A Comparative Study of the Reliability and Validity of China's Postgraduate Enrollment System

Lan Yu*

School of Public Affairs, University of Science and Technology of China,
No.96 Jin Zhai Road, Hefei 230000, China
* E-mail of the corresponding author: 1833684865@qq.com

Abstract
This study aimed to evaluate the validity and utility of knowledge test and situational judgement test (SJT) for postgraduate selection. In a sample of 310 applicants, this study shows that SJT significantly related to creative performance and the SJT had incremental validity over the knowledge test. At last, the study proposes countermeasures and suggestions for the effective evaluation of candidates' innovative ability and academic potential in postgraduate selection.
Keywords: creative performance, situational judgment test, declarative knowledge, procedural knowledge
DOI: 10.7176/PPAR/9-2-01

1. Introduction

Postgraduate students are high-level talents trained by colleges and universities. They are also the main force of innovative talents in China. Their innovative ability and innovation level are of great significance to promoting the construction of world-class universities and world-class subject in China (Liu Dan *et al*. 2015). Accurate assessment of candidates' innovative ability and future academic development potential in the postgraduate entrance selection process is the key to improving the validity of postgraduate admission and the quality of postgraduate training (Chen Qian 2016). At present, the current selection of graduate students in China is mainly based on knowledge tests. Knowledge tests more focus on declarative knowledge of concepts, propositions, rules, theorems, and theories, but less on the specific behaviors and procedural knowledge of "how to do." As a result, the selected students are likely to have the phenomenon of "high scores and low energy". In recent years, the Ministry of Education has clearly stated that it is necessary to deepen the reform of the postgraduate enrollment mechanism and strengthen the investigation of graduate students' innovative ability. Domestic scholars Wang Chuanyi and Cheng Zhe (2017) suggested that situational and case-based topics should be included in the selection of graduate students. Lu Xiefeng and Shu Wenhui (2017) also clearly pointed out that the situational judgment test can provide standardized evaluation tools for graduate selection.

The Situational Judgment Test (SJT) is a low-fidelity simulation test. According to the behavioral judgment and behavior choice of the subject in the context directly related to the work it can judge whether the subject has the relevant knowledge and skills required for the job (Christian *et al*.2010). A large number of studies at home and abroad have shown that SJT can comprehensively examine the declarative and procedural knowledge of the subjects and effectively predict the performance of the subjects at work (O'Connell *et al*.2007). In terms of student selection, Lievens (2013) and Song (2015) have shown that SJT can better predict students' academic performance. However, at present, SJT has relatively few theoretical and empirical studies on graduate selection. Although Lu Xiefeng and Shu Wenhui (2017) developed the graduate competency situational judgment test, they did not verify its validity in graduate selection. At the same time, the existing research also lacks comparison of the validity of SJT and traditional knowledge test (graduate entrance examination) in the same environment and sample on predict the future performance of graduate students. Therefore, the purpose of this study is to develop a situational judgment test in the selection of graduate students, and to test the relative validity and incremental variance in graduate innovation performance over knowledge test.

2. Theory and Hypothesis

The innovation of graduate students is mainly a knowledge innovation. On the basis of knowledge acquisition, accumulation and processing, it constantly proposes new concepts, explores new laws and verifies them through literature research and experimental research (Wang Ying *et al*.2014). The research by Zhu Hong *et al* (2011) shows that the innovation achievements of graduate students during their school years largely depend on the construction of their knowledge systems. According to modern cognitive psychology, knowledge can be divided into two categories: declarative knowledge and procedural knowledge (Anderson *et al*.1985). Procedural knowledge refers to abstract knowledge about concepts, attributes, and general principles that can be transformed into word form, such as concepts, theories, propositions, and theorems. Procedural knowledge, also called operational knowledge, is a set of operational steps and is about "how to do". Past research on employee selection suggests that declarative and procedural knowledge should also have an important impact on graduate students' innovation ability and innovation performance (Lievens & Patterson 2011). Specifically, the storage of

declarative knowledge is an important premise and foundation for the formation of innovative thinking; while procedural knowledge has the characteristics of rapid, unique and flexible migration, which is conducive to the development of innovative capabilities and the realization of innovative results.

2.1 SJT and graduate innovation performance

SJT is a simulation test, and its prototype first appeared in the Social Intelligence Test of the University of Washington in 1926 (McDaniel *et al*.2001). In the 1990s, Motowidlo *et al* (1990) scholars found that SJT can effectively measure individual competencies, which has caused widespread concern about SJT. The Behavioral Consistency proposed by Schmitt and Ostroff (1986) provides theoretical support for the predictive validity of the situational judgment test. That is to say, the individual's behavior and behavior style have certain similarity and stability, and the performance of the individual in the work simulation scenario can predict its performance in actual work well. Research in student selection scenarios also shows that SJT has a good predictive power on students' academic performance and potential. For example, Lievens and Coetsier (2002) empirically based on student samples, SJT can effectively predict the academic performance of medical students in the first year (r = 0.23, p <0.01). Lu Xiefeng and Shu Wenhui (2017) showed that SJT can effectively evaluate the ability of postgraduate task processing and interpersonal communication through the measurement of 346 graduate students. Based on the SJT test principle and the existing empirical research, we propose the following assumptions:

Hypothesis 1: There is a positive correlation between SJT scores and graduate students' innovation performance.

2.2 SJT and knowledge test

First, as a type of paper and pencil test, knowledge tests and SJT can be used for the assessment of declarative knowledge such as basic knowledge and skills. Specifically, by filling in the blanks, judgments, choices, questions and answers, the knowledge test mainly examines the declarative knowledge such as professional concepts, principles, methods, etc. The meta-analysis of Schmidt and Hunter (1998) shows that knowledge test can effectively predict individual training and job performance. SJT mainly examines the subjects by setting up scenarios in which the subjects need to be identified and judged through professional knowledge. For example, the SJT item for doctor testing "You are a doctor, a rheumatoid arthritis patient asks you if you can give up taking methotrexate, what do you do then?" The disease and pharmacology of the subjects were investigated through specific doctor-patient situations.

Secondly, Motowidlo (2006) and other scholars' knowledge determinant theory suggests that SJT can measure the procedural knowledge of the subject in addition to the specific declarative knowledge. That is, whether the subject can effectively judge the effectiveness and ineffectiveness of each behavioral response described in SJT based on his or her general knowledge and specific work knowledge, thereby selecting the most effective behavior in this scenario. Based on the SJT example in which the above patients refused to take the drug, when the subjects determined that they could meet the patient's request based on their declarative knowledge, they also needed to use procedural knowledge to determine how to effectively communicate their decisions to the patient. At the same time, Lievens and Coetsier (2002) through the selection test of 941 medical students also showed that SJT can effectively supplement the traditional knowledge test and improve the predictive validity of students' academic performance. The postgraduate selection mainly uses the postgraduate entrance examination to test the knowledge level of the candidates. Therefore, we propose the following research hypotheses:

Hypothesis 2: SJT has incremental validity for knowledge test (graduate entrance test scores), which can improve the predictive validity of postgraduate innovation performance in selection.

3. Method and Process

3.1 Development of Postgraduate Innovation Situational Judgment Test

Preparation of items. According to the research of Scott and Bruce (1994), the generation of innovation performance can be divided into three stages: the generation of innovative ideas, the improvement of innovative ideas, and the realization of innovative ideas. In order to more fully understand the main tasks and behavioral activities of graduate students in each stage of innovation performance. We randomly selected 5 teachers and 25 graduate students from our university for key event interviews. Respondents are asked to describe the most successful and unsuccessful events they believe to be in the three stages of innovation. By reviewing the relevant research materials and collating the interview materials, we have compiled a total of 24 postgraduate innovative situational judgment test questions, and each situational contains 4-6 response items.

Screening of scenarios and response options. Two professors and five graduate students in the human resources assessment major were invited to form a team of seven experts. The panel of experts used the method of subject expert assessment to further screen and improve the prepared scenarios. First, experts judged and screened the 24 scenarios based on whether the scenario was related to the innovation activities of graduate

students, whether they were typical and common, and whether there were duplications between scenarios. Secondly, the panel used independent evaluate all the reaction options under each scenario using the Likert 6-point scale, with 1 indicating "very disagree" and 6 indicating "very agree". We summarize the results of the expert's scoring and calculate the average and standard deviation of each response option. Delete the response options where the expert evaluation scores are inconsistent, that is, the standard deviation is greater than 1. So far, four reaction options have been screened under each scenario. Finally, the experts revised the remaining 17 scenarios and 68 response options to form a preliminary version of the graduate innovation situational judgment test covering the main tasks of each stage of postgraduate innovation.

Table 1. Example of test questions

Example: You are a graduate student. When you read the literature, you can see some research concepts or research methods that you don't understand. What do you do then?

A Find relevant materials and learn this research concept

1 2 3   4   5 6

B Ask classmates to see if they can answer for yourself

1 2 3 4 5 6

C Record it first, then learn slowly later.

1 2 3 4 5 6

D Skip the concepts and methods you don't understand and see what you can understand.

1 2 3 4 5 6

Test. The behavior-oriented guidance was chosen for the test. It requires the participants to imagine themselves as a graduate student, what will happen when the problem setting situation occurs in the process of scientific research and innovation? According to their own practice, the subjects used the Likert 6-point scale to judge the four reaction options listed in each scenario from "1 very disagree" to "6 very agree". The example below shows. Using the research of scholars such as McDaniel (2011), the dichotomous consensus was used to calculate the scores of the subjects on each reaction option.

The result. A sample of 166 graduate students from a university in Anhui Province was used as a sample to test the postgraduate innovation situation judgment test. A sample of 166 graduate students from a university in Anhui Province was used as a sample to test the postgraduate innovation situation judgment test. From the perspective of grade distribution, research one accounted for 38.6%, research two accounted for 38.6%, and research three accounted for 22.9%.We processed the collected sample data using Spss23.0. By analyzing the correlation between each reaction option score and the total test score, we examined the consistency of each reaction option. The results show that there are four reaction options with correlation coefficients less than 0.4, such as "putting a lot of time and effort into community activities and taking leadership positions", so they need to be deleted. The meta-analysis of scholars such as McDaniel (2011) shows that the overall reliability of the situational judgment test is between 0.43 and 0.94. The overall reliability of this test is 0.73, which is consistent with the previous research conclusions, indicating that the test scale is consistent in concept and the measurement results are more reliable.

3.2 Hypothesis testing

In order to expand the scope of the sample, we selected 310 graduate students from Shandong, Henan, Hebei and other places for testing. In terms of gender, men accounted for 47.1% and women accounted for 52.9%. From the perspective of grade distribution, research one accounted for 38.7%, research two accounted for 34.2%, and research three accounted for 27.1%.

Postgraduate innovation situation judgment test. Using the self-developed postgraduate innovation situation judgment test scale, it contains 17 scenarios and 68 response options, each of which is evaluated by the Likert 6-point scale (1 means "very disagree" and 6 means "very agree" ").

Knowledge test. The scores of the graduate entrance examinations were selected as the scores of the graduate knowledge test. In order to analyze and compare with other variables, according to the previous research, the results of the postgraduate entrance examination results were counted and scored by the Likert 5 point method. Among them, 1 means ">300 points", 2 means "300~350 points", 3 means "350~370 points", 4 means "370~400 points", and 5 means ">500 points".

Innovation performance of postgraduate. The evaluation of others plus objective indicators, taking the total score of the two as the actual innovation performance score. In the evaluation section of others, the participants are asked to find a teacher or classmate to evaluate their own scientific research. we used the 11-item scale compiled by Scott and Bruce (1994). The sample items include "Does he/she often communicate with the tutor and promote his research ideas in research work?". The internal consistency of the scale is 0.83. The objective evaluation part requires the participants to fill out the level and number of the published academic papers and

conference papers. Referring to the previous study, each type of objective indicator is divided into different levels and assigned corresponding scores (Jin Lingzhi 2011).

Control variables. In addition, previous studies have shown that demographic variables such as gender, grade, and professional relevance may have an impact on graduate innovation performance (Guo Guimei & Duan Xingmin 2007). Therefore, we measured these variables simultaneously and analyzed them as control variables in the regression analysis.

Using statistical software such as Spss23.0 and Amos24.0, we performed descriptive statistics, correlation analysis and regression analysis on the sample data collected by the test to verify the predictive validity and incremental validity of the graduate innovation situational judgment test.

4. Results
4.1 Relevant analysis
The mean, standard deviation and correlation coefficient of Knowledge test scores, SJT scores, innovation performance and other control variables are shown in Table 2.

Table 2. Mean, standard deviation and correlation coefficient of each variable

Predictor	M	SD	1	2	3	4	5	6
1 Gender	1.64	0.48	1					
2 Grade	2.17	0.92	-0.02	1				
3 Professional relevance	3.50	1.10	0.48	-0.10	1			
4 Knowledge test	3.31	1.28	0.15**	0.05	0.17**	1		
5 SJT	1.74	0.49	0.06	0.10	0.13*	0.22**	1	
6 Innovation performance	3.48	0.59	-0.10	0.02	0.17**	0.20**	0.37**	1

Note.***$P<0.001$;** $P<0.01$.

It can be seen from the correlation analysis results that there is a significant positive correlation between the knowledge test scores ($r=0.20$, $p<0.01$) and SJT scores ($r=0.37$, $p<0.01$) and the actual innovation performance of graduate students. And the correlation between SJT score ($r=0.37$, $p<0.01$) and innovation performance is higher than knowledge test. Through the evaluation of declarative knowledge and procedural knowledge, knowledge test, postgraduate innovation situation judgment test can predict the innovation ability and innovation performance of graduate students to a certain extent, and the postgraduate innovation situation judgment test has stronger predictive validity. The developed postgraduate innovation scenario judgment test has good validity and can effectively predict the actual innovation performance of graduate students. Hypothesis 1 is verified.

4.2 Regression analysis
In order to further explore whether the developed postgraduate innovation scenario judgment test has incremental validity to the knowledge test, we conducted a multi-layer regression analysis on the predictor variables. The results of specific regression analysis are shown in Table 3. First, put the control variable gender, grade, and professional relevance into the first layer of regression analysis. Second, put the knowledge test score into the second layer of regression analysis. It can be seen that the knowledge test scores ($\beta=0.19$, $p<0.01$) have a significant positive impact on innovation performance. In the case of considering control variables, the knowledge test explained the innovation performance by 8%. Third, by placing the SJT score on the third level of the regression analysis, it can be seen that the SJT score ($\beta=0.33$, $p<0.001$) has a significant positive impact on innovation performance. Under the premise of controlling the variable of knowledge test, the SJT score increased the interpretation of innovation performance by 11%. It can be seen that the postgraduate innovation scenario judgment test is effective and unique in predicting the innovation performance of graduate students, and has incremental validity to the knowledge test method. Hypothesis 2 is verified.

www.iiste.org

Table 3. Regression analysis of knowledge test scores, SJT scores and innovation performance

Predictor	Innovation performance		
	first step	Second step	third step
Gender	-0.10	-0.13*	-0.14*
Grade	0.04	0.03	-0.01
Professional relevance	0.19**	0.15**	0.12*
Knowledge test		0.19**	0.13*
SJT scores			0.33***
R2	0.04	0.08	0.19
ΔR^2	0.04	0.04	0.11
F	4.37**	6.08***	12.84***
ΔF	4.37**	10.77**	36.85***

Note.***P<0. 001;** P<0. 01;* P<0. 05.

5. Discussion

In this study, we introduce a method of scenario judgment test to explore ways to improve the validity of graduate selection. The empirical results show that: (1) The situational judgment test developed based on the graduate selection scenario has good reliability ($\alpha=0.73$) and criterion-related validity ($r=0.37$, $p<0.01$), which can effectively predict the innovation performance of graduate students; (2) Situational judgment test has incremental validity to the existing selection method of postgraduate knowledge test, which can improve the interpretation of graduate innovation performance by about 11%.

Based on the above research conclusions, in order to improve the quality and innovation performance of graduate students, we can get the following inspirations when conducting graduate selection and training:

First, in the selection of graduate students, it is necessary to pay attention to the examination of the declarative knowledge, but also to the inspection of procedural knowledge. This will help to assess candidates' innovation potential more comprehensively and effectively. Declarative knowledge is the foundation of individual innovation. Only with good knowledge reserves and professional skills can we constantly discover and propose new problems and new ideas. Procedural knowledge is an important driving force for individual innovation, which enables individuals to make effective behavior choices in the process of innovation and promote the promotion and realization of innovative ideas. In the selection of graduate students, it is necessary not only to strengthen the testing of declarative knowledge, such as: basic knowledge, professional knowledge, etc., but also to strengthen the testing of procedural knowledge in scientific research practice and scientific research activities.

Second, universities should deepen the reform of the graduate enrollment mechanism. They should introduce more diversified assessment methods such as SJT, and improve the effectiveness of graduate selection by complementing the advantages of different assessment methods. The selection of graduate students is the starting point of postgraduate education and the key to determining the quality and training level of graduate students. At present, in the selection of graduate students in knowledge testing and structured interviews in China, the following drawbacks exist: for example, the knowledge test is too broad and lacks pertinence, and the subjective interference in the interview is too strong and difficult to quantify. As a simulation test, SJT can measure candidates' procedural knowledge needed for innovative work and evaluate candidates' future innovation potential through behavioral judgment and behavioral choices of key situations in graduate innovation activities. Therefore, the introduction of SJT in the selection of graduate students can make up for the shortcomings of the existing graduate selection methods, and it is an effective supplement to the existing selection methods for graduate students.

References

Chen Qian. (2016), "Discussion on the Path of Optimizing the Selection of Doctoral Candidates", *Jiangsu Higher Education* **6**, 39-42.

Christian, M.S., Edwards, B.D., & Bradley, J.C. (2010), "Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities". *Personnel Psychology* **63**(1), 83-117.

Guo Guimei, Duan Xingmin. (2007), "A Review of the Research on the Influence of Individual Characteristics on Employees' Creativity". *Science & Technology Progress and Policy* **10**, 206-210

Jin Lingzhi, He Wei. (2011), "Recognition of Innovation Performance Influencing Factors of Doctoral Students in Science and Engineering". *Research in Higher Education of Engineering* **3**, 133-139.

Lievens, F., & Coetsier, P. (2002), "Situational tests in student selection: An examination of predictive validity, adverse impact,and construct validity". *International Journal of Selection and Assessment* **10**(4), 245-257.

Lievens, F., & Patterson, F. (2011), " The validity and incremental validity of knowledge tests, low-fidelity

simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection". *Journal of Applied Psychology* **96**(5), 927-938.

Lievens, F. (2013), "Adjusting medical school admission: assessing interpersonal skills using situational judgement tests". *Medical Education* **47**(2), 182-189.

Liu Dan, Wang Fei, Wang Zongxia. (2015), "Analysis of Influencing Factors of Graduates′ Innovation Performance and Promotion Methods". *Scientific Management Research* **33**(4), 102-105.

Lu Xiefeng, Shu Wenhui. (2017), "Development of Situational Judgment Test for Postgraduate Competency". *Educational Measurement and Evaluation* **12**, 43-49.

McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., *et al*. (2001), "Use of situational judgment tests to predict job performance: A clarification of the literature". *Journal of Applied Psychology* **86**(4), 730-740..

McDaniel M A, Psotka J, Legree P J, *et al*. (2011), "Toward an understanding of situational judgment item validity and group differences". *Journal of Applied Psychology* **96**(2), 327-335.

Motowidlo, S.J., Dunnette, M.D., & Carter, G.W. (1990), " An alternative selection procedure: The low-fidelity simulation". *Journal of Applied Psychology* **75**(6), 640-647.

Motowidlo, S.J., Hooper, A.C., & Jackson, H.L. (2006), "Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items". *Journal of Applied Psychology* **91**(4), 749−761.

O'Connell, M.S., Hartman, N.S., McDaniel, M.A., *et al*. (2007), "Incremental validity of situational judgment tests for task and contextual job performance". *International Journal of Selection and Assessment* **15**(1), 19-29.

Schmitt, N., & Ostroff, C. (1986), "Operationalizing the behavioral consistency approach : selection test development based on a content-oriented strategy". *Personnel Psychology* **39**(1), 91-108.

Scott, Bruce R A. (1994), "Determinants of innovative behavior: A path model of individual innovation in the workplace". *Academy of Management Journal* **37**(3), 580— 607.

Schmidt, F.L., &Hunter, J.E. (1998), "The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings". *Psychological Bulletin* **124**(2), 262-274.

Song, Y. (2015), "Development and Validation of a Situational Judgment Test for a Competencey of Public Ethics: A Case of an University in South Korea". *Journal of the Korea Academia-Industrial cooperation Society* **16**(7), 4478-4488.

Wang Chuanyi, Cheng Zhe. (2017), "Measuring Methods on Non-cognitive Abilities in Postgraduate Entrance Examination: Concept, Practice and Prospect". *Journal of Graduate Education* **5**, 67-72, 78.

Wang Ying, Liang Wei, Lei Xinghui. (2014), "Research on the Context Factors Influencing the Knowledge Innovation Performance". *Science and Technology Management Research* **34**(14), 17-22, 30.

Zhu Hong, Li Wenli, Zu Zuzu. (2011), "The Status Quo and Influence Mechanism of Graduate Students' Innovation Ability in China". *Higher Education Research* **32**(2), 74-82.