

Pseudolongitudinal investigation on Chinese students' categorization of kinematics and mechanics problems

Guangtian Zhu¹ and Jue Wang²

¹College of Teacher Education, East China Normal University, Shanghai 200062, China

²Department of Physics, East China Normal University, Shanghai 200062, China

(Received 8 May 2017; published 21 September 2017)

Students' categorization of physics problems reflects their expertise in problem solving. We conducted a pseudolongitudinal study to investigate the development of students' categorization ability. Over 250 Chinese students from grade 10 to grade 12 were asked to categorize 20 problems of kinematics and mechanics into suitable categories based on the similarity of solutions. We compared the categories made by the students in different grades and found that, although students in all three grades performed as novices, their expertise in categorization gradually developed. The results also suggested that the training in problem solving may affect students' categorization, especially for those in grade 12.

DOI: 10.1103/PhysRevPhysEducRes.13.020118

I. INTRODUCTION

The study of the differences between experts and novices has attracted the attention of both educators and researchers. Cognitive scientists and psychologists have investigated how knowledge is stored and organized in different patterns by experts and novices [1–5]. In a classic research conducted in 1981, Chi, Feltovich, and Glaser found that experts (graduate students in physics) categorized physics problems based on underlying physical principles, while novices (introductory physics students) were often influenced by certain surface features [3]. For example, about half of the novices sorted the problems in Chi's questionnaire as “inclined planes” or “pulleys,” but none of the experts used such categories. The conceptual categories reflected people's knowledge structure of physics. As described by Chi, the experts employed a schema, an appropriate principle-oriented knowledge structure, when solving physics problems. The experts first conducted a qualitative analysis of the question to determine the proper schema. Once the schema was confirmed, the knowledge contained in this schema provided the proper quantitative equations for the solution [3].

The fact that the novices have difficulty fitting problems into appropriate conceptual categories has been confirmed for other subjects beyond mechanics [6,7]. Savelsbergh, de Jong, and Ferguson-Hessler found that the physics students often failed to identify the proper solutions to electromagnetism questions. However, in the domain of electricity and magnetism, students' categorization was strongly affected by the “domain and task formats” rather than the “surface

features,” such as shapes or objects, in mechanics [6]. The categorization of the advanced physics contents was more complex even for experts. The physics professors made more diverse categories of problems for quantum mechanics than for introductory physics. Additionally, faculty cannot reach a common agreement on the best categories for the upper-level quantum mechanics questions [8].

Mason and Singh conducted a large-scale investigation to assess the distribution of categorization expertise among 400 introductory physics students [9]. Half of the participants were in an algebra-based class, while the other half were in a calculus-based class. The “degree of novice” for the algebra-based physics students was not as high as that reported in Chi's study, and the calculus-based physics students had a large overlap in categorization with the graduate students in the physics department. The results revealed a wider distribution of expertise among students in different levels of physics courses. Moreover, not all graduate students should be taken for granted as “experts.” In a study conducted by Singh, graduate students were asked to categorize introductory physics questions twice, i.e., from their own perspective and from the introductory physics students' perspective. Many graduate students were not aware of the significance of the categorization tasks even if they were the teaching assistants for the introductory physics courses [10].

Faculty's opinion are always treated as criteria in expertise, and most of time they appropriately categorize introductory physics questions according to the underlying principles that lead to the solutions. However, faculty may have “compiled knowledge” and sometimes assign false categories to certain types of problems [11]. Reif and Allen invited physics professors to answer questions related to the acceleration of a pendulum swinging to its lowest position. Many professors mistakenly claimed that the acceleration was zero without noticing the centripetal acceleration [11]. Further interviews suggested that the professors sorted the

Published by the American Physical Society under the terms of the *Creative Commons Attribution 4.0 International* license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

questions of “pendulum” and “springs” into the category of “simple harmonic motion.” For a standard model of one-dimensional simple harmonic oscillator, e.g., springs, the acceleration in the balance position was indeed zero. Hence, faculty may neglect the unique properties of a pendulum if they were simply asked to categorize the questions instead of explicitly solving the problems. Singh found that when dealing with nonintuitive introductory level questions that were familiar, physics faculty encountered difficulties under the time constraints [12]. However, the hierarchical knowledge structure still enabled the professors to apply a systematic approach in solving these nonintuitive problems.

The expertise of problem solving can be probed through categorizing tasks [13–16]. Veldhuis applied a cluster analysis to study the categorization of physics problems, and the statistical results indicated that the advanced novices used both surface and deep structure [13,14]. Hardiman, Dufresne, and Mestre assigned similarity judgment tasks to students and found that novices with the same level of experience may have divergent categorization behavior [15]. The better problem solvers among the novices reasoned consistently on principle, while others relied primarily on surface features [15]. This conclusion was also supported by another study by Chi on how students made explanatory statements when they were trying to interpret physics problems [16].

Noticing the relation between how students categorize problems and how they solve problems, educators began to apply categorizing tasks to improve students’ problem-solving abilities [7,17–20]. Bunce, Gabel, and Samuel provided an experimental group of students three 40-min training sessions on how to categorize chemistry problems. The control group used the same time to review the explicit method of problem solving. The experimental group students performed significantly higher than the control group students on the achievement test [7]. Mestre and collaborators trained the treatment group students with five 1-h sessions using a hierarchical analysis tool that emphasized the structure of physics principles and procedures. The treatment group outweighed the control group in both categorization tasks and problem-solving performance [19]. Similar effects were also observed in other disciplines such as mathematics [20].

The distribution of expertise is wide for novices. In previous research on categorization that has evaluated expertise based on the two levels of expertise (i.e., novices and experts), researchers have usually chosen first-year college students in an introductory physics course as the novice group, while the faculty or graduate students acted as the experts. Some research further distinguished graduate students with an intermediate level of expertise [10]. However, there is usually a huge gap between introductory physics students and physics graduate students. Most physics graduate students complete not only calculus-based

introductory physics but also upper-level physics courses, such as analytical mechanics or electrodynamics in their junior and senior years. However, many undergraduate students in introductory physics classes may never touch a physics textbook again after their freshman year. Hence, the comparison between undergraduate and graduate students is not sufficient to represent the progress of novices in learning physics. There is a lack of longitudinal research that probes the development of novices’ expertise in introductory physics as they learn more physics content.

In this Letter, we discuss a pseudolongitudinal study on novices’ categorization of introductory physics problems. Previous studies have already found that graduate students are better at categorization than students in introductory physics. However, it is unclear whether their advantage results from solely the advanced physics courses (e.g., analytical mechanics) or from gradual progress as students learn more introductory physics. This study aims to investigate whether the novices’ ability in categorizing physics problems develops gradually with their content knowledge growth in introductory physics. By collecting and analyzing the categories of different physics problems created by the students, we cannot only rank the expertise level of the students who had learned physics for different years but also identify the factors that may influence the choices of categories for different levels of novices. Since students’ problem-solving strategy is cued by their categorization of physics problems [15], the results of this study can help physics instructors better understand what is missing from their students’ knowledge structure. Hence, instructors can accordingly employ suitable teaching methods for different students to improve their conceptual understanding as well as problem-solving skills in introductory physics.

In contrast to previous studies of categorization that surveyed undergraduate students enrolled in an introductory level physics curriculum, we administered categorization tasks to high school students from grade 10 to grade 12 in China. The sample selection was based on the following considerations:

- (i) Undergraduate students in China have outstanding performance in introductory physics. A large-scale comparison of students’ understanding of introductory physics in China and in the United States revealed that Chinese students’ performance in FCI peaked near the 90% score [21]. Having survived the rigorous training in high school and the highly competitive national college entrance exam, most freshmen in Chinese universities have a solid understanding of the conceptual contents of introductory physics. Therefore, compared to Chinese undergraduate students, Chinese high school students from grade 10 to grade 12 are a better sample to represent the continuous growth of knowledge in algebra-based introductory physics. Their categorization of physics problems can reflect some common features of how students learn

algebra-based introductory physics in both high schools and universities.

- (ii) High school students in China have a consistent background of learning physics. In China, the K–12 education system consists of primary schools, middle schools, and high schools. The year span of these three levels of K–12 schools is either 5-4-3 or 6-3-3 years in different provinces. In either case, students start to learn physics in middle school in grade 8, and they must take a physics exam (along with other courses, such as Chinese, English, and Mathematics) at the end of grade 9 for high school admission. Physics is also a compulsory course in high school. All students take algebra-based physics courses in grade 10 and grade 11. If students select physics as one of their college entrance exam subjects, they continue to learn and review physics contents in grade 12. Since the Chinese students have similar prerequisite knowledge when learning algebra-based introductory physics in high school, the influence on students' prior knowledge before learning physics can be reduced.
- (iii) High school students in most provinces of China use the same curriculum standard of physics. Specifically, the grade 10 physics course starts with kinematics and continues to mechanics. The grade 11 physics course focuses on electricity and magnetism. In grade 12, students may study the optional topics of “thermodynamics,” “wave and optics,” or “modern physics,” depending on the requirement of their province. In most Chinese high schools, 12th graders study the new content for less than one semester and use the remaining time to review all the physics content introduced in grades 10 to 12. Since all students follow the same syllabus of introductory physics, we can safely combine the responses to the categorization tasks collected in different schools.

II. METHODOLOGY

The distributed questionnaire contained 20 questions about kinematics and mechanics, since all students from grade 10 to grade 12 had learned these topics. Seven questions were available from Chi's paper [3], and we selected six of them. One question about angular momentum conservation was ignored because the content of angular momentum was not covered in the physics curriculum of Chinese high schools. We also selected 6 questions from Singh's problem set [9]. According to Singh's categorization, the six questions from Chi's paper involved the primary categories of “rotational kinematics,” “mechanical energy conservation,” “Newton's second law” and “work-energy theorem.” The six questions we chose from Singh's work were mostly about “impulse momentum,” “2D kinematics,” and “momentum conservation plus energy conservation.” The other 8 questions in our questionnaire were selected from

the past papers of the Chinese college entrance exams of physics (we call it “Gao Kao”) [22]. To balance the number of questions in each primary category, the Gao Kao questions mostly covered the contents of rotational kinematics, 2D kinematics, and momentum conservation plus energy conservation. The full questionnaire can be found in the Supplemental Material [23].

We administered the questionnaire to 277 students (96 grade 10 students, 92 grade 11 students, and 89 grade 12 students) in two high schools. The students had three to four 45-min algebra-based introductory physics classes in school every week. The tests were given to the students in the second semester of a school year when the 10th graders had learned all the related content involved in the questionnaire. The students needed to answer the questionnaire on their own in a 45-min recitation class. They were given enough time to finish the questionnaire, and most students submitted their answers in about 20 min.

The front page of the questionnaire contained instructions that informed the students to categorize the questions based on their similarity. The instruction was adapted from Mason and Singh's study [9] and translated into Chinese. No further hints were provided to the students beyond the instructions printed on the questionnaire. On the answer sheet, students needed to write down the name of their categories, the question numbers binned in each category, and the reason why they grouped these questions into one category.

After we organized the responses to the questionnaire, we found that certain types of category names commonly existed in students' answers. We also found some novel category names that had not been reported in the previous studies. Sixteen students were interviewed to help us better interpret the origin of these categories. In the interviews, we first provided the students with their original answer sheet and asked them to recall how they had created each category and why they had binned each question into the corresponding category. Then, we asked the students to elaborate on the categories we were interested in. Students needed to solve one or two questions in these categories in order for us to probe their understanding of the related physical principles. For example, when we found that some students used the category name “force composition” instead of “Newton's second law” or “Newton's third law,” we asked them to solve the corresponding questions and observed whether they mentioned Newton's laws. The problem-solving process was monitored in the think-aloud protocol [24]. Students needed to speak out their reasoning process while solving the problem. We also interviewed the high school teachers about their curriculum arrangement related to kinematics and mechanics.

We evaluated students' categorization for each question as good, moderate, or poor. The good categories and one or two common moderate or poor categories for each question are listed in Table I. To be consistent with the standards of

TABLE I. Examples of good, moderate, and poor categories for each question. The family name in parentheses represents the source of the question. The item N/A represents that few of the students' answers were evaluated as moderate for the corresponding questions.

Question	Good categories	Moderate	Poor
1 (Chi)	Rotational kinematics	Friction	Angular velocity
2 (Chi)	Mechanical energy conservation, work-energy theorem, definition of work, or Newton's second law and kinematics	Energy or Hooke's law	Spring
3 (Singh)	Impulse-Momentum theorem	Momentum	Force
4 (Singh)	Impulse-Momentum theorem	Momentum	Force or collision
5 (Chi)	Mechanical energy conservation, work-energy theorem, definition of work, or Newton's second law and kinematics	Energy	Velocity or pulley
6 (Chi)	Newton's second law	N/A	Tension
7 (Zhu)	Newton's second law and rotational kinematics	Friction	Angular velocity
8 (Zhu)	Mechanical energy conservation, work-energy theorem, or 2D kinematics	Energy	Velocity or speed
9 (Chi)	Work-energy theorem, definition of work, or Newton's second law and kinematics	Energy	Friction or ramp
10 (Singh)	2D kinematics	N/A	Force or cliff
11 (Singh)	Mechanical energy conservation and momentum conservation	Energy or momentum	Collision
12 (Zhu)	Mechanical energy conservation and impulse-momentum theorem	Energy or momentum	Force or tension
13 (Zhu)	2D kinematics	Velocity	Ramp
14 (Singh)	Mechanical energy conservation and momentum conservation	Energy or momentum	Velocity
15 (Zhu)	Newton's second law and rotational kinematics	Friction	Rotating disk
16 (Zhu)	2D kinematics	N/A	Velocity
17 (Zhu)	Mechanical energy conservation and Momentum conservation	Energy or momentum	Velocity
18 (Chi)	Work-energy theorem, definition of work, or Newton's second law and kinematics	Energy or Friction	Velocity or ramp
19 (Zhu)	2D kinematics	N/A	Velocity or circular motion
20 (Singh)	Mechanical energy conservation	Energy	Velocity

Chi and Singh, we defined a category as “good” if it was named by the physical principles suitable for solving the corresponding questions. The categories named by the surface features of each problem were treated as “poor” categories. The common poor categories included “pulley,” “free fall,” “force,” and “speed”. Falsely assigning an irrelevant principle to a question was also considered poor. Hence, some categories could be considered good for certain questions but poor for other questions. For example, the category of Newton’s second law and rotational kinematics was good for question 7 since this question could be solved by analyzing the centripetal force. However, in question 19, although the object was moving along with the rotating table before it flew off, the information of circular motion did not contribute to solving this question. Hence, the category of rotational kinematics was binned as poor for Q19. The context of Q7 and Q19 is shown in Fig 1.

There are two types of “moderate” answers. First, if the name of the category was a general concept that was related to the question but was too broad, and the students did not

specify their rationale for placing the problem into that category, we considered the category moderate. For example, the category of “energy” was evaluated as moderate if the student did not mention “work-energy theorem” or “energy conservation” in their explanation. Second, if solving a particular question required two physical principles simultaneously, the students’ answer would be treated as moderate when they binned the question in a category of only one principle. For example, to be considered good for Q17, the students needed to either put Q17 in both categories of “momentum conservation” and energy conservation or create a separate category, “momentum conservation plus energy conservation.” If a student only binned Q17 as momentum conservation, then his or her answer was treated as moderate.

Some questions could be solved in different ways using different principles. For example, Q18 could be solved by applying either the energy conservation principle or Newton’s second law. Then, the students only needed to place the question in one suitable category (either energy conservation or Newton’s second law) to receive a good

7. An object with mass m is rotating with the disk. Suppose the distance between the object and the center of the disk is r . The coefficient of kinetic friction is μ , and the maximum static friction is equal to the kinetic friction. Find the maximum angular velocity of the disk to keep the object rotating with the disk without sliding.

19. A rotatory table with radius $R = 0.5\text{m}$ and height $H = 0.8\text{m}$ was rotating. An object on the edge of the table moved with the table without sliding. When the angular velocity increased, the object was thrown out and finally hit the ground. The horizontal distance of the object is $s = 0.4\text{m}$ away from the launching point and the maximum static friction was equal to the kinetic friction. Suppose the gravity was $g = 10\text{m/s}^2$. Find the initial velocity of the object when it was thrown out.

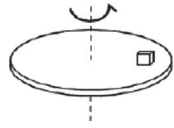
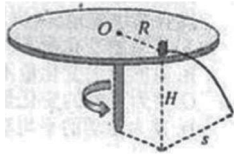



FIG. 1. Q7 and Q19 in the questionnaire. The category of rotational kinematics was considered good for Q7. However, the same category was evaluated as poor for Q19 because Q19 was actually a question about projectile motion, and the information about circular motion made no contribution to solving this problem.

evaluation. As shown in Table I, the questions requiring two simultaneous principles were described in the “A and B” format, and the questions suitable for different categories were described in the A or B format.

III. RESULTS

The questionnaire contained 20 questions, and the category for each question was evaluated as good, moderate, or poor. Two graduate students studying physics education independently evaluated all students’ answers based on the standards discussed above. The agreement between their ratings was more than 90%. For the situations that the two graduate students had a disagreement between “good or moderate” or “moderate or poor,” we presented these answers to the professors in the physics department to obtain the final rating.

The participating students answered all 20 questions on the questionnaire. Figure 2 shows the number of questions binned in the good categories created by the participants. The horizontal axis of Fig. 2 is the number of good answers, and the vertical axis is the percentage of students in different grades. For example, the leftmost blue bar indicates that 63% of the 10th graders (60 out of 96

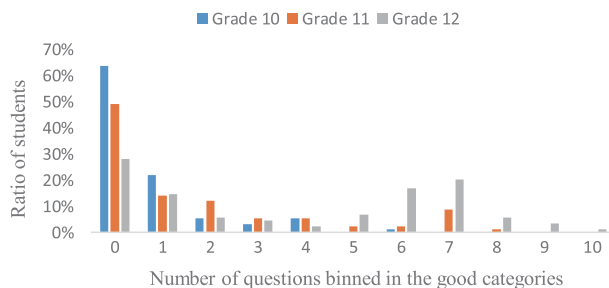


FIG. 2. Percentage of good categories created by students in different grades. The horizontal axis is the number of questions in the good categories, and the vertical axis is the percentage of students in different grades (grade 10-blue, grade 11-orange, grade 12-gray).

students) had zero good answers. The second blue bar to the left indicates that about 20% of the 10th graders (20 out of 96 students) had only one good answer. Nevertheless, across all three grades, none of the students achieved more than 10 good answers. For better readability, the maximum value of the horizontal axis in Fig. 2 is kept to 10 instead of 20.

Students in different grades were shown in different colors in Fig. 2. The 12th graders performed better than the 10th and 11th graders. About 45% of the 12th graders had successfully binned six or seven questions into good categories. In contrast, only 1% of the 10th graders and 12% of the 11th graders had put more than six questions in the good categories.

Figure 2 counted only in the good categories, but the moderate categories also reflect students’ understanding of the questions. We assigned 1 point to each question binned in the good categories, 0.5 point to the moderate categories, and 0 points to the poor categories. The highest possible score on the 20 question categorization task would be 20 (if all questions were properly put into good categories). The average score for the 10th graders, 11th graders, and 12th graders were 1.70, 3.05, and 5.49, respectively.

Among the 20 problems in the questionnaire, the participating students performed better on Q9, Q14, and Q20. All three questions involved a surface feature of an object moving on an incline. The good category for Q9 was related to the work-energy theorem or Newton’s second law and kinematics, and the good categories for both Q14 and Q20 were related to mechanical energy conservation. About 20% of the 10th graders, 40% of the 11th graders and 60% of the 12th graders had binned these three questions into good categories.

None of the students in all three grades categorized Q19 (as shown in Fig. 1) into good or moderate categories. Some students were misled by the pictorial representation of a rotatory table in Q19 and mistakenly thought this question required rotational kinematics. Other students directly named their categories with the variable “velocity”

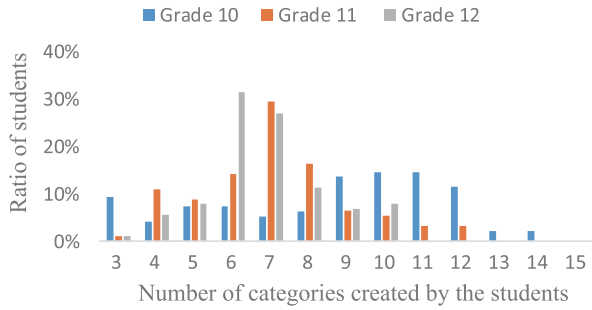


FIG. 3. The distribution of the number of categories created by students in different grades. The horizontal axis is the number of categories created by the students, and the vertical axis is the percentage of students in different grades (grade 10-blue, grade 11-orange, grade 12-gray).

asked in the question. Both the “circular motion” and velocity categories were treated as poor since they did not reflect the physical principle of 2D kinematics, which can be used to solve this problem. Students also had difficulty categorizing the questions related to the impulse-momentum theorem, such as Q4 and Q12. Most students categorized Q4 and Q12 into the poor categories of “collision,” “force,” or “tension.” A small portion of students noticed the physical concept of momentum but ignored the role of “impulse” involved in these two questions. Hence, their answers were rated as moderate.

Different students binned the 20 questions into different numbers of categories (including good, moderate, and poor categories). For example, one student may have used 6 categories, such as energy conservation, momentum conservation, force, velocity, spring, and ramp, to describe the 20 questions, while another student may have binned the 20 questions into 8 categories. We found that the students in higher grades created fewer numbers of categories in the categorization task. On average, the 10th graders created 8.7 categories, the 11th graders created 7.0 categories, and the 12th graders created 6.7 categories for the 20 questions. Figure 3 shows the distribution of the number of categories made by the students in different grades. The number of categories made by 10th graders was much larger than that made by 11th and 12th graders. Specifically, the lower and upper quartiles (25%–75%) of the number of categories were [6.25, 11.75] for 10th graders, [6.25, 8.75] for 11th graders, and [6.5, 8.5] for 12th graders.

IV. DISCUSSIONS

A. High school students’ categorization ability gradually developed

Our data showed that the scores on the categorization task increased for students from grade 10 to grade 12. The 12th graders’ scores were significantly higher than the 10th graders’ (p value < 0.01). Since the enrollment standard of the students in the same school was stable across different

years, this result suggested that the high school students’ ability of categorizing physics problems gradually developed as they took more physics courses. Further evidence of the improvement in categorization is the number of categories created by the students in different grades. On average, the 10th graders created 8.7 categories, while the 12th graders only created 6.7 categories. A lower number of categories indicated a better hierarchical knowledge structure of the physics concepts [25].

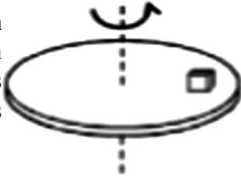
We noted that the 10th graders had a tendency to directly name a category with the variable asked in the question. For example, Q5 and Q6 used the same context but asked the students to solve for different variables. Q5 asked the students to find the velocity, while Q6 asked about the tension in the string. About 40% of the 10th graders binned Q5 in the category of velocity and Q6 in the category of tension, which were evaluated as poor categories. Compared to the 10th graders, the 11th and 12th graders used a more generalized name to describe Q5 and Q6. Only 12% of the 11th graders and 12% of the 12th graders directly categorized Q5 as velocity. Furthermore, 4% of the 11th graders and 8% of the 12th graders directly categorized Q6 as tension. The good category for Q6 should be Newton’s second law”. Forty percent of the 11th graders and 50% of the 12th graders used force composition to name the category of Q6, which was evaluated as a moderate category.

In the interviews, we asked three 12th grade students who used the category name force composition to speak out their reasoning process while solving Q6. One student drew a free body diagram but failed to apply Newton’s second and third laws. The other two students correctly applied Newton’s second and third laws to solve this problem and obtained the right answer. When we asked them to elaborate on the physical concepts and principles involved in their category of force composition, they claimed that it contained both Newton’s second and third laws. We agreed that when the question involved multiple objects, students usually applied Newton’s second and third laws simultaneously. However, since not all of the students who used the category force composition expressed a clear understanding of Newton’s laws in the interview, we still treated this category as moderate instead of good in our data analysis.

B. Pictures in the problems affected students’ categorization

We found that the pictures in the problems affected students’ categorization of the problems, especially for the 10th grade students. Q7 and Q15 described a similar scenario of the motion of objects on a rotating plate. There was a picture of a rotating plate in Q7 but not in Q15. Both questions mentioned friction in the context. The categories created by the students for Q7 and Q15 are listed in Table II and Table III. Sixty-two percent of the 10th

TABLE II. The percentage of the top 3 categories for Q7 created by the students in different grades.

<p>7. An object with mass m is rotating with the disk. Suppose the distance between the object and the center of the disk is r. The coefficient of kinetic friction is μ, and the maximum static friction is equal to the kinetic friction. Find the maximum angular velocity of the disk to keep the object rotating with the disk without sliding.</p>		
		
Grade 10	Grade 11	Grade 12
Angular velocity (62%)	Circular motion (67%)	Circular motion (70%)
Circular motion (27%)	Angular velocity (17%)	Angular velocity (20%)
Centripetal force (8%)	Friction (6%)	Velocity (7%)

graders categorized Q7 as angular velocity, which is a direct reproduction of the variable asked in the question, and 27% of the 10th graders binned Q7 in the category of circular motion. Therefore, altogether, about 90% of the 10th graders categorized Q7 according to the scenario of rotation, and none of the 10th graders used “friction” as the category name for Q7. Similarly, 85% of the 11th graders and 90% of the 12th graders binned Q7 in the categories related to rotation (i.e., angular velocity or circular motion) and few of them emphasized friction. However, for Q15 without the pictorial representation, the percentage of categories related to rotation dropped. Only 46% of the 10th graders, 56% of the 11th graders, and 68% of the 12th graders used angular velocity or circular motion as the category name. More students in the higher grade maintained a consistent judgment of the categories for Q15 and Q7. This result indicated that the influence of the pictures in the question reduced as the students gained more problem-solving experience. Similar findings were also reported in the area of mathematical education when the students were asked to categorize word problems [26].

Furthermore, in Q15 without the pictorial representation, 48% of the 10th graders, 13% of the 11th graders, and 8% of the 12th graders used friction as the category name. In the interview, we asked the students how they created the categories. Two students in grade 10 claimed that they tried to find a keyword when they were solving the physics problems, and they named their category with the keyword. The picture in Q7 acted as the keyword, so both of them directly related Q7 with rotation. However, since there was no picture in Q15, these two students had to search for the keyword in the context. One of the two students chose friction as the keyword and the category name because Q15 mentioned both “coefficient of friction” and “slide.” The other student chose motion as the keyword because Q15 asked the students to “describe the motion of the two

TABLE III. The percentage of the top 3 categories for Q15 created by the students in different grades.

<p>15. Two blocks A and B with equal masses were placed on a rotating disk. The coefficient of kinetic friction between the two blocks and the disk was the same. The two blocks were connected to each other by a string along the direction of the radius. The rotating speed of the disk gradually increased to a value such that the two blocks just started to slide. If the string was burned at this moment, describe the motion of the two blocks A and B.</p>		
Grade 10	Grade 11	Grade 12
Friction (48%)	Circular motion (54%)	Circular motion (68%)
Circular motion (40%)	Friction (13%)	Constant speed motion (13%)
Angular velocity (6%)	Velocity (8%)	Friction (8%)

blocks.” Then, this student used circular motion as the category name. When we asked him why he did not use motion as the category name, he replied that “motion was too broad for a category name... it should be circular motion because the blocks were placed on a rotating disk.” Their strategy of keyword searching also applied in Q2, as shown in Table IV. About 20% of the 10th graders binned Q2 in the category of “Hooke’s law” because they found the keywords “compressing the spring” and “the force constant of the spring” in the question. In contrast, none of the 12th graders used Hooke’s law as the category for Q2, and half of them binned this question in the good category of “work-energy theorem.”

C. Students’ categorization may be affected by their training in problem solving

The students in the 12th grade created some unique categories that the 10th and 11th grade students did not. For example, 10% of the 12th graders binned Q16 in the

TABLE IV. The percentage of the top 3 categories for Q2 created by the students in different grades.

<p>2. A 2-kg block is forced against a horizontal spring of negligible mass, compressing the spring by 15 cm. When the block is released, it moves 60 cm across a horizontal tabletop before coming to rest. The force constant of the spring is 200 N/m. What is the coefficient of friction between the block and the table?</p>		
Grade 10	Grade 11	Grade 12
Friction (70%)	Friction (50%)	Work-energy theorem (53%)
Hooke’s law (21%)	Work-energy theorem (20%)	Friction (22%)
Work-energy theorem (6%)	Hooke’s law (13%)	1D kinematics (10%)

category of “decomposition and superposition of vectors.” About 5% of the 12th grade students binned Q15 and Q17 in the category of “connected objects.” The high school physics teachers mentioned that during the review session for the 12th graders before the college entrance exam, the teachers sometimes organized the questions with similar scenarios together for the students to practice. The scenarios of decomposition and superposition of vectors and connected objects happened to be two of the reviewed topics. This finding reminded us that students’ training in problem solving may affect their style of categorization.

Another example also indicated that training in problem solving influenced students’ categorization. Two students in grade 12 created a category named “novel questions” for the questions they were not familiar with. None of the 10th and 11th graders made such categories. In the interview, the 12th graders confirmed that they had received intensive training in problem solving. When sitting for an exam, they could judge whether a problem was a routine question. These students’ behavior satisfied the characteristics of the “routine expert,” as described by Schwartz, Bransford, and Sears [27]. Hence, although the 12th grade students were more expertlike in the tasks of categorization, whether they were routine experts remains a question worth further investigation in the future.

We also administered the questionnaire to 60 first-year undergraduate students at the end of a calculus-based physics course. Note that none of the undergraduate students taking this survey had attended the high schools involved in our investigation. Hence, the results of the undergraduates should not be treated as part of the pseudolongitudinal research. However, when analyzing the categories created by the undergraduate students, we found some common features attributed to the intensive training of high school physics. Therefore, we include the results of the undergraduate students here as a supplementary data for the categorization tasks.

Six out of the 60 undergraduate students mentioned connected objects as a category when performing the categorization task. Moreover, three undergraduate students mentioned the “whole system method” as a category for questions Q5, Q6, and Q17. The instructor of their university physics curriculum had never taught such topics in lectures or recitations. We interviewed these undergraduate students, who had come from different provinces. They confirmed that their high school teachers used to emphasize the method of treating the system as a whole while solving dynamic problems that involved connected objects.

The experts we interviewed had relatively positive opinions about whether the whole system method should be treated as a good category. Some professors in the physics department thought that although the whole system method did not involve any physics principles, it somehow reflected the virtue of “thinking like a physicist.” The high school teachers provided an explanation for the category of

the whole system method from the perspective of problem solving. When evaluating students’ solutions to physics problems, the Chinese high school teachers often focused on three aspects: “skills,” “methods,” and “principles.” For example, to solve question Q5 (shown below), the students needed to grasp the skills of sketching the free-body diagram, choose the method of whether to analyze the two objects as a whole system or separately, and apply the principles of energy conservation or Newton’s second law. The appropriate method may provide some shortcuts to solve certain questions, which would be an advantage on exams with time limitations. Therefore, the high school students paid considerable attention to the method selection in the problem-solving process.

Q5. A man of mass M_1 lowers himself to the ground from a height X by holding onto a rope passed over a massless, frictionless pulley and attached to another block of mass M_2 on the other side. The mass of the man is greater than the mass of the block. With what speed does the man hit the ground after falling through a distance X ?

V. SUMMARY

We conducted a pseudolongitudinal study to investigate the development of novices’ expertise in categorizing kinematics and mechanics problems. Our data showed a gradual progression in categorization for the high school students as they took more physics courses. Although the high school students in all three grades performed as novices in conducting the categorization tasks, the 12th graders were closer to being experts than the 10th graders.

Compared to Chi’s studies, we found that although some of the novices categorized the questions according to surface features, such as pulley or connected objects, this may not always indicate a lack of physics knowledge. Intensive training in problem solving affected students’ categorization of the questions. The students categorized the questions with similar surface features together at first glance. Then, they chose a proper method from their organized problem-solving strategies that corresponded to these problems. To some extent, these students behaved as routine experts instead of novices. In future studies, new questionnaires of categorization tasks need to be carefully designed in order to effectively differentiate novices, routine experts, and “adaptive experts.”

ACKNOWLEDGMENTS

This work is sponsored by the Shanghai Sailing Program (Grant No. 16YF1415600), the National Natural Science Foundation of China (Grant No. 61473196), and the Key Research Institutes in Humanities and Social Sciences authorized by MOE (Grant No. 16JJD880023).

- [1] J. D. Bransford, A. L. Brown, and R. R. Cocking, *How People Learn: Brain, Mind, Experience, and School* (National Academy Press, Washington, DC, 1999).
- [2] A. H. Schoenfeld and D. J. Herrmann, Problem perception and knowledge structure in expert and novice mathematical problem solvers, *J. Exp. Psychol.* **8**, 484 (1982).
- [3] M. T. Chi, P. J. Feltovich, and R. Glaser, Categorization and representation of physics problems by experts and novices, *Cogn. Sci.* **5**, 121 (1981).
- [4] W. G. Chase and H. A. Simon, Perception in chess, *Cogn. Psychol.* **4**, 55 (1973).
- [5] D. L. Medin, Concepts and conceptual structure, *Am. Psychol.* **44**, 1469 (1989).
- [6] E. R. Savelsbergh, T. de Jong, and M. G. Ferguson-Hessler, Choosing the right solution approach: The crucial role of situational knowledge in electricity and magnetism, *Phys. Rev. ST Phys. Educ. Res.* **7**, 010103 (2011).
- [7] D. M. Bunce, D. L. Gabel, and J. V. Samuel, Enhancing chemistry problem-solving achievement using problem categorization, *J. Res. Sci. Teach.* **28**, 505 (1991).
- [8] S. Lin and C. Singh, Categorization of quantum mechanics problems by professors and students, *Eur. J. Phys.* **31**, 57 (2010).
- [9] A. Mason and C. Singh, Assessing expertise in introductory physics using categorization task, *Phys. Rev. ST Phys. Educ. Res.* **7**, 020110 (2011).
- [10] C. Singh, Categorization of problems to assess and improve proficiency as teachers and learners, *Am. J. Phys.* **77**, 73 (2009).
- [11] F. Reif and S. Allen, Cognition for interpreting scientific concepts: A study of acceleration, *Cognit. Instr.* **9**, 1 (1992).
- [12] C. Singh, When physical intuition fails, *Am. J. Phys.* **70**, 1103 (2002).
- [13] G. H. Veldhuis, Ph.D. Dissertation, Iowa State University, 1986, <http://lib.dr.iastate.edu/rtd/8502/>.
- [14] G. H. Veldhuis, The use of cluster analysis in categorization of physics problems, *Sci. Educ.* **74**, 105 (1990).
- [15] P. T. Hardiman, R. Dufresne, and J. P. Mestre, The relation between problem categorization and problem solving among experts and novices, *Memory & Cognition* **17**, 627 (1989).
- [16] M. T. H. Chi, M. Bassok, M. W. Lewis, P. Reimann, and R. Glaser, Self-explanations: How students study and use examples in learning to solve problems, *Cogn. Sci.* **13**, 145 (1989).
- [17] J. L. Docktor, J. E. P. Mestre, and B. H. Ross, Impact of a short intervention on novices' categorization criteria, *Phys. Rev. ST Phys. Educ. Res.* **8**, 020102 (2012).
- [18] R. L. Keith, Ph.D. Dissertation, University of Minnesota, 1993, <https://elibrary.ru/item.asp?id=5759325>.
- [19] J. P. Mestre, R. J. Dufresne, W. J. Gerace, P. T. Hardiman, and J. S. Touger, Promoting skilled problem-solving behavior among beginning physics students, *J. Res. Sci. Teach.* **30**, 303 (1993).
- [20] M. C. Lovett and J. R. Anderson, Effects of solving related proofs on memory and transfer in geometry problem solving, *J. Exp. Psychol.* **20**, 366 (1994).
- [21] L. Bao *et al.*, Learning and scientific reasoning, *Science* **323**, 586 (2009).
- [22] G. Zhu and Y. Shu, Research on classification method of high school students to the curve movement problems, *Phys. Engin.* **25**, 25 (2015).
- [23] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.13.020118> for the questionnaire used in the categorization tasks.
- [24] M. T. Chi, Quantifying qualitative analyses of verbal data: A practical guide, *J. Learn. Sci.* **6**, 271 (1997).
- [25] R. N. Giere, The cognitive structure of scientific theories, *Philos. Sci.* **61**, 276 (1994).
- [26] S. K. Reed, A structure-mapping model for word problems, *J. Exp. Psychol.* **13**, 124 (1987).
- [27] D. L. Schwartz, J. D. Bransford, and D. Sears, Efficiency and innovation in transfer, *Transfer of Learning from a Modern Multidisciplinary Perspective* (Information Age, Greenwich, CT, 2005), pp. 1–51.