Development of a Game-Based Assessment to Measure Creativity
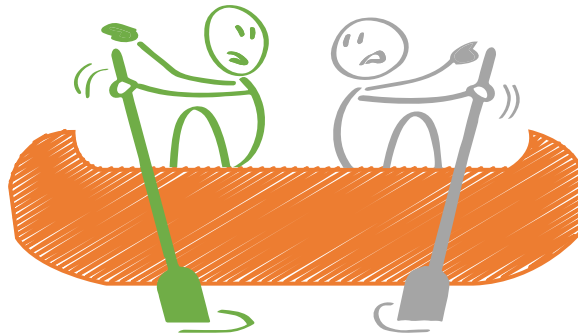
Fang Luo
Faculty of Psychology Beijing Normal University

Assessment based on virtual dynamic situation

Game-Based Assessment

High-fidelity test situation

Stimulate individual initiative

Recorded process data

High test concealment

High test validity

Traditional Assessment

Low-fidelity test situation

Highly structured

Examine superficial knowledge

Boring test format

Virtual dynamic game environment provides almost real task background, students can apply knowledges and a variety of abilities to solve the game tasks (Dede 2005; DiCerbo and Behrens 2012; Quellmalz et al. 2012)

Game-Based Assessment : Core Elements

Interactive Problem Solving

In the dynamic interaction with the system, individuals learn rules and solve problems.

1

Game Rules

Individuals achieve mission goals within the rules set by the game.

2

Adaptive Challenges

The difficulty of the game task is reasonable, matching with individuals' ability level.

3

Real-Time Feedback

Provide appropriate feedback on the individual's performance in the game, which will promote the subject's performance.

4

Rich Sensory Stimulation

The game has rich stimulating elements to attract individuals to solve the tasks in the game

5

Shute, V. J., & Ke, F. (2012). Games, Learning, and Assessment. Assessment in Game-Based Learning. Springer New York.

1 The Definition of Creativity

The concept of creativity is generally thought to encompass two dimensions：

➢ Novelty： The novelty, uniqueness and originality of creative ideas and products

➢ Usefulness： New things are useful, suitable and appropriate to problem or the environment.

The Complexity and Multidimensionality of Creativity

Torrance(1966) defines creativity as a series of cognitive skills：

➢ Be sensitive to problems, vulnerabilities, knowledge gaps, and inconsistent content；

➢ Discover challenges；

➢ Search for solutions, make guesses, and develop assumptions about deficiencies；

➢ Inspection, correction and testing；

➢ Communicate results.

2 Creative Problem Solving

Creative problem solving combines the creative thinking skills and problem solving：

➢ The process of finding creative solutions to problems in a novel, flexible, and original way

➢ It emphasizes idea generation in the problem solving process and the role of divergent and convergent thinking.(Liang, Proctor, & Salvendy, 2011; Newell, Shaw, & Simon, 1962)
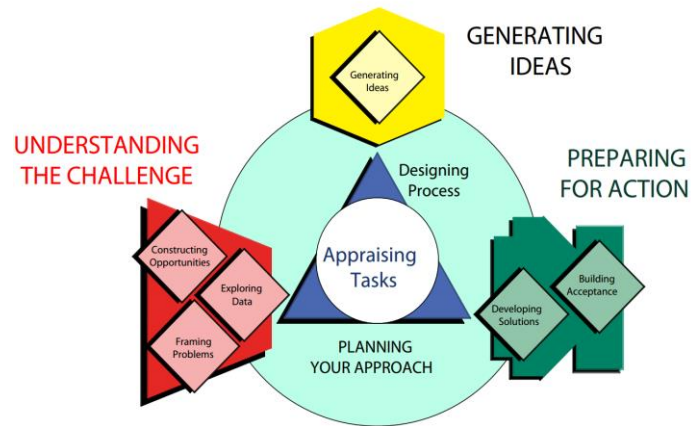
The Creative Problem Solving Framework
(CPS Version 6.1™)
© 2010, Center for Creative Learning, Inc. and Creative Problem Solving Group, Inc.

(Treffinger & Isaksen, 2010)

3 The Measurement of Creativity

➢ Torrance and Goff(1989) counted no less than 255 creativity tests in the literature

➢ However, most measurements have problems such as low psychometric qualities, unclear construct, and failure to update the norm in time (Treffinger et al., 2002), which lags behind other areas.

Divergent Thinking Test

Attitude and Interest batteries

Study of Outstanding Figures

Personality Inventories

Biographical inventories

Judgements of productions

Peer and teacher nomination

Self-reported Creative Activities and Achievements

(Hocevar & Bachelor, 1989)

Divergent Thinking Test

Divergent Thinking & Convergent Thinking

Divergent Thinking：Generate a large amount of responses, following different perspectives, in original ways of thinking.

➢ Fluency: fluency in thinking

➢ Flexibility: the ability to switch and bound in thinking

➢ Novelty: the ability to come up with new, unique ideas

Convergent Thinking：By thinking systematically about a particular problem, cognitive processes point to a single correct solution.

Insight Problem

Insight Process

Insight：In the face of a problem impasse, more effective new ideas burst into the brain at the moment when the individual can view the problem from a new perspective, accompanied by instant inspiration and thinking leaps (Mednick,1962).

The Cognitive Mechanism of Insight:

➢ Representation Transformation Theory (Kaplan & Simon, 1990)

➢ Process Monitoring Theory (MacGregor, Ormerod, & Chronicle, 2001)

➢ Prototype heuristic Theory (Zhang, 2004)

Open Questions

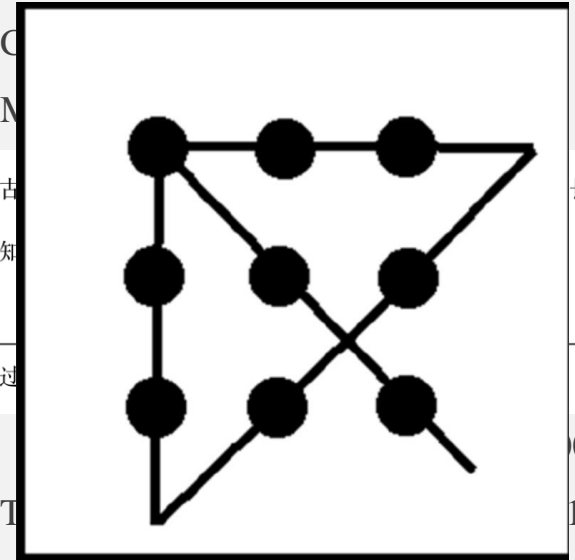BALANCING ACT

NEW FRIEND

(b).

arge?
home
tual

Insight Test

Ill-structured Problems

Laboratory Tasks
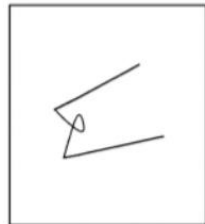
➤ C

➤ M

3. 有人拿一枚古   长

想都不想就知

答：

是否之前看过

➤ T

2)

08)

11)

Divergent Thinking Test

Limitations

- **Ignore the "Usefulness" dimension.** Many bizarre and impractical responses are rated as highly creative;
- **The task is too simple and abstract,** detached from the problem situation;
- It isn't focused on the whole creative problem solving process. **Too much emphasis is placed on divergent thinking** at the expense of other creative thinking abilities.
- **Manual scoring is difficult and subjective,** and needs to compare the norm.

Insight problems

Limitations

- The insight items are demanding， and **hard to develop**；
- **Insight items can not be reused**, easy to expose; If a person knows the answer in advance, it can only be treated as missing value;
- **There is only one index of the insight test**, which can only be evaluated for right and false answers, and it is difficult to evaluate the cognitive process.

Research objective: to develop a game-based assessment to measure creativity

1. Considering usefulness in the creativity measurement

➢ Design real and meaningful problem situations, and set condition restrictions in the process of solving problems, so as to avoid the weird answer and be difficult to evaluate.

➢ Pre-design creative problem-solving steps and answer sets in advance.

2. Measuring creative problem solving processes

➢ ill-defined problems: the information of the initial state of the problem is incomplete, which induces the creative problem solving process and examines the multiple thinking abilities of creativity.

➢ Divergent thinking: think about different categories of solutions

➢ Convergent thinking: the process of exploring and executing each specific solution

Divergent Thinking Test & Insight problems

➢ Ill-defined open questions

➢ A closed set of answers

3. Problem situations can trigger creative thinking

➢ Simulate complex problem scenarios

➢ Response mode of human-computer interaction

➢ Strong interest help improve the subjects' motivation and involvement in answering

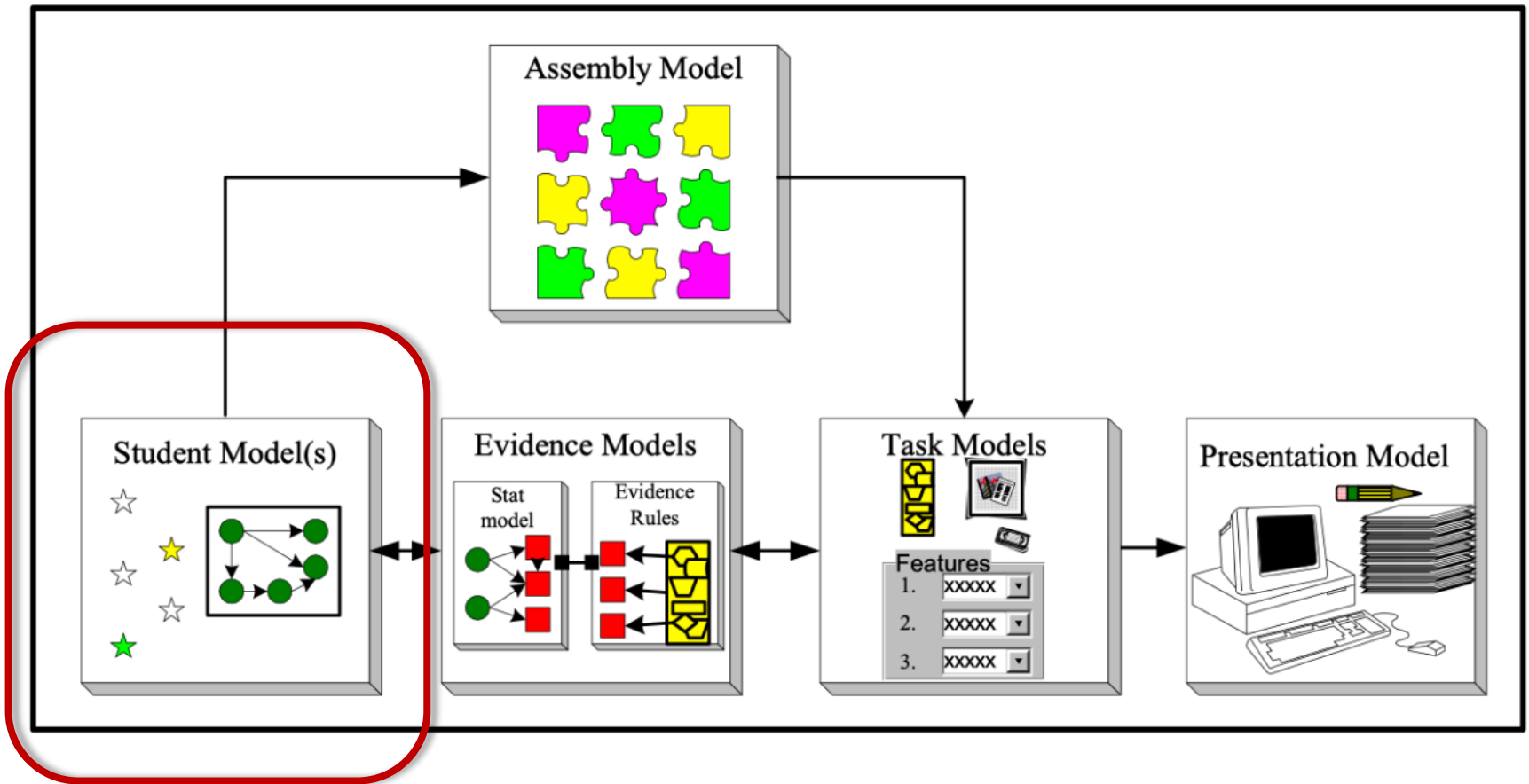## 4. Extracting process indicators and scoring

➢ Establish well-informed log files for process data

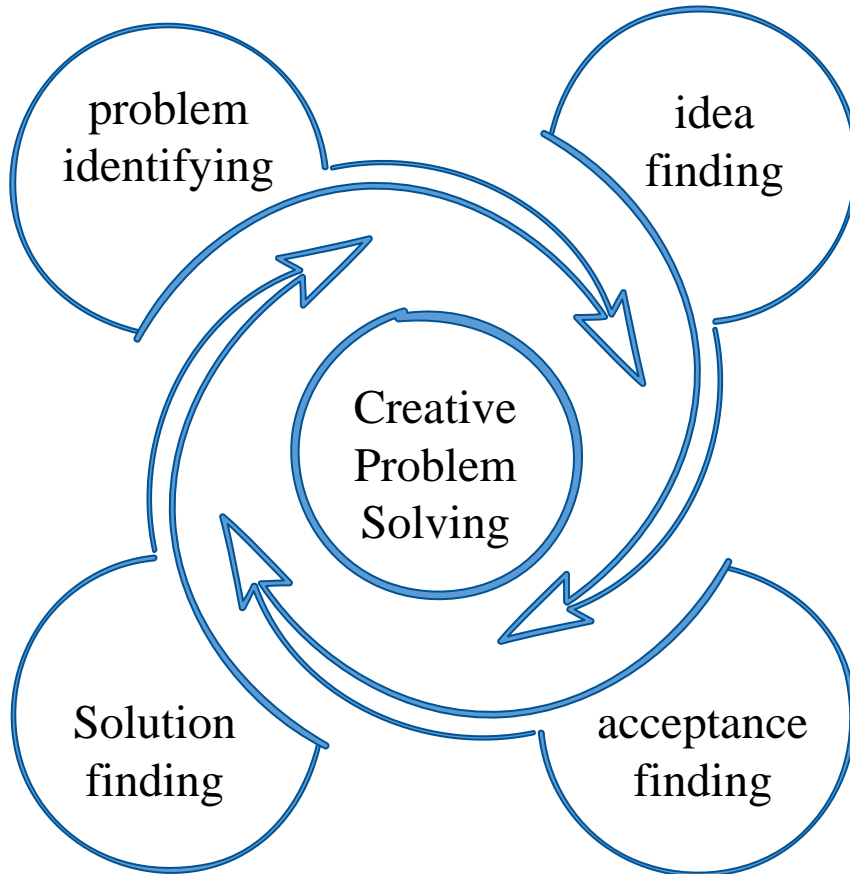➢ Extract outcome and process measurement indicators

Game-based Assessment

Evidence-Centered Test Design

(Mislevy, Sternberg, & Almond, 2003)

Competence Model

problem identifying

idea finding

Solution finding

acceptance finding

Creative Problem Solving

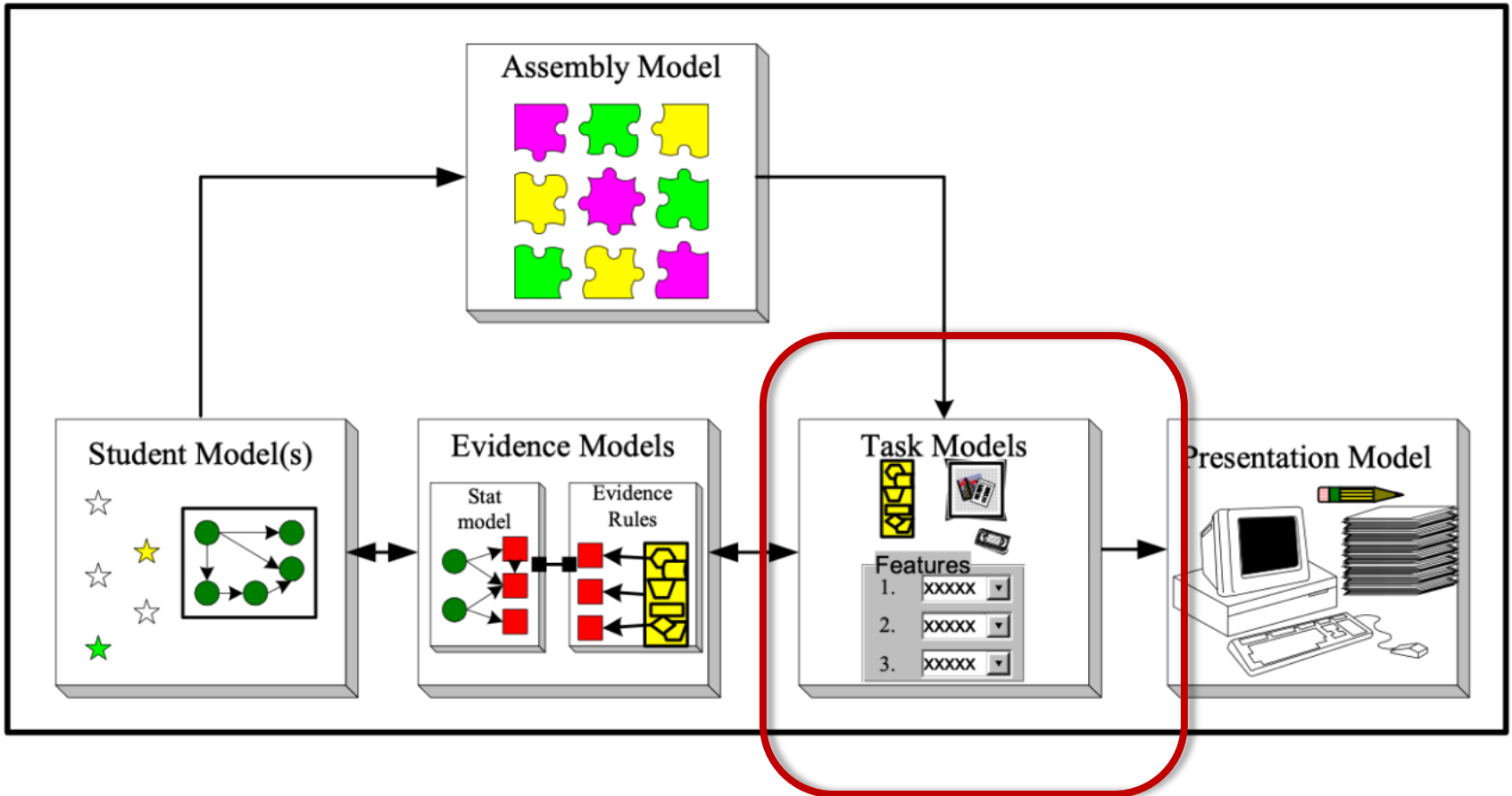**Creative Problem Solving**

➢ Based on the theoretical model of creative problem solving CPS 3.0™, and combined with the test tasks, the creative thinking process is divided into four cognitive stages with dynamic cycles.

➢ Each sub-process involves both divergent and convergent thinking processes.

(Treffinger & Isaksen, 2004)

Evidence-Centered Test Design

(Mislevy, Sternberg, & Almond, 2003)

Task Model

Creativity game test design

Problem situation: primitive life

➢ Level 1: Hunting (10 min)

➢ Level 2: Crossing a river (15 min)

➢ Level 3: Picking bananas (20 min)

Mission goals：

➢ Find solutions according to the mission goals

➢ The more solutions you find, the better you are

➢ The more novel the method, the better it is

Game Rules

Card Combination: Three Types of Feedback

Make a Combination Card

The cards are combined to produce a new combined card that incorporates the card elements.

Animation Feedback

No cards are synthesized, but clues to creative solutions are provided through animation.

Unable to Combine

If an individual synthesized a combination outside the preset, the system would give a prompt "unable to combine."

High Creativity Path (Level 3)
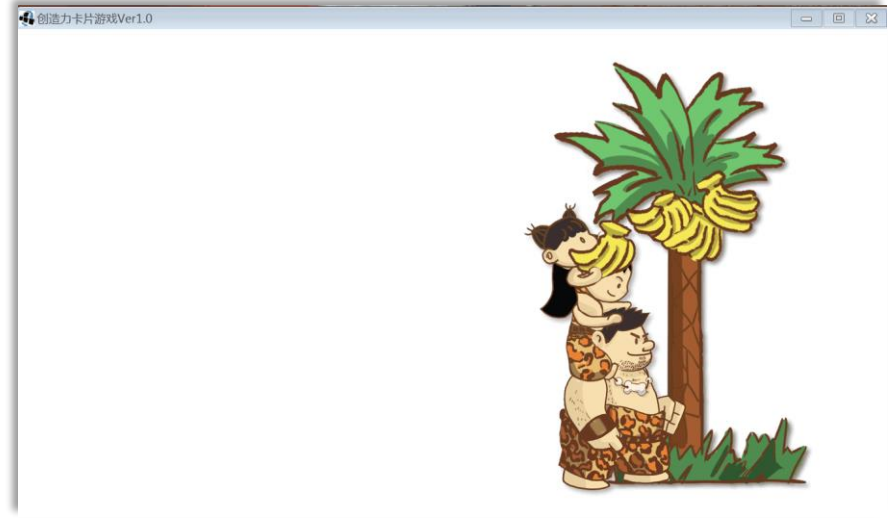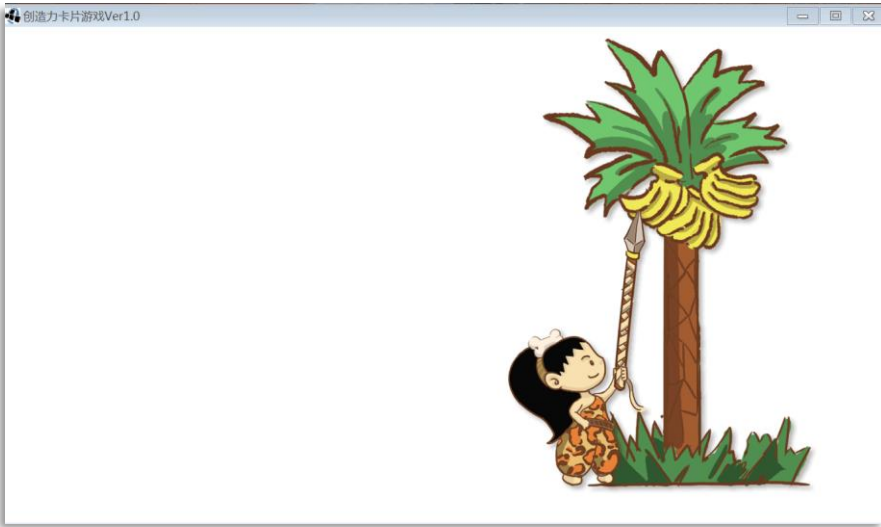
Test Development Flow
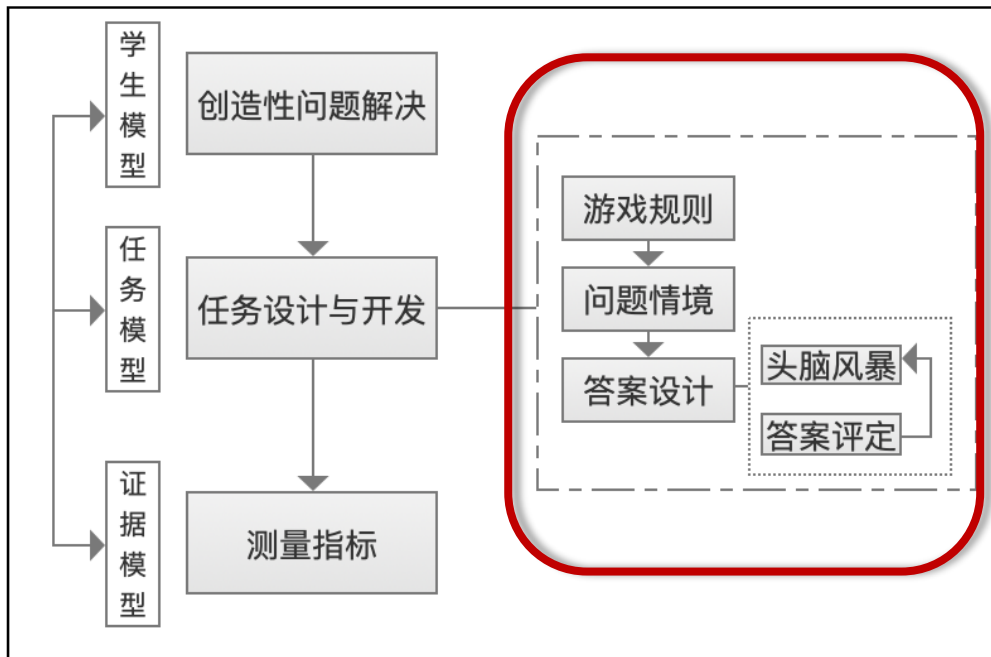
The Design of The Answer Path

Step 1 Brainstorm (6 to 7 people)

➢ Delay criticism

➢ Free association

➢ Encourage more ideas

➢ Integration and Improvement

   (Osborn, 1963)

Step 2 Path Evaluation and Selection

➢ Novelty

➢ Suitability/reasonableness

➢ A sense of humor

Evidence-Centered Test Design
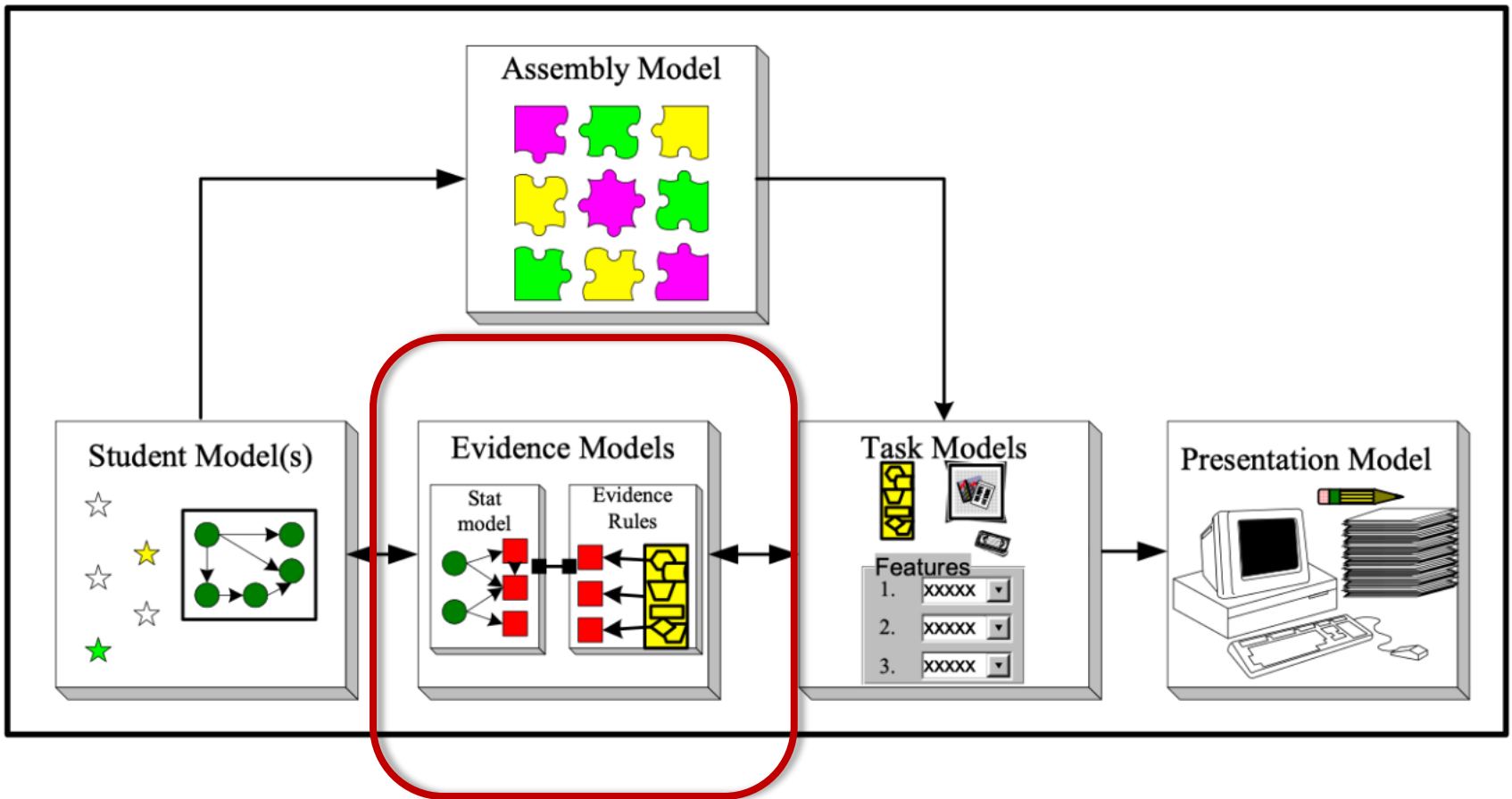
(Mislevy, Sternberg, & Almond, 2003)

Evidence Model

Time	ID	Event No.	Action	Object	Start Point	Target Point
			Drag	Q	0	2

Reaction Zone1	Reaction Zone2	Reaction Zone3	Event label
A	Q	0	Move the card successfully

Evidence Model

Total number of paths

Within each level, the total number of highly creative paths prespecified by the researcher was found.

solutions

Total number of small schemes

Within each level, the total number of combined cards

paths

Evidence Model

Ideas

Total number of ideas

The total number of combinations of cards, that is, the amount of ideas formed by a subject

Ideas

Clues

Total number of clues

The total number of clue animations watched when combining cards, without synthesizing combination cards.

Clues

Unexpected Ideas

Total number of ideas beyond the preset

The total number of schemes beyond the preset that a subject give

Unexpected Ideas

▶ The Research Methods

Subject

A total of 515 first-year university student in Shanxi Province were included in the study. There were 255 males and 241 females, with an average age of 18.67±0.88 years.

Measurement Tools

- Background Questionnaire
- Divergent thinking tests (Verbal)
- Divergent thinking tests (Figure)
- Insight Problem
- Remote Association Test
- Raven's Progressive Matrices
- Williams Creativity Scale

Research Results

▶ Study One / **Psychometric Properties**

/ Difficulty, Discrimination, Reliability

/ Criterion Validity

▶ Study Two / **Validity evidence Based On Bayesian Network**

/ Insight Ability As a Predictor of The Path Total Score

/ Conditional Probability Analysis of Bayesian Networks

Test Difficulty
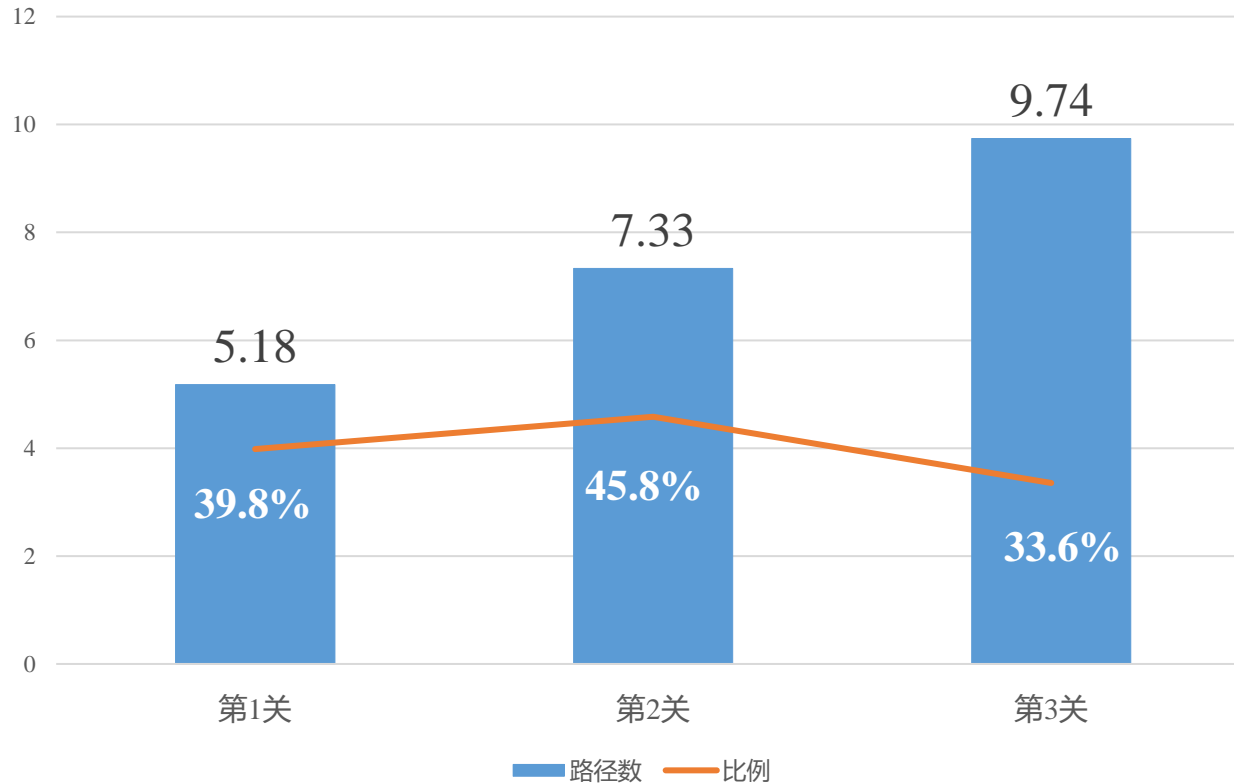
Note: There are 13, 16 and 29 kinds of paths in the three levels respectively.

Level One: Difficulty And Discrimination

表 4-2 第 1 关路径的难度和区分度（$N = 515$）

编号	卡片数	步骤数	通过率	标准差	区分度
路径 1	3	2	0.74	0.44	0.18***
路径 2	3	2	0.15	0.36	0.40***
路径 3	4	3	0.16	0.36	0.36***
路径 4	4	4	0.22	0.41	0.25***
路径 5	4	2	0.33	0.47	0.31***
路径 6	3	2	0.97	0.18	0.12**
路径 7	4	3	0.15	0.36	0.39***
路径 8	5	3	0.25	0.43	0.27***
路径 9	3	2	0.56	0.5	0.23***
路径 10	3	2	0.96	0.2	0.08+
路径 11	3	2	0.48	0.5	0.28***
路径 12	3	2	0.17	0.37	0.30***
路径 13	4	3	0.05	0.22	0.30***

注：*$p<.05$, **$p<.01$, ***$p<.001$

Level Two: Difficulty And Discrimination

表 4-3 第 2 关路径的难度和区分度（$N = 515$）

编号	卡片数	步骤数	通过率	标准差	区分度
路径 1	4	2	0.07	0.26	-0.02
路径 2	5	5	0.20	0.40	0.25***
路径 3	4	4	0.77	0.42	0.20***
路径 4	5	5	0.02	0.14	0.01
路径 5	4	2	0.40	0.49	0.38***
路径 6	3	3	0.31	0.46	0.29***
路径 7	4	2	0.22	0.41	0.26***
路径 8	4	3	0.46	0.50	0.42***
路径 9	3	2	0.97	0.16	0.19***
路径 10	3	2	0.78	0.41	0.15***
路径 11	4	3	0.39	0.49	0.39***
路径 12	4	3	0.87	0.34	0.29***
路径 13	3	2	0.19	0.4	0.26***
路径 14	3	2	0.72	0.45	0.39***
路径 15	3	3	0.35	0.48	0.34***
路径 16	5	5	0.60	0.49	0.37***

注：*$p<.05$, **$p<.01$, ***$p<.001$

Level Two: Difficulty And Discrimination

表 4-4 第 3 关路径的难度和区分度（$N = 515$）

编号	卡片数	步骤数	通过率	标准差	区分度
路径 1	4	3	0.07	0.26	0.12***
路径 2	4	3	0.91	0.28	0.22***
路径 3	3	2	0.35	0.48	0.24***
路径 4	4	3	0.54	0.50	0.31***
路径 5	4	4	0.22	0.42	0.18***
路径 6	4	3	0.62	0.49	0.27***
路径 7	3	2	0.71	0.46	0.35***
路径 8	4	3	0.65	0.48	0.34***
路径 9	4	2	0.76	0.43	0.31***
路径 10	4	2	0.67	0.47	0.34***
路径 11	3	1	0.06	0.24	0.13**
路径 12	4	3	0.53	0.50	0.31***
路径 13	4	2	0.17	0.38	0.32***
路径 14	4	3	0.61	0.49	0.27***
路径 15	5	4	0.50	0.50	0.37***
路径 16	3	2	0.93	0.26	0.16***
路径 17	4	4	0.09	0.28	0.18***
路径 18	4	2	0.02	0.13	0.03
路径 19	5	3	0.04	0.18	0.09*
路径 20	4	3	0.57	0.50	0.29***
路径 21	5	4	0.11	0.31	0.20***
路径 22	3	1	0.08	0.27	0.00
路径 23	5	3	0.28	0.45	0.32***
路径 24	4	3	0.05	0.21	0.00
路径 25	5	4	0.12	0.33	0.05
路径 26	6	5	0.03	0.17	0.02
路径 27	6	5	0.01	0.12	0.05
路径 28	5	4	0.02	0.15	0.06
路径 29	6	6	0.04	0.19	0.02

注：*$p<.05$, **$p<.01$, ***$p<.001$

Criterion Validity: Divergent Thinking

表 4-7　创造力游戏测验测量指标总分的效标关联效度（$N = 507$）

创造力游戏测验	远距离联想测验	顿悟性问题测验	新编创造思考图形测验			新编创造思考语文测验		
			流畅力	变通力	独创力	流畅力	变通力	独创力
路径总分	0.14**	0.30***	0.09*	0.16***	0.14**	0.06	0.10*	0.06
观点总分	0.06	0.21***	0.12**	0.29***	0.16***	0.06	0.07+	0.08+
小方案总分	-0.01	0.21***	0.09*	0.19***	0.09*	0.02	0.02	0.03
线索总分	-0.01	0.13**	0.07	0.14**	0.09+	-0.01	0.02	0
设定外的观点总分	0.05	0.05	0.05	0.23***	0.12**	0.03	0.09+	0.04

注：*$p<0.05$, **$p<.01$, ***$p<.001$. 指标总分的合成方式为每关对应指标 z 分数之和

➤ Compared with the language test, the game test has more significant correlation with the graph test.

➤ The correlation between the total score of viewpoint and each dimension of the graph test is the highest among all tests, especially the flexibility dimension.

➤ On the whole, the correlations between the measurement indexes and the flexibility dimension are higher than that of the fluency and originality dimensions.

Criterion Validity: Insight Ability

表 4-7 创造力游戏测验测量指标总分的效标关联效度（$N = 507$）

创造力游戏测验	远距离联想测验	顿悟性问题测验	新编创造思考图形测验			新编创造思考语文测验		
			流畅力	变通力	独创力	流畅力	变通力	独创力
路径总分	0.14^{**}	0.30^{***}	0.09^{*}	0.16^{***}	0.14^{**}	0.06	0.10^{*}	0.06
观点总分	0.06	0.21^{***}	0.12^{**}	0.29^{***}	0.16^{***}	0.06	0.07^{+}	0.08^{+}
小方案总分	-0.01	0.21^{***}	0.09^{*}	0.19^{***}	0.09^{*}	0.02	0.02	0.03
线索总分	-0.01	0.13^{**}	0.07	0.14^{**}	0.09^{+}	-0.01	0.02	0
设定外的观点总分	0.05	0.05	0.05	0.23^{***}	0.12^{**}	0.03	0.09^{+}	0.04

注：$^{*}p<0.05$, $^{**}p<.01$, $^{***}p<.001$. 指标总分的合成方式为每关对应指标 z 分数之和

Criterion Validity: Convergent Thinking

表 4-9 创造力游戏测验每关测量指标的效标关联效度（$N = 507$）

指标	瑞文	威廉斯测验总分	冒险性	好奇性	想象力	挑战性
路径总分	0.15***	0.17***	0.15***	0.12**	0.16***	0.12*
小方案总分	0.06	0.08+	0.11*	0.01	0.09*	0.05
观点总分	0.08+	0.15**	0.10*	0.10*	0.15***	0.13**
线索总分	0.06	0.09*	0.12**	0.03	0.09*	0.06
设定外的观点总分	0.04	0.10*	0.01	0.11*	0.12**	0.09*

注：*$p<0.05$, **$p<.01$, ***$p<.001$.

Inter-criterion Correlation Results

	1	2	3	4	5	6	7	8	9	10
远距离联想测验	1									
顿悟性问题测验	0.185***	1								
创造思考图形测验										
流畅力	0.031	0.084*	1							
变通力	0.006	0.110**	0.611***	1						
独创力	-0.016	0.067	0.833***	0.514***						
创造思考语文测验										
流畅力	0.001	0.083*	0.476***	0.262***	0.391***	1				
变通力	-0.006	0.062	0.390***	0.274***	0.342***	0.851***	1			
独创力	-0.027	0.064	0.455***	0.217***	0.405***	0.877***	0.695***	1		
瑞文	0.221	0.259	0.173***	0.148***	0.138***	0.070+	0.041	0.071+	1	
威廉斯测验	0.042	0.142***	0.198***	0.184***	0.205***	0.123**	0.129**	0.126**	0.017	1

注：*$p<0.05$, **$p<.01$, ***$p<.001$.

Study Two

Insight prediction of the total number of paths

➢ Hierarchical regression was established to explore the independent effect of the insight ability on the total number of game paths

➢ Results: After controlling for demographic variables (gender, age) and differences in gaming experience, as well as divergent and convergent thinking abilities, the insight ability was still a significant predictor of the total number of pathways $(\beta = 0.242,\ p < 0.000, \Delta R^2 = 0.053)$

The Construction of Bayesian Networks

➢ To verify the construct validity of creativity game test and explore the effect of insight levels on key response steps.

➢ The Bayesian network algorithm was used to model and analyze the individual response process, taking the intermediate steps (combined cards) of each path as the network nodes, and the combining relationship between cards as the network adjacent edges.
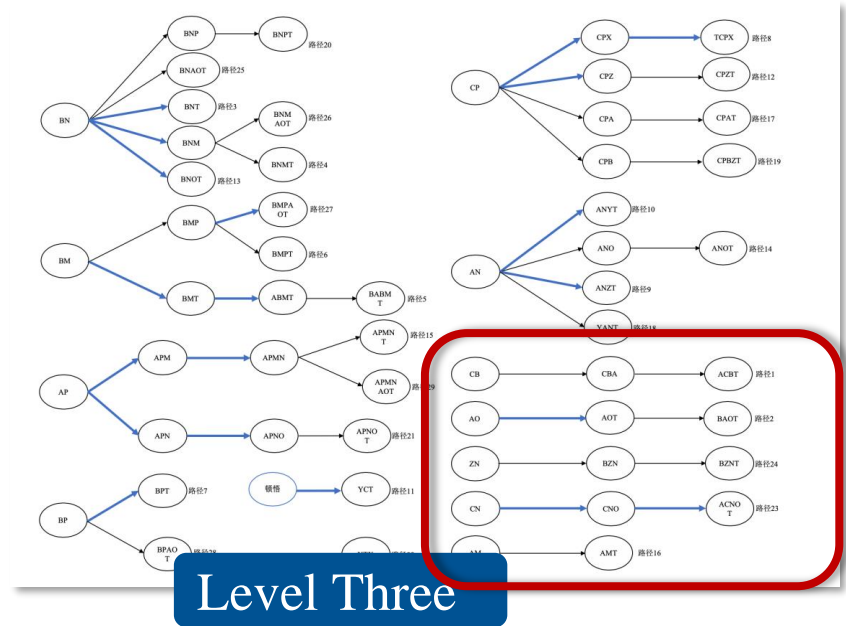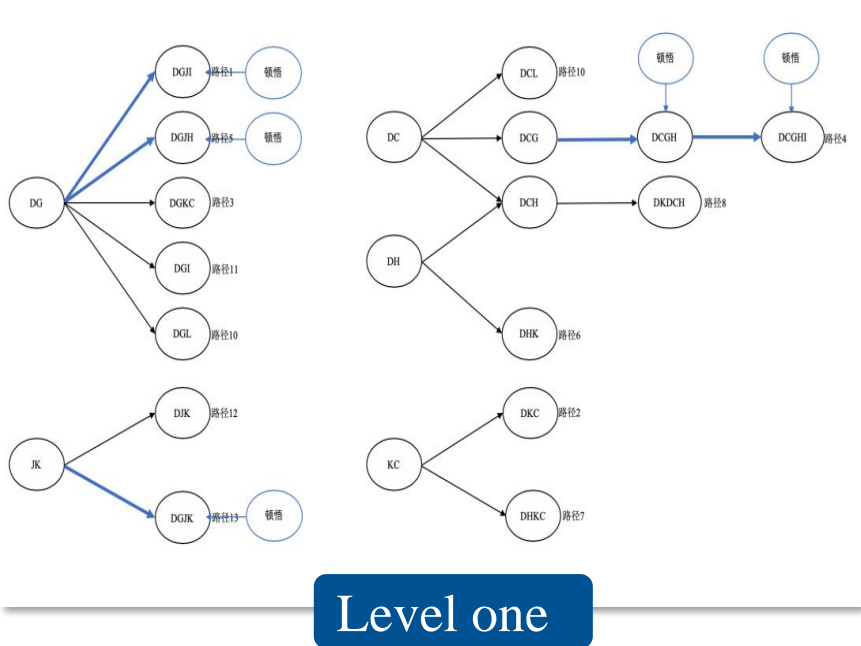
Bayesian Network Diagram

Group of The Insight Ability

表 5-2 等宽度离散顿悟分组（$N = 489$）

顿悟分组	人数	占比
高顿悟水平组	67	13.70%
中顿悟水平组	317	64.80%
低顿悟水平组	105	21.50%

➢ Apply equal-width discrete method

➢ According to the score of the insight question test, the subjects were divided into low, middle and high three groups.

Bayesian Network of level 1 and 2

Level one

Level Three

Group I Combination Cards：
Node Edge Probability Table

Level One： Hunting

表 5-3 I 类组合卡节点边缘概率表（第 1 关）

节点	DG		JK		KC		DH		DC	
元素名称	爸爸拿着火把		山羊猛犸象相撞		山羊吃白菜		爸爸拿藤蔓		爸爸拿白菜	
取值	0	1	0	1	0	1	0	1	0	1
P	0.002	0.998	0.790	0.210	0.646	0.354	0.020	0.980	0.010	0.990

Level Two： Cross a river

表 5-5 I 类组合卡节点边缘概率表（第 2 关）

节点	AJ		AP		FR		AG	
名称	姐姐拿着斧头		姐姐折断柳条		河马游入河中		姐姐抓着蛇	
取值	0	1	0	1	0	1	0	1
P	0.002	0.998	0.190	0.810	0.646	0.354	0.110	0.890

节点	AR		AO		AK	
名称	姐姐拿水草		姐姐拿石块		姐姐拿着匕首	
取值	0	1	0	1	0	1
P	0.634	0.366	0.014	0.986	0.002	0.998

Effect of the insight level on conditional probability of group II combination cards

Level One：Hunting

表 5-4 II 类以上组合卡条件概率分布表（第 1 关）

节点		DGJI		DGJH		DGJK		DCGH	
名称		爸爸追赶大象 进水塘		藤蔓引火烧大象		爸爸拿火 捉晕倒的羊		爸爸用藤蔓 拴住虫子	
取值		0	1	0	1	0	1	0	1
顿悟	低	0.403	0.597	0.746	0.254	0.769	0.231	0.250	0.750
	中	0.252	0.748	0.681	0.317	0.735	0.265	0.235	0.765
	高	0.173	0.827	0.587	0.413	0.727	0.273	0.108	0.892

节点		DCG		DKDCH		DCGHI	
名称		爸爸烤出虫子		藤蔓白菜陷阱 捉山羊		爸爸用虫子钓鱼	
取值		0	1	0	1	0	1
顿悟	低	0.758	0.242	0.841	0.159	0.167	0.833
	中	0.674	0.326	0.640	0.360	0.103	0.897
	高	0.648	0.352	0.481	0.519	0.090	0.909

	DG	
	爸爸拿着火把	
0		1
0.002		0.998

爸爸拿着白菜
D:爸爸
C:白菜

DC

顿悟

爸爸拿白菜在
火上烤出虫子
G:火堆

DCG

		0	1
顿悟	低	0.758	0.242
	中	0.674	0.326
	高	0.648	0.352

顿悟

爸爸把虫子拴
在藤蔓上
H:藤蔓

DCGH

		0	1
顿悟	低	0.250	0.750
	中	0.235	0.765
	高	0.108	0.892

顿悟

爸爸用合成的
工具钓鱼
I:水塘

DCGHI

		0	1
顿悟	低	0.167	0.833
	中	0.103	0.897
	高	0.09	0.909

图 5-5 第一关路径 4 的贝叶斯网络结构图

Discussion: The Effect of Insight on The Response Steps

The Cognitive Mechanism of Insight

Path 4: Fishing (Difficulty 0.22)

Represent the task objectives and problems

➤ Shallow representation: animals in the inventory

➤ Deep representation: fish in a pond

Overcome the influence of **thinking set**

➤ Vines: fishing line

➤ The role of cabbage

Appropriate **prototype activation** and key information inspiration

Bayesian Knowledge Tracking

Knowledge Tracing (KT) represented as a static Bayesian Network

Latent

Observed

Node representations
K = Knowledge node
Q = Question node

Node states
K = Two state (0 or 1)
Q = Two state (0 or 1)

- Knowledge tracking: use the answer sequence (behavior sequence) to train the model parameters and update the knowledge mastery level (ability level).

- Model parameters: estimate students' initial state of knowledge mastery, learning parameters, and final state of knowledge mastery. In the game test, the initial ability of the student, the ability development between steps, and the final ability are estimated.

- In this study, the knowledge tracking model is used to find that: students are easy to master the card combination method, easy to acquire, and the first step of synthesis is easy. However, in the process of finding a solution, the ability to synthesize a complete creative solution is weak and harder to learn.
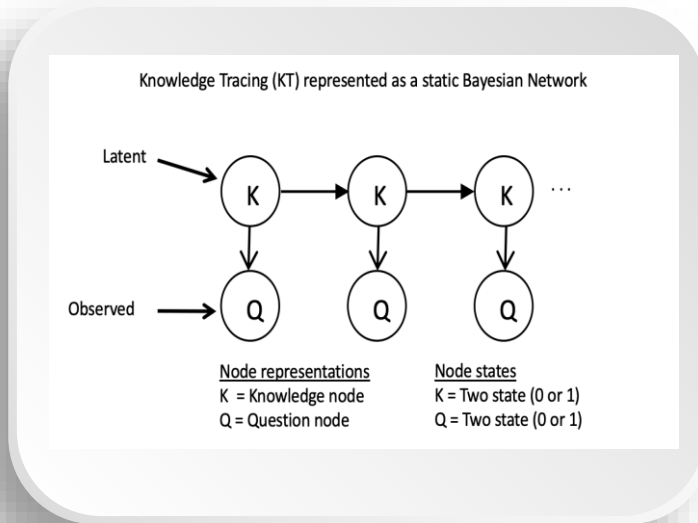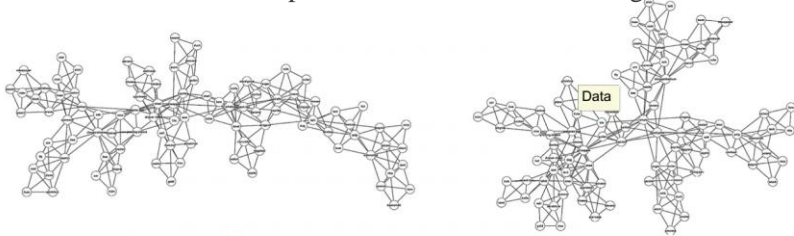
Try to Use Network Analysis Method to Model Process Data in The Game
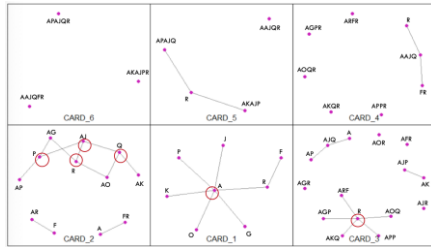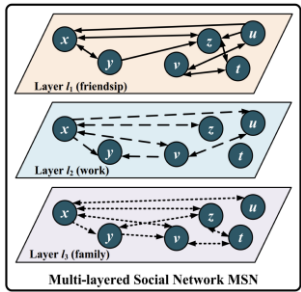
The creative problem solving network based on the process data from the creative game

Low creative

High creative

Semantic memory networks of high and low creatives based on free association tests

Complex network analysis of creativity

In the creativity test, the complex network method was used to construct the creative problem solving process network of the responders based on the key steps in the process data.
Through this method, not only can the individual creative thinking process be visualized, but also can explore the network characteristics of high and low creativity.
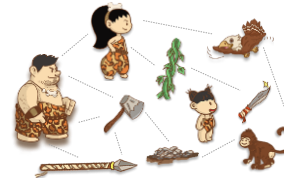
1 Network topology trait analysis

Network distance, centrality, community and other important network statistical indicators were used to describe the specific characteristics of the creativity sub-skills of the respondents.

2 Automated creativity assessment

Combining network characteristics into the measurement framework of creativity, exploring the path and method of automatic evaluation of creativity and its various dimensions.

THANK YOU FOR YOUR LISTENING