Journal of Science Education

ISSN 0124-5481

# Evaluation of thinking ability in physical science of senior high school Students

Na Cui[a], Qiye Zhang[b], Weiming Zheng[c], Jianxin Li[d*]

[a] College of Physics and Electronic Technology, Liaoning Normal University, Dalian 116029, P.R. China
[b] Jiangsu Haizhou Senior High School, Lianyungang 222000, P.R. China
[c] Zhuhai Experimental Middle School, Zhuhai 519090, P.R. China
[d] Beijing Hongzhang Flower Data Technology, Beijing 102445, P.R. China

A R T I C L E I N F O

*Keywords:*
scientific thinking ability
high school physics
rasch model

A B S T R A C T

The core quality of physics in senior high school includes physical concept, scientific thinking, scientific inquiry, scientific attitude and responsibility. Among them, scientific thinking is an important training goal in high school physics curriculum. Therefore, taking the students of a middle school in Zhuhai as the research object, this paper makes an empirical study on the students' scientific thinking ability by using the Rasch model and other tools such as tests and questionnaires. The results show that the students' scientific thinking ability is good on the whole, the best performance in scientific reasoning and scientific demonstration, and the poor performance in scientific modeling ability.

## 1. Research background

"Scientific thinking" is a way of understanding the essential attributes, internal laws and mutual relations of objective things from the perspective of physics, and is the key to cultivating students' core physical literacy. Therefore, the cultivation of scientific thinking should not start from higher education, but should attract the attention of educators at least in the high school stage. Weiping Hu and Chongde Lin (Weiping Hu and Chongde Lin, 2003) pointed out that "scientific thinking ability of adolescents is a special ability, an organic combination of general thinking ability and science subject, a concrete expression of general thinking ability in science subject, a crystallization of development of general thinking ability and science education, and the core of scientific ability. In 2016, China proposed that basic education should focus on the cultivation of core qualities, requiring students to master the necessary knowledge and key abilities necessary for future development. In 2018, China promulgated the Physics Curriculum Standards for Senior High Schools (2017 edition)(Xiaotong Liu, 2018), in which discipline literacy is divided into four dimensions: physical concept, scientific thinking, experimental inquiry, scientific responsibility and attitude, among which scientific thinking is an important part of the core literacy. At the same time, scientific thinking also includes four elements: "model construction", "scientific reasoning", "scientific demonstration" and "questioning innovation" (The PRC Ministry of Education,2017). Therefore, in the study of physics, high school students should not only master physics knowledge, but also help students improve their comprehensive quality form scientific thinking through the development of their own psychological ability on the basis of what they have learned(Haiyang Gao, Meifen Yang, 2020) .This study tries to analyze and evaluate abstract and complex scientific thinking by means of test questions and computer software to examine students' thinking performance.

## 2. Design of evaluation tools

### *2.1 Research objects and tools*

The research object of this study is the students of Zhuhai high school, 158 people were measured, and 143 valid questionnaires were collected, the effective rate is 90.5%,61.5% (88 people) are boys, 38.4% (55 people) are girls. There are certain differences among student groups, which can be used as the research object of this test.

In order to avoid the influence of unfamiliar knowledge on the assessment results, the topics designed in this study are all important content that students are familiar with Celestial body motion, elastic potential energy and electromagnetic induction are mainly used as knowledge carriers for situational questions. Written test tasks are designed and scoring standards are designed for each question. The problems in the context are evaluated from nine dimensions: raising questions and making assumptions, designing experiments and generating data, interpreting data and drawing conclusions, viewpoints, facts and theoretical basis, reasoning and refutation, model construction and use, model testing and correction, modeling meta-cognition and meta-modeling knowledge. According to the

* Corresponding author: Jianxin Li
Email: lijianxinbd@126.com
Accepted 1 October 2022, Available online 10 July 2023

students' judgment and whether the answer is right or wrong, it is divided into three levels: Level1, Level 2 and Level 3. In general, this test can be used as a tool to measure the level of scientific thinking of students.

### *2.2 Quality inspection of assessment tools*

Referring to the existing empirical studies of Winsteps software and Rasch model in science education, this study presents the analysis results from five aspects, such as overall quality of work, unidimensionality, grading structure, project fitting and project-subject response.

143 sample data were imported into Winsteps software, and 27 items were estimated with no missing values in the sample data. The Rasch model will automatically set the average difficulty estimate value of the item to 0. If the difficulty estimate value of the subject is greater than 0, it means that the ability of the subject in the test is high or the difficulty of the item is low. If the difficulty value of the subject is less than 0, it indicates that the ability of the subject in the test is low or the difficulty of the item is high. As can be seen from Table 1, the estimated difficulty value of the subjects in this test is -0.37logits, and the average ability value of the samples of the subjects is less than 0, indicating that the performance of the subjects' academic thinking ability is poor or the assessment tools are difficult. The Error represents the difference between the theoretical model and the actual observed value and can be used to measure the accuracy of the Rasch measurement. From the results, the error values of subjects and items are 0.35 and 0.14 respectively, which are both within the acceptable range, indicating that the sample data is consistent with the measurement model. The MNSQ values of the subjects and the project Infit and Outfit are within the acceptable range of 0.7 ~ 1.3, and both Infit ZSTD and Outfit ZSTD are close to or equal to 0, indicating a high degree of fitting between the measured data values and the Rasch theoretical model. In this study, the separation degree of the subjects is 2.95, greater than 2; Item separation degree is 6.20, greater than 3; The reliability of the subjects is 0.90, greater than 0.8; The item reliability is 0.97, greater than 0.9, which is within the acceptable range, indicating that the assessment tool has good reliability and can distinguish the level of physical science thinking ability of the subjects.

**Table 1** Overall quality analysis of work

| | Measure | Error | Infit | | Outfit | | Seperation | Reliability |
|---|---|---|---|---|---|---|---|---|
| | | | MNSQ | ZTSD | MNSQ | ZTSD | | |
| Person | -0.37 | 0.35 | 1.01 | 0.0 | 1.04 | -0.1 | 2.95 | 0.90 |
| Item | 0.00 | 0.14 | 0.99 | -0.1 | 1.04 | 0.1 | 6.20 | 0.97 |

unidimensionality is a fundamental assumption of the Rasch model. According to the test results, the total variance explained by the test accounted for 45.3%, which was higher than the recommended value of 40% (Linacre, 2006). The first unexplained eigenvalue is 5.3, greater than 3, and the proportion of this eigenvalue is 10.8%, less than 15%. The ratio of the variance explained by the test to the variance of the first eigenvalue not explained is about 4.19, which is greater than 3. In summary, the single dimensionality of this test is good and in line with the Rasch model analysis hypothesis.

In terms of the fitting degree of projects, the difficulty values of 27 projects are in the range of -1.80~+2.26. The project with project number 9 is the most difficult, and the project with project number 22 is the least difficult. The Standard error of all project difficulty estimates is between 0.13~0.22, which is acceptable. Combining the above indicators, it can be seen that the reliability of the developed test tools is high.

In the scoring structure of the test questions, the scoring curve of each dimension has obvious peak and flat value, and covers a certain range in the horizontal coordinate, and the performance is good.

By converting the original score into logit score, the Rasch model enables the ability value and the item difficulty value of the subject to be compared on the same scale. As shown in Figure 1, the item-person map visually shows the relationship between the subject and the item. The central axis in the figure is a logarithmic scale, using the logit scale, where the numbers 2, 1, 0, -1, and -2 represent logit scores. The left side of the scale shows the distribution of subjects with different abilities. The ability of subjects gradually increases from bottom to top. A "#" indicates 3 subjects, and a "·" indicates 1 or 2 subjects. The right side of the scale shows the distribution of the difficulty of the item, increasing gradually from the bottom to the top. M on the left is the average value of the subjects, and M on the right is the average value of the difficulty of the item of the assessment tool. The difference is less than 1logits, indicating that the assessment tool is well-targeted (Bond and Fox, 2007) and can be accurately estimated. The difficulty of the 27 projects is basically evenly distributed, and there is no "bunching" phenomenon.

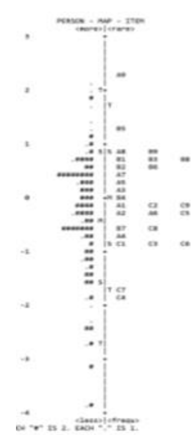

**Figure 1** Wright diagram

## 3. Students' scientific thinking performance

In order to further reflect the concrete manifestation of the scientific thinking ability of the subjects, this study uses SPSS software to make descriptive statistics in nine dimensions , the mean distribution of each dimension (Figure 2) and the percentage distribution of scores in each dimension (Figure 3). From Figure 2, it can be observed that the mean is within the range of 1.58-2.22, with a small standard deviation, and the overall performance of scientific thinking is good. Among them, students have the best performance in viewpoint, model construction and use, and scientific demonstration ability, with the average (standard deviation) of 2.22, 2.13, and 1.96, respectively. However, there are some deficiencies in model testing and revision, modeling meta-cognition and original modeling knowledge, with average values of 1.68 and 1.58 respectively, indicating that students do not apply empirical facts and existing theoretical basis well, and it is difficult to analyze phenomena and interpret data by building models. Students need to improve their ability in this aspect. From the percentage distribution of scores in each dimension, most of the subjects concentrated on level 1, which may be caused by the difficulty of the questions set.

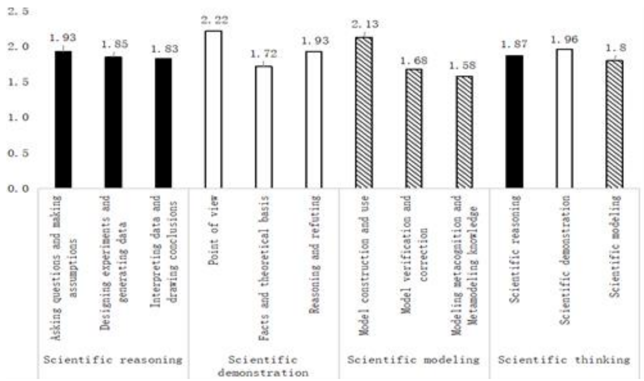


**Figure 2** Mean distribution of students' scientific thinking performance

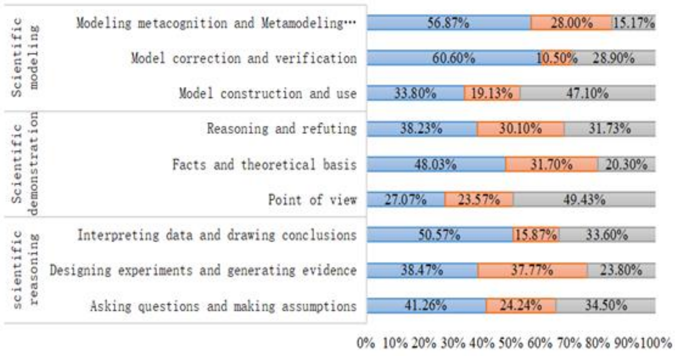


**Figure 3** Percentage distribution of scores in each dimension

SPSS software was used to draw the scientific thinking ability into the following standard P-P chart (Finger 4) to further test the normal distribution of the measured scores. The normal P-P chart examined the degree of agreement between the theoretical cumulative probability expected by the normal distribution and the actual cumulative probability. The scattered points are roughly clustered around the straight line, approximately a diagonal straight line, indicating that the data presents normality. In general, the distribution of scientific thinking ability of the subjects is close to normal distribution, and the sampling has good representativeness.

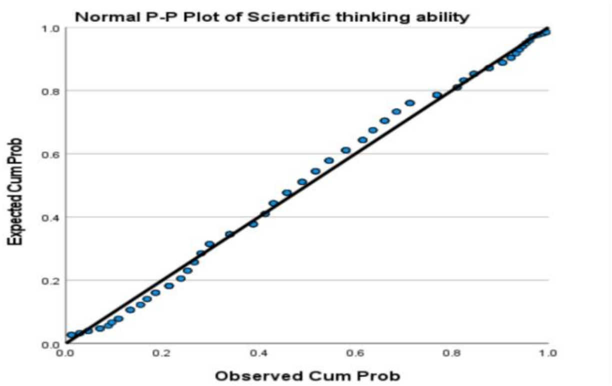


**Figure 4** Normal P-P graph of scientific thinking ability

## 4. The relationship between scientific thinking ability and gender

In order to explore the differences in scientific thinking ability of subjects of different genders, an independent sample t-test was conducted on the scientific thinking ability of students of different genders, and the results were shown in Table 2. As can be seen from the table, the average score rate of male subjects in the three dimensions of "scientific reasoning", "scientific modeling" and "scientific thinking" is slightly higher than that of female students, and the average score rate in the dimension of "scientific demonstration" is slightly lower than that of female students. The results of independent sample t-test shows that there are no statistically significant differences between the male group and the female group in the total score rate and score rate of each dimension of scientific thinking ability assessment, indicating that the level of scientific thinking has nothing to do with gender.

**Table 2** Gender differences in the total score of scientific thinking ability and scoring rate of each dimension

| | Scientific reasoning | | scientific demonstration | | scientific modeling | | scientific thinking | |
|---|---|---|---|---|---|---|---|---|
| group | male | female | male | female | male | female | male | female |
| M | 1.88 | 1.85 | 1.95 | 1.97 | 1.81 | 1.78 | 1.88 | 1.87 |
| SD | 0.47 | 0.50 | 0.46 | 0.51 | 0.45 | 0.47 | 0.43 | 0.45 |
| t | 0.340 | | -0.183 | | 0.327 | | 0.170 | |
| p | 0.481 | | 0.406 | | 0.975 | | 0.627 | |

## 5. The path analysis of the influence of students' scientific thinking

### *5.1 Correlation analysis of study variables and common method deviation test*

After correlation analysis of questionnaire results by IBM SPSS Statistics 27.0 software, variables with too little correlation are removed, and the following variables are finally selected for analysis. The correlation among variables is shown in Table 3. Among them, the correlation between teacher-student relationship and teacher teaching is the most significant (0.580).

In addition, this study uses principal component analysis to process the data, and calculates the total variance interpretation and its cumulative value. The results showed that the cumulative variance was interpreted as 34.258%. This suggests that some variation in the original variable can be explained by a small number of principal components.

**Table 3** Correlation analysis of research variables

| | Physics results | Teacher-student relationship | Teacher teaching | Learning skills | Learning motivation | Self-management | Scientific thinking ability |
|---|---|---|---|---|---|---|---|
| Physics results | | | | | | | |
| Teacher-student relationship | 0.119 | | | | | | |
| Teacher teaching | 0.163 | .580** | | | | | |
| Learning skills | .308** | .278** | .313** | | | | |
| Learning motivation | 0.136 | .274** | .192* | 0.134 | | | |
| Self-management | .181* | .187* | .379** | .500** | .168* | | |
| Scientific thinking ability | 0.195* | 0.033 | 0.123 | 0.068 | 0.108 | 0.061 | |

### *5.2 Model test and results of path analysis*

According to relevant literature, the influence path constructed in this study is as follows (Figure 5) and tested:

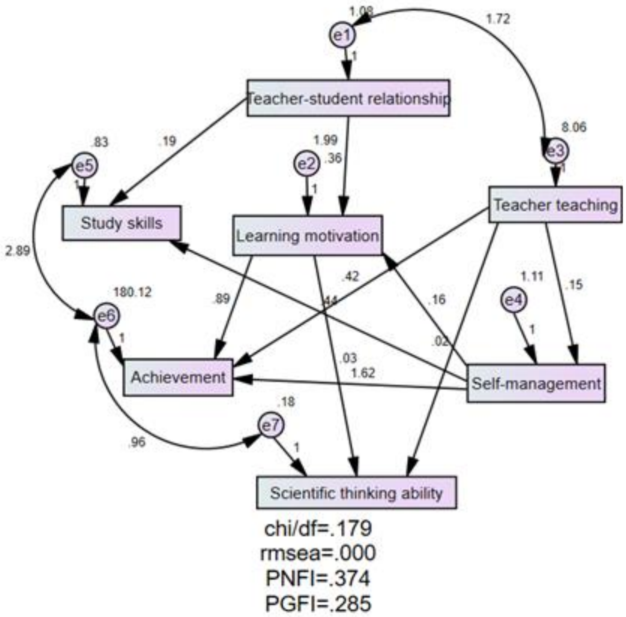


**Figure 5** Path analysis model

The model fit indicators obtained in this study include chi/df, RMSEA, PNFI, and PGFI. Specifically, the chi/df value is 0.179, which is less than 3; RMSEA value is less than 0.05; PNFI value is 0.374, less than 0.5; The PGFI value is 0.285, less than 0.5, and they are all within an acceptable range, indicating that the model performs well overall in terms of fit.

As shown in Table 4, several significant positive effects were found in this study. Firstly, teacher teaching has a significant positive impact on self-management (P<0.001), which means that teachers' guidance and attention can promote the improvement of students' self-management ability in the teaching process. Secondly, teacher-student relationship has a significant positive impact on learning motivation (P<0.05), which indicates that harmonious teacher-student relationship can stimulate students' learning motivation. Third, self-management has a significant positive effect on learning skills (P<0.001), which further emphasizes the importance of self-management in students' learning skills. Fourth, teacher- student relationship also has a significant positive impact on learning skills (P<0.05).

In addition, there are some direct and indirect effects of this model. Teacher teaching has direct positive effects on both self-management and scientific thinking ability, and the direct effect sizes are 0.380 and 0.104, respectively, indicating that teacher teaching can improve students' self-management ability and scientific thinking ability. The teacher-student relationship has a direct positive effect on learning motivation and learning skills, and its direct effect size is 0.251 and 0.187, indicating that a good teacher-student relationship can stimulate students' learning motivation and improve learning strategies. Self-management has a direct positive effect on learning motivation, learning skills and physics scores, and its direct effect sizes are 0.120, 0.463 and 0.134, respectively, indicating that the improvement of self-management ability can promote the improvement of students' learning motivation and skills, and also have a positive impact on physics scores. Learning motivation also has a positive effect on physics scores, indicating that the improvement of students' learning motivation is conducive to the improvement of their physics scores. Among them, the indirect effect size of teacher teaching on learning skills is 0.176, indicating that teacher teaching can affect the improvement of students' learning skills through the mediating effect of self-management.

**Table 4** Path analysis results

| | | | Estimate | S.E. | C.R. | P | Label |
|---|---|---|---|---|---|---|---|
| Self-management | <--- | Teacher teaching | .152 | .031 | 4.881 | *** | par_8 |
| Learning motivation | <--- | self-management | .156 | .108 | 1.445 | .149 | par_2 |
| Learning motivation | <--- | Teacher-student relationship | .358 | .118 | 3.025 | .002 | par_13 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Physical achievement | <--- | Teacher teaching | .417 | .424 | .985 | .325 | par_1 |
| Physical achievement | <--- | Self-management | 1.619 | 1.062 | 1.524 | .128 | par_3 |
| Physical achievement | <--- | Learning motivation | .893 | .760 | 1.175 | .240 | par_4 |
| Study skills | <--- | Self-management | .441 | .070 | 6.315 | *** | par_7 |
| Scientific thinking ability | <--- | Learning motivation | .026 | .025 | 1.040 | .298 | par_10 |
| Scientific thinking ability | <--- | Teacher teaching | .016 | .013 | 1.226 | .220 | par_11 |
| Study skills | <--- | Teacher-student relationship | .195 | .075 | 2.593 | .010 | par_12 |

## 6. Research results and teaching suggestions

In this study, Winsteps and SPSS were used to analyze 143 valid questionnaires, and the following results were obtained: Students performed better in scientific reasoning and scientific argumentation; Poor performance in scientific modeling dimensions, especially in model testing and revision, modeling meta-cognition and original modeling knowledge, indicating that students have difficulty in analyzing phenomena and interpreting data by building models; Secondly, the overall distribution of scientific thinking ability of the subjects is close to normal distribution, and the performance is good. Moreover, the level of scientific thinking ability has nothing to do with gender, and the excellent teaching of teachers can also promote the improvement of students' scientific thinking ability.

Based on this, this paper puts forward the following suggestions to improve the teaching and learning of scientific thinking ability: First, make use of physical phenomena in life to train students to build models. In the process of solving physical problems, students can try to combine real physical phenomena with physical knowledge, which not only stimulates students' interest, but also allows students to experience ways to find and solve problems from life. This knowledge model summarized from real situations is more meaningful than directly telling students how to construct. At the same time, it can be expanded on the basis of the original physical problem, using analogy and other methods to reproduce the establishment of this model and the exploration process, so as to realize the transfer of thinking ability. Second, students do hands-on experiments and experience the process of building models. Through personal experimental perception, students can improve their ability to test and revise models, and experience the joy of successful exploration, thus generating good learning motivation and effectively improving their scientific thinking ability. In addition, design some open, progressive questions that can cause students' cognitive conflicts, jump out of the question sea tactics, and develop students' questioning and innovation ability, reasoning ability and argumentation ability. Finally, the history of physics is used to promote the development of students' scientific thinking ability. The history of physics reveals the way of thinking of exploring and solving problems, which is an important means to cultivate students' scientific thinking ability. At the same time, the spirit of perseverance is cultivated, which will not only improve students' scientific thinking ability but also promote the improvement of physics teaching quality in middle school.

## References


Weiping Hu , Chongde Lin . Research on Scientific thinking ability of adolescents, J. Educational Research,2003(12):19-23.

Xiaotong Liu . Research on high school Physics Teaching Design based on Physics Core Literacy, D. Dalian: Liaoning Normal University,2018,1-22.

Ministry of Education of the People's Republic of China. Physics Curriculum Standards for senior high schools, M. Beijing: People's Education Press, 2017:3-5.

Haiyang Gao , Meifen Yang. Analysis of Core Literacy of high school physics curriculum standards, J. Mathematical and Physical Problem Research,2020(27):40-41.