A COMPARATIVE STUDY ON THE COMPREHENSIVE DIFFICULTY OF JUNIOR HIGH SCHOOL NATIONAL EXAMINATION

Xiaohong Wu¹, Ying Zhou¹ Zongzhao Mo²
¹ Department of mathematics and statistics, Guangxi Normal University, China
² Department of Psychology, Shenzhen University, China
Email: zhouying66@mailbox.gxnu.edu.cn

Abstract
Mathematics’ National examination is conducted annually in all countries, including Indonesia and China. But there are only a few researchers who make observations about the difficulty of the national exams in each country. There is not even a paper comparing the national exams in 2 countries. So in this study, This article compares the comprehensive difficulty of the Nanjing and Nanning test mathematics’ National examination questions from the five difficulty factors of background factors, cognitive level, reasoning ability, calculation level, and knowledge content. The method in this research is a mix method, using qualitative and quantitative for maximum data analysis results. The results show that, in terms of background factors, the test questions of Nanning paper The comprehensive difficulty is higher than that of the Nanjing paper. In the other four factors, the comprehensive difficulty of the Nanjing paper is slightly higher than that of the Nanning paper. The comprehensive difficulty coefficient of the two sets of test papers is slightly larger in terms of the calculation level and the cognitive level. The set of test papers basically maintains a balance of difficulty in the five factors.

Keywords: comprehensive difficulty coefficient; Junior High School National examination, mathematics test questions; comparative study

INTRODUCTION
National exams can be used as material for evaluating mathematics lessons in schools. From the results of the national exams, the government can set the right strategy to improve students' mathematical abilities and arrange study plans for the next academic year. With this evaluation, the government can continue to improve the quality of education in Indonesia. As an acceptance test in the nine-year compulsory education stage, the National examination has attracted the attention of teachers and students (Supriyono, Siroj, & Zulkardi, 2012). Teachers cannot do without the study of the difficulty of the middle school examination questions in daily teaching. Therefore, the difficulty of the test questions affects the difficulty of the course to a certain extent. The difficulty of middle
school examination questions is "too high" or "too low", which is not conducive to the development of students. Only when the difficulty of the test questions is control at an appropriate level of difficulty can the students develop better. Nanjing and Nanning are the capitals of Jiangsu and Guangxi, Are also the "leaders" in the development of basic education in Jiangsu and Guangxi respectively. According to the existing mature comprehensive difficulty model, through comparative analysis of the difficulty of the mathematics test questions in Nanjing and Nanning, aiming to be able to provide mathematics test questions in the two places. The order system and optimization of the test paper structure provide data reference, learn from each other’s strengths, and promote the better development of basic education.

In analyzing national exam questions, researchers must look at various aspects such as whether there is a relationship between the questions and students' daily lives? (Dini, Wijaya, & Sugandi, 2018) is there a high order thinking ability in every question? (Surya, Zulfah, Astuti, Marta, & Wijaya, 2020) What is the content knowledge aspect of a question, Computing level, and most importantly the cognitive level (Jonsson, Norqvist, Liljekvist, & Lithner, 2014). There have been many studies conducted on textbook analysis (Aditya, Wijaya, Dewi, & Zulfah, 2020), student ability analysis (Kulsum, Hidayat, Wijaya, & Kumala, 2019; Zuyyina, Wijaya, & Senjawati, 2018), development of learning modules, analysis of student practice questions. but there was no comparison of the national exams between the 2 countries. The results of this study can be used as an evaluation material for each country in forming national exam questions.

METHOD

The research method used in this research is the mix method (Mukminin et al., 2019). Select the 2019 Nanjing Paper test questionsPaper and 2019 Nanning City Paper test questionsPaper as the research objects. In particular, the 2019 Nanning Paper test questionsPaper title is "2019 Guangxi Bay" Mathematics papers of the junior high school level examination".

According to the specific situation of the mathematics test paper of the senior high school paper test questions, on the basis of the difficulty model of the scholar, we study the comprehensive difficulty model of the college paper test questions questions. The adapted comprehensive difficulty model contains background Factors, cognitive level, reasoning ability, calculation level, knowledge content five difficulty factors, and each difficulty factor is divided into different levels, and each indicator is attached with an explanation, and the different levels of each difficulty factor are assigned Certain weight coefficients, codes, etc. Finally formed the comprehensive difficulty coefficient model level division and coding table shown in Table 1, and constructed a comprehensive difficulty coefficient model suitable for the math test questions of the high school paper test questions. The formula for calculating the comprehensive difficulty coefficient is:
\[ d_i = \frac{\sum n_{ij} d_{ij}}{n} \quad (\sum n_{ij} = n, i = 1, 2, 3, 4, 5; j = 1, 2, \ldots) \]

Among them, in turn are the values of the five difficulty factors: background factors, cognitive level, reasoning ability, calculation level, and knowledge content. It is the weight of the first level of difficulty factor, which means that this group of questions belongs to the first level. The number of questions at the first level of the difficulty factor, the sum of which is equal to the total number of questions in the group.

Tabel 1

*The level division and coding table of the comprehensive difficulty coefficient model based on the middle exam questions*

<table>
<thead>
<tr>
<th>Difficulty factor</th>
<th>Level division</th>
<th>Descriptive</th>
<th>Code</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>No background</td>
<td>Did not create a situation, just start the problem of mathematical knowledge itself</td>
<td>A1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Life background</td>
<td>Eliciting problems in the context of real life is related to solving real life problems</td>
<td>A2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Scientific background</td>
<td>Expand with knowledge of other subjects, including backgrounds such as mathematical graphics and images</td>
<td>A3</td>
<td>3</td>
</tr>
<tr>
<td>Cognitive level</td>
<td>Memorize</td>
<td>Problems can be solved directly by the memory of mathematical facts, concepts, formulas, rules, and properties</td>
<td>B1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>understanding</td>
<td>Able to select mathematical knowledge and methods reasonably, and simply decompose or transform the problem, so as to answer the problem</td>
<td>B2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>application</td>
<td>Be able to discover new relationships, use and organize existing concepts and operations in new situations</td>
<td>B3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Explore</td>
<td>Be able to expand the knowledge of already learned mathematics, establish mathematical models, form mathematical conjectures and create and use mathematical strategic knowledge</td>
<td>B4</td>
<td>4</td>
</tr>
</tbody>
</table>
### RESULTS AND DISCUSSION

According to the definition of different difficulty factors in Table 1, the Nanjing Paper test questions and Nanning Paper test questions were separately coded and counted separately. The Nanjing paper has 27 questions, 6 multiple-choice questions, 10 fill-in-the-blank questions, each with 2 points, a total of 32 points, 11 answers, 7-11 points each, a total of 120 points; Nanning volume has 12 multiple-choice questions, fill-in the blanks There are 6 questions, 3 points for each question, 8 questions for answering questions, 6-10 points for each question, a total of 120 points. Although the number of questions in the two sets of test papers is different from the distribution of scores, the total score is the same, in order to balance the number of questions and points For the error caused by the
value, we code the multiple-choice and fill-in-the-blank questions once for each question, and the answer questions once for each small question. In particular, the 24th and 25th questions of the Nanjing Volume have 8 points for each question. These two questions perform double coding; Question 21 (1) (2) of the Nanning Volume is coded once, and question (3) is coded once. This question is coded twice. According to the above method, the two sets of test papers can be coded. Similarly, each set of test papers has 36 codes in total. The coded data is statistically sorted and the comprehensive difficulty coefficient of each factor is calculated using formula (1). Finally, the quantification of the difficulty factors of the Nanjing and Nanning papers as shown in Table 2 Indicator statistics table.

Tabel 2

Statistical Table of Quantitative Indexes for Each Difficulty Factor of Nanjing Volume and Nanning Volume

<table>
<thead>
<tr>
<th>Difficulty factor</th>
<th>Level division</th>
<th>No background</th>
<th>Nanjing</th>
<th>Nanning</th>
<th>Percentage</th>
<th>Comprehensive difficulty factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>Life background</td>
<td>13</td>
<td>6</td>
<td>36.11%</td>
<td>16.67%</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>Scientific background</td>
<td>13</td>
<td>15</td>
<td>36.11%</td>
<td>41.67%</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>Memorize</td>
<td>10</td>
<td>15</td>
<td>27.78%</td>
<td>41.67%</td>
<td></td>
</tr>
<tr>
<td>Cognitive level</td>
<td>understanding</td>
<td>4</td>
<td>6</td>
<td>11.11%</td>
<td>16.67%</td>
<td>2.56</td>
</tr>
<tr>
<td></td>
<td>application</td>
<td>12</td>
<td>13</td>
<td>33.33%</td>
<td>36.11%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Explore</td>
<td>16</td>
<td>12</td>
<td>44.44%</td>
<td>33.33%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No reasoning</td>
<td>4</td>
<td>5</td>
<td>11.11%</td>
<td>13.89%</td>
<td></td>
</tr>
<tr>
<td>Reasoning ability</td>
<td>Simple reasoning</td>
<td>7</td>
<td>10</td>
<td>19.44%</td>
<td>27.78%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Complex reasoning</td>
<td>17</td>
<td>16</td>
<td>47.22%</td>
<td>44.44%</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>No operation</td>
<td>12</td>
<td>10</td>
<td>33.33%</td>
<td>27.78%</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Research results and analysis

Based on the statistical data in Table 3, based on the percentage data of the five difficulty factors: background factors, cognitive level, reasoning ability, calculation level, and knowledge content, five broken-line statistical graphs were made, and based on the comprehensive difficulty of each of the five difficulty factors The coefficients are used to make a comprehensive difficulty coefficient radar chart. The following will compare and analyze the various factors of Nanjing Paper test questions and Nanning Paper test questions according to the line chart, and compare the comprehensive difficulty coefficients between the factors according to the radar chart.

1. Background factor

The best mathematical problems are those related to everyday problems (Aminah, Wijaya, & Yuspriyati, 2018; Bernard & Chotimah, 2018; Tommy Tanu Wijaya & Afrilianto, 2018). Therefore, it is hoped that the presentation of questions related to people, science, or students' daily lives is expected (Dewi, Mediyani, Hidayat, Rohaeti, & Wijaya, 2019; T.T. Wijaya, Sukma, Purnama, & Tanuwijaya, 2020; Tommy Tanu Wijaya, Purnama, & Tanuwijaya, 2020). As can be seen from Figure 1, in terms of life background, Nanjing Paper is 36.11%, Nanning Paper is 41.67%, Nanning Paper’s living background level is 5.56% higher than Nanjing Paper; in terms of scientific background, Nanjing Paper has 27.78 questions. %, the Nanning paper accounted for 41.67% of the test questions.
which was 13.89% higher than the Nanjing paper. It can be seen that in terms of background factors, there are differences between the Nanjing paper and the Nanning paper. The Nanning paper has higher background factors than the Nanjing paper.

![Figure 1. Line chart of changes in background factors at different levels](image)

2. Cognitive level

The cognitive level theory can be seen in the bloom theory (Anderson, 2001; Babinčáková, Ganajová, Sotáková, & Bernard, 2020; Duval, 2014). In this concept, it is hoped that questions containing explore and application can be higher than other types of questions (Koh, 2006; Weiler, 2005). It can be seen from Figure 2 that the two sets of test papers are relatively similar in the cognitive level of the three cognitive levels of memorization, comprehension, and inquiry, but there are obvious differences in the application level. The Nanjing paper focuses on application. The Nanning paper focuses on the examination of the level of comprehension; the Nanjing paper accounts for 44.44% of the application level examination questions, which is close to half of the whole set of examination papers, while the Nanning paper accounts for 33.33%. The Nanjing paper has a higher application level than the Nanning paper. It is 11.11% higher. The examination of the two sets of examination papers focuses on the level of comprehension and application. From the overall trend, the assessment of the cognitive level of the Nanning paper is relatively balanced.
3. Reasoning ability

Mathematical reasoning ability is one of the High order thinking skills (HOTS), and is the core and target of mathematics lessons in schools (Chotimah, Wijaya, Aprianti, Akbar, & Bernard, 2020; Hendriana, Putra, & Hidayat, 2019; Wahyu Hidayat, 2017). Reasoning ability is also one of the abilities that must be mastered by students (W. Hidayat, Wahyudin, & Prabawanto, 2018; Soifer, 2017). It can be seen from Figure 3 that both papers focus on the simple reasoning ability test, and the test intensity is close to half of the total number of questions, but the Nanjing paper has a 2.78% higher test intensity for simple reasoning than the Nanning paper; the Nanning paper has no reasoning ability and complex reasoning The test of ability accounted for 27.78%, and the test of non-reasoning ability in Nanjing paper was 19.44%, which was lower than 8.34% of Nanning paper, and the test of complex reasoning ability was 33.33%, which was higher than that of Nanning paper 5.55%. It can be seen that the trends of the two sets of test papers in the three levels of reasoning ability are relatively similar. It can be seen in Figure 3, the national exam questions in China contain questions with a reasoning ability above 30%.

Figure 2. Line chart of changes in different levels of cognition
Figure 3. Line chart of changes in different levels of reasoning ability

4. Computing level

It can be seen from Figure 4 that both volumes emphasize the examination of the level of simple numerical calculations. The proportions of the two levels of no calculations and numerical calculations are similar, but there are differences in the two levels of simple symbolic calculations and complex symbolic calculations. At the level of simple symbolic calculations, Nanjing volume accounted for 36.11%, Nanning volume accounted for 44.44%, at this level Nanjing volume was lower than Nanning volume; in complex symbolic calculation, Nanjing volume accounted for 22.22%, Nanning volume accounted for 13.89%, and Nanjing volume accounted for 22.22%. This level is higher than Nanning paper test questions.
5. Content knowledge

It can be seen from Figure 5 that the Nanjing paper focuses on the examination of 2 knowledge points, while the Nanning paper focuses on the examination of 1 knowledge point; the two volumes have the same examination force on 3 knowledge points; the Nanning paper focuses on the examination of more than 4 knowledge points. The strength is greater than the Nanjing Paper test questions, which shows that the Nanning Paper test questions pays more attention to the examination of comprehensive knowledge application than the Nanjing Paper test questions.
Figure 5. Line chart of changes of Content knowledge

Comprehensive difficulty analysis

It can be seen from Figure 6 that except for the background factors, the comprehensive difficulty coefficients of the Nanjing Paper test questions are slightly higher than those of the Nanning Paper test questions, and the gap between the two volumes in factors other than the background factor is small, the gap is 0.06 ~ 0.14. In terms of background factors, there is a significant difference in the comprehensive difficulty coefficient of the two volumes. The Nanning volume is 2.25 and the Nanjing volume is 1.92. The Nanning volume is 0.33 higher than the Nanjing volume. It can be seen that the Nanning volume is in The background factor is more difficult than the Nanjing paper test questions, and in other factors, the Nanjing paper test questions is slightly more difficult than the Nanning paper test questions.
The test questions focus on creating problem situations to promote the improvement of students' mathematics literacy (Dinni, 2018; Fachrudin et al., 2019; Retnawati & Wulandari, 2019). Through research, it is found that both papers focus on the creation of test problem scenarios, especially the creation of life problem scenarios. For example, the subway line 3 day created in Nanning paper test questions The average passenger flow created in the Nanjing paper test questions, weather conditions, summer social practice, rectangular square expansion, road construction The shortest path and other problem scenarios, combined with local, social and national hot topics, not only cultivate students' awareness of applied mathematics, promote the improvement of students' mathematics literacy, but also promote students' attention to social hot topics and cultivate patriotism.

 Appropriately adjust the difficulty of the test questions and implement a test-setting system based on the curriculum standards. The research results show that the calculation level and cognitive level of the two sets of test papers are relatively difficult, and the test questions can be set according to the cognitive level requirements of the curriculum standards. Appropriately reduce the difficulty of the two factors. Reasonable design and implementation of written tests can help comprehensively examine students' mathematics academic achievements, provide timely feedback on teaching effectiveness, and continuously improve teaching quality.
CONCLUSION

Based on the results of this study, it shows that the national exam in China has many factors and a difficult level of difficulty. This study compares 2 national exam questions in 2019 in the Nanjing region and Nanning China. it can be seen that, in terms of background factors, the comprehensive difficulty of the Nanning paper is higher than that of the Nanjing paper, while in the other four factors, the comprehensive difficulty of the Nanjing paper is slightly higher than that of the Nanning paper; two sets The comprehensive difficulty coefficient of the test papers on the two factors of calculation level and cognitive level is slightly larger, but the two sets of test papers. Suggestion for Future studies can compare the difficulty of exam questions between countries. for example the comparison of the Indonesian and Chinese national exams.

REFERENCES


https://doi.org/10.1088/1742-6596/1417/1/012042


