research*, 6, 2010.

[26] Lauren M Aycock, Zahra Hazari, Eric Brewe, Kathryn B H Clancy, Theodore Hodapp, and Renee Michelle Goertzen. Sexual harassment reported by undergraduate female physicists. *Physical Review Physics Education Research*, 15, 2019.

[27] D. W. Johnson and R. T. Johnson. Cooperative learning and achievement. In *Cooperative learning: Theory and research*, pages 173–202. New York, 1990.

[28] Lu Hong and Scott E Page. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Technical report, 2004.

[29] Alexa Lamm, Catherine Shoulders, Grady Roberts, Tracy Irani, Lori Snyder, and Joel Brendemuhl. The Influence of Cognitive Diversity on Group Problem Solving Strategy. *Journal of Agricultural Education*, 53(1):18–30, 2012.

[30] Jason J B Harlow, David M Harrison, and Andrew Meyertholen. Effective student teams for collaborative learning in an introductory university physics course. *Physical Review Physics Education Research*, 12, 2016.

[31] Linda C Hodges. Contemporary Issues in Group Learning in Undergraduate Science Classrooms: A Perspective from Student Engagement. *CBE - Life Sciences Education*, 17(3):1–10, 2018.

[32] Lucy Appert, Christine Simonian Bean, Amanda Irvin, Amanda M. Jungels, Suzanna Klaf, and Mark Phillipson. Guide to Inclusive Teaching at Columbia. Technical report, Columbia University, 2018.