

# 青少年科学创造力测验的编制\*

申继亮<sup>1</sup> 胡卫平<sup>2</sup> 林崇德<sup>1</sup>

(1. 北京师范大学发展心理研究所 100875; 2. 山西师范大学课程与教学研究所, 临汾 041004)

**摘要:**基于科学创造力的结构模型、青少年科学创造力的表现及托兰斯的创造性测验,编制了青少年科学创造力测验,并用该测验施测于英国的 1087 名中学生和中国的 1087 名中学生,结果表明:(1)《青少年科学创造力测验》具有较高的信度,Cronbach 系数、评分者信度、重测信度均达到心理测验学要求的水平;(2)《青少年科学创造力测验》具有较高的结构效度。

**关键词:**青少年; 科学创造力; 测验

## 1 问题提出

由于研究者对创造力的定义不同,理论依据不同,所采用的测量方法也不相同,有关创造力的测量量表多达百余种。既然有如此多的创造力测验,为什么还要编制青少年科学创造力测验呢?首先,科学研究不仅需要科学知识,而且需要创造力,这一点已经被科学发展的历史所证实。如著名美籍华裔物理学家朱棣文<sup>[1]</sup>认为,创造精神是最重要的,创造精神强而天资差一点的学生往往比天资强而创造精神不足的学生能取得更大的成就;其次,研究者<sup>[2]</sup>普遍认为,领域知识和技能是影响创造力的一个主要因素,没有一定的某一领域的知识和技能,是不可能在该领域取得创造成果的。为了在某一领域中进行创造,人们必须有一定的关于该领域的知识,需要知道已经取得了那些成果、需要研究什么、如何进行研究、如何让别人理解自己的想法。不同领域创造力之间的相关系数只有 0.37;科学(包括物理、化学、生物等)各个分支具有许多共同的特点,如科学观念、科学素养、科学方法、科学能力等,因此可以作为一个独立的领域进行研究;第三,已有的科学创造力测验不能用于研究较大年龄范围内青少年科学创造力的发展趋势。Friedlander<sup>[3]</sup>设计了一份青少年科学创造力测量量表,143 名高中学生参加了测

试,测试问题主要出自植物和动物领域,要求学生处理数据收集、问题解决、假设提出和实验设计过程中的一些发散思维问题。Sinha 和 Singh<sup>[4]</sup>考察了科学创造力的概念,并设计了有 84 个项目的量表,用于测量青少年的科学创造力,该量表有灵活性、新颖性、观察能力、想象力、分析能力和转换能力六个维度。尽管这些量表可用于青少年科学创造力的测量,但由于包含过多的学科知识,只能测量特定年龄青少年的科学创造力。

鉴于以上情况,本研究旨在以科学创造力的结构模型、青少年科学创造力的表现及 Torrance 的创造性思维测验<sup>[5]</sup>为基础,编制适合东西文化背景的青少年科学创造力测验,检验其信度和效度,为今后开展相关的研究提供一个测量工具,并确定中英青少年科学创造力的常模。

## 2 研究方法

### 2.1 被试

初测:样本一为英国 Lady Margaret School 的 60 名女生。其中:12 岁的 30 名,14 岁的 30 名;样本二为山西一普通中学的 42 名中学生(初中二年级),其中:男生 25,女生 17。

内部一致性信度及结构效度的确定:样本一为英国 Trinity School 的 160 名学生(年龄及性别分布见表 1);样本二为山西省某一普通中学的

\*全国教育科学“十五”规划重点课题(课题号:ABA010012)、全国教育科学“十五”规划课题(课题号:FBB011041)和山西省归国留学课题“儿童青少年创造力的发展研究”的部分研究成果。

作者简介:申继亮(1965-),男,河南封丘人,北京师范大学发展心理研究所教授,博士生导师。

349 名中学生(年龄及性别分布见表 2)。

表 1 第二次施测被试的年龄及性别分布(英国)

年龄	12	13	15
男生	26	20	24
女生	32	29	29

表 2 第二次施测被试的年龄及性别分布(中国)

年龄	13	14	15	16	17
男生	41	37	32	29	37
女生	35	33	44	33	28

重测信度的确定:样本一为英国 Trinity School 的 160 名学生(年龄及性别分布见表 1);样本二为山西一所普通中学(城镇中学)12 岁的学生 76 名,其中男生 41 名,女生 35 名;

常模的确定:中国的被试来自山西省一所普通中学(城镇中学)和浙江省一所重点中学(城镇中学)的 1087 名学生,年龄及性别分布见表 3。

表 3 中国被试的年龄及性别分布

年龄	12	13	14	15	16	17	18
男生	41	95	88	89	196	55	73
女生	35	68	85	75	115	37	35

英国的被试来自 6 所城镇中学的 1087 名学生,年龄及性别分布见表 4。

表 4 英国被试的年龄及性别分布

年龄	11	12	13	14	15
男生	154	149	147	121	79
女生	101	110	101	73	52

## 2.2 研究程序

第一,根据我们提出的科学创造力的结构模型及对青少年科学创造力的表现的研究,参照已有的创造力量表,特别是 Torrance 创造性思维测验,编制了初步的青少年科学创造力测验,并进行了初测。根据试测结果,对测验及指导语进行了调整和修改。

第二,将修改后的量表施测于 Trinity School 的 160 名中学生和山西省某一普通中学的 349 名中学生,根据学生的得分,首先对测验的项目进行了分析,计算了项目之间以及项目与总量表之间的 Pearson 积差相关系数;其次,计算了 Cronbach 系数,确定了量表的内部一致性信度;再次,计算了当某一项目被删除后项目与总量表之间的相关系数以及 Cronbach 系数;最后,对量表进行探索性因素分析,确定其结构效度。

第三,两个评分者分别对 50 个被试的试卷按照评分标准进行独立评分,计算两组分值的 Pearson 积差相关系数,确定评分者信度。

第四,将修改后的量表施测于英国 Trinity School 的 160 名学生和山西一所普通中学初一(12 岁)的 76 名学生,确定量表的重测信度。前后两次测验相隔 3 个多月。参加测试的学生对测验有极大的兴趣,未发现特殊的问题。

第五,调查科学教育研究人员、中学科学教师及学生对量表的看法,确定量表的表面效度及学生对量表的接受程度。

第六,将修改后的量表施测于英国 6 所学校的 1087 名中学生和中国 2 所学校的 1087 名中学生,确定中英青少年科学创造力的常模。

## 2.3 研究工具

根据我们提出的科学创造力的结构模型及对青少年科学创造力的表现的研究,编制了青少年科学创造力测验。科学创造力的结构模型包括三个维度,即创造性的过程(主要包括思维和想象)、创造性的产品(包括科学问题、科学现象、科学知识、技术产品)和创造性的品质(包括流畅性、灵活性、独创性),构成 24 种成分。基于这 24 种成分,我们设计了调查问卷,以了解青少年科学创造力的主要表现。50 位科学教育研究者和中学科学教师填写了问卷,结果表明,5 个因素是最主要的,即观察与实验、科学知识的学习、科学问题的提出、科学问题的解决和科学创造活动。在对测试题目的选择中,主要考虑了 3 个方面,即被试的年龄、科学的特点和测验的管理。最初的测验有 9 个问题,Lady Margaret School 的 60 名中学生和山西一普通中学的 42 名中学生参加了初测。测试结束后,去掉了 2 个难度较大的题目,并对题目的顺序进行了调整,增加了答题示范,修改了指导语。

修改后的青少年科学创造力测验共有 7 个题目,每个题目后面给出一个答题示范,以帮助学生理解答题要求。第一题是物体应用,这一题目是参照托兰斯非常用途测验设计的,旨在考查青少年列举物体的科学应用时所表现出来的科学创造力的流畅性、灵活性和独创性;第二题是问题提出,以测量青少年对科学问题的敏感性。因为人们普遍认为,对科学问题的敏感性是科学创造力很重要的一个方面。第三题是产品改进,目的是考查青少年对技术产品创造性的改进能力,类似于托兰斯的产品改进测验,只不过这里除考察产品的新颖性、独特性外,还要考察其在科技方面的合理性。第四题是科学想象,以测量青少年创造

性的科学想象能力,这是科学创造力的核心成分之一。第五题是问题解决,旨在测量青少年创造性的问题解决能力,同时,该题还能反映青少年的空间想象能力,这种能力在科学创造中也是必不可少的。第六题是科学实验,用于考查青少年创造性的实验设计能力。第七题是产品设计,以测量青少年创造性的产品设计能力。我们设计科学实验和产品设计这两个问题,不仅是因为它们是青少年科学创造力的主要表现,而且还因为这种真实世界中的问题更能有效地测量青少年的科学创造力。本量表属于纸笔测验,所有问题都是开放性的,既适合于个体测试,又适合于集体施测,测验时间为60分钟。施测条件与平时考试相同,

监考人员要激发学生充分发挥其创造力,认真答题。评分按照答案所反映的被试在解答问题的过程中思维和想象的流畅性、灵活性和独创性来进行。

## 2.4 数据分析

数据采用 SPSS10.0 for Windows 进行统计和分析。

# 3 结果及分析

## 3.1 项目分析

项目分析包括项目之间的相关系数及项目与总量表的相关系数。我们计算了项目之间以及项目与总量表之间的 Pearson 积差相关,结果见表4。

表4 项目之间及项目与总量表之间的相关系数

项目	1	2	3	4	5	6	7	总量表
1 英	1.000							
中	1.000							
2 英	0.729	1.000						
中	0.702	1.000						
3 英	0.615	0.587	1.000					
中	0.583	0.570	1.000					
4 英	0.661	0.638	0.638	1.000				
中	0.477	0.557	0.527	1.000				
5 英	0.540	0.421	0.578	0.526	1.000			
中	0.515	0.350	0.420	0.457	1.000			
6 英	0.479	0.502	0.510	0.558	0.345	1.000		
中	0.409	0.440	0.423	0.475	0.346	1.000		
7 英	0.570	0.564	0.632	0.596	0.440	0.641	1.000	
中	0.402	0.432	0.483	0.425	0.441	0.458	1.000	
总量表								
英	0.829	0.825	0.815	0.835	0.654	0.751	0.800	1.000
中	0.793	0.746	0.684	0.616	0.598	0.695	0.800	1.000

注:采用双尾检验方法,检验相关系数的显著性水平,对所有相关系数, $p < .01$ 。

英国: $N = 160$ ;中国: $N = 160$ 。

从表4可以看出,对于英国被试,项目之间的相关系数在0.345(第5题与第6题之间)到0.729(第1题与第2题之间)之间,且大多数项目之间的相关系数在0.40以上,项目与总量表之间的相关系数在0.654到0.829之间,所有相关系数的显著性水平均在0.01以上。对于中国被试,项目之间的相关系数在0.346(第5题与第6题之间)到0.702(第1题与第2题之间)之间,且大多数项目之间的相关系数在0.40以上,项目与总量表之间的相关系数在0.598到0.800之间,所有相关系数的显著性水平均在0.01以上。这一结果表明,该量表具有较高的内部一致性,所有项目及总量表测量被试的同一特质,即青少年在科学思维与想象中的流畅性、灵活性和独创性。

## 3.2 信度分析

本测验收集了三个方面的信度指标。第一,Cronbach 系数。基于 Trinity School 的160名学生的分值,其信度系数为0.893,基于山西省某一普通中学的349名中学生的分值,其信度系数为0.775。该数据表明,本测验具有较高的内部一致性。在此基础上,我们又计算了当某一项目被删除后项目与总量表之间的相关系数以及 Cronbach 系数,结果见表5。表中数据表明,所有修正的项目与总量表之间的相关系数以及 Cronbach 系数均小于项目未被删除时的值,这一结果说明,对于青少年科学创造力的测量,量表中的每一个项目都是必要的。

第二,评分者信度。两个评分者分别对50个

被试的试卷按照评分标准进行了独立评分,评分者之一是研究者本人,另一位评分者是与本研究无关的中学科学教师,我们计算了两组分值的 Pearson 积差相关系数,结果见表 6。从表 5 可以看出,两组分值的 Pearson 积差相关系数在 0.793 到 0.913 之间,整个量表的重测信度为 0.916,表明评分标准具有客观性。

表 5 某一项目被删除后项目与总量表之间的相关系数及 Cronbach 系数

项目	修正后的项目与总量表之间的相关系数		项目被删除后的值	
	英	中	英	中
1	0.761	0.701	0.869	0.714
2	0.727	0.653	0.875	0.723
3	0.741	0.617	0.871	0.734
4	0.766	0.593	0.868	0.749
5	0.576	0.516	0.892	0.766
6	0.630	0.601	0.888	0.765
7	0.731	0.712	0.873	0.750

注:英国:  $N = 160$ ; 中国:  $N = 349$ 。

第三,重测信度。我们利用修改后的量表对英国的 160 名中学生和山西省一所普通中学初一的 76 名学生进行了两次测试(测试的时间间隔都在 3 个月以上),然后,计算了两次测试成绩之间的 Pearson 积差相关系数,结果见表 7。由表可看出,该量表具有较高的重测信度。

### 3.3 效度分析

本研究对测验的结构效度进行了分析。Guilford(1950)认为,对创造力测验的有效性检验的第

表 6 评分者信度( $N = 50$ )

项目	1	2	3	4	5	6	7	总量表
信度	0.880	0.897	0.913	0.875	0.867	0.823	0.793	0.916

表 7 重测信度

项目	1	2	3	4	5	6	7	总量表
信度 英	0.871	0.820	0.835	0.762	0.879	0.866	0.805	0.849
中	0.864	0.793	0.876	0.748	0.910	0.892	0.795	0.857

注:英国:  $N = 160$ ; 中国:  $N = 76$ 。

表 8 测验项目的因素负荷

项目	1	2	3	4	5	6	7
因素负荷 英	0.836	0.809	0.827	0.840	0.685	0.724	0.805
中	0.773	0.754	0.797	0.747	0.628	0.596	0.636

注:英国:  $N = 160$ ; 中国:  $N = 349$

### 3.4 确定常模

在对量表的项目、信度、效度等进行分析的基础上,我们分别选取了英国 6 所学校的 1087 名被

试和中国两所学校的 1087 名被试,确定了英国及中国青少年科学创造力的常模,详细情况见表 9 和表 10。

一步就是进行因素分析,计算其结构效度。本研究对 7 个项目进行探索性因素分析,结果仅获得一个因素,并且不能旋转(见表 8 所示)。对英国被试,因素负荷从 0.685 到 0.840,解释率为 63%。对中国被试,因素负荷从 0.596 到 0.797,解释率为 56%。Horn 和 Cattell(1966)对各种能力测验进行了因素分析,结果表明:创造力是一个被称为“一般流畅性”的能力因素,这一能力因素主要落在思想流畅性(ideational fluency)、联想流畅性(association fluency)和遥远联想测验(irrelevant associations test)。Kline(1993)认为,对于一种单一因素能力的测验项目进行因素分析,各测验项目应落在一个因素上,且因素负荷应大于 0.3。我们认为,创造力是根据一定的目的,运用一切已知信息,在独特地、新颖地且有价值地(或恰当地)产生某种产品的过程中,表现出来的智能品质或能力。创造力既是一种心理过程,但更重要的是与个性特征相联系,表现出个性差异,亦即流畅性、灵活性、独特性等智能品质。虽然科学创造力测验包括物体应用、问题提出、产品改进、创造想象、问题解决、实验设计和产品设计 7 个方面,但都测量被试的流畅性、灵活性、独创性等智能品质。在评分过程中,我们只考虑流畅性、灵活性、独创性三种品质的总分,因此,可以将此认为是一个因素。从这个角度来讲,各个测验项目落在同一个因素上是合理的,本测验具有较好的结构效度。

表 9 英国青少年科学创造力常模 (N = 1087)

等级	累积百分数	各年龄阶段的得分					T 分数
		11	12	13	14	15	
E	1	9.56	10.00	18.45	14.95	13.60	27
	3	10.00	13.00	26.00	17.85	17.96	31
	5	13.80	16.00	28.00	20.00	22.20	34
	10	16.60	19.00	31.00	23.00	27.00	37
	15	19.00	24.00	33.00	26.00	32.00	40
D	20	23.00	27.00	38.00	28.00	37.00	42
	25	25.00	30.00	40.00	30.00	40.00	43
	30	27.00	33.00	42.00	33.00	42.00	45
	35	29.00	34.00	45.00	37.00	45.00	46
	40	31.00	36.00	46.00	41.00	48.00	47
C	45	33.00	40.00	48.00	44.00	53.00	49
	50	36.00	43.00	50.00	46.00	60.00	50
	55	38.00	46.00	52.00	48.00	61.00	51
	60	40.00	46.00	54.00	50.00	65.00	52
	65	42.00	50.00	58.00	51.00	66.00	54
B	70	44.00	53.00	61.00	55.00	68.00	55
	75	47.00	55.00	62.00	58.00	73.00	57
	80	50.00	59.00	65.00	61.00	75.00	58
	85	53.00	65.00	69.00	65.00	80.00	60
	90	57.00	69.00	74.00	73.00	82.00	63
A	95	68.00	76.00	83.00	77.00	90.00	66
	97	74.24	79.00	89.00	88.00	92.00	69
	99	85.08	92.00	104.00	94.00	108.00	73
	平均数	36.937	43.602	52.097	46.273	56.573	
	标准差	16.448	18.351	16.915	18.167	21.279	
	人数	255	259	248	194	131	

注:A、B、C、D、E代表分数等级,其中A代表前10%学生的分数等级,B代表前30%-10%学生的分数等级,C代表前70%-30%学生的分数等级,D代表后30%-10%学生的分数等级,E代表后10%学生的分数等级。T分数栏中的每一个值代表与累积百分数及测验分相对应的T分数值。下同。

表 10 中国青少年科学创造力常模 (N = 1087)

等级	累积百分数	各年龄阶段的得分							T 分数
		12	13	14	15	16	17	18	
E	1	5.00	7.20	9.00	12.60	10.72	19.00	14.08	27
	3	6.31	10.92	15.22	20.95	20.36	30.95	27.27	31
	5	7.85	13.00	17.00	24.00	23.00	33.00	30.00	34
	10	9.00	17.00	20.40	28.50	30.20	37.00	34.90	37
	15	11.10	19.00	22.00	32.00	35.00	39.90	39.00	40
D	20	12.00	20.80	23.80	33.00	38.00	43.60	40.80	42
	25	13.00	23.00	25.00	35.25	41.00	45.25	42.00	43
	30	14.10	26.00	27.00	37.00	44.60	47.00	43.70	45
	35	17.00	28.00	28.00	38.00	46.00	50.00	46.15	46
	40	19.80	30.00	30.00	39.00	48.00	51.20	48.00	47
C	45	21.65	32.00	31.00	42.00	50.00	52.85	52.05	49
	50	23.00	34.00	32.00	43.00	53.00	54.00	55.00	50
	55	25.35	35.00	33.70	46.00	56.00	56.15	56.00	51
	60	27.00	37.00	36.00	48.00	58.00	59.00	58.00	52
	65	29.00	38.00	38.00	50.00	61.00	60.00	59.00	54
B	70	30.00	41.00	38.80	52.00	64.00	63.00	61.00	55
	75	32.00	44.00	41.00	55.75	68.00	66.00	64.75	57
	80	32.60	48.20	45.00	59.00	70.00	69.80	67.20	58
	85	35.45	53.40	49.90	61.00	73.00	75.00	78.10	60
	90	45.30	57.00	54.00	67.50	78.00	79.00	86.55	63
A	95	59.35	60.00	65.00	76.75	83.00	82.00	95.65	66
	97	68.38	69.08	67.34	84.15	87.00	83.63	95.65	69
	99	82.00	83.44	84.12	95.60	99.68	92.00	105.64	73
	平均数	26.145	35.055	34.988	46.055	53.752	55.967	54.880	
	标准差	15.091	15.583	14.051	15.803	18.132	14.963	16.986	
	人数	76	163	173	164	311	92	108	

## 4 研究结论

(1)《青少年科学创造力测验》具有较高的信度,Cronbach 系数、评分者信度、重测信度均达到心理测验学要求的水平。

(2)《青少年科学创造力测验》具有较高的结构效度。

### 参考文献:

[1]朱棣文.最重要的是创新精神.人民日报(海外版),2000

年7月25日.

[2] Alexander P. A. Domain knowledge: Evolving themes and emerging concerns. *Educational Psychology*, 1992,27(1),33 - 51.

[3]Friedlander M. A natural science creativity test as a prediction of creative thinking in science. *Creative Child & Adult Quarterly*, 1983,8, 211 - 215.

[4]Sinha A. K. Singh C. Measurement of scientific creativity. *Indian Journal of Psychometry & Education*, 1987,18, 1 - 13.

[5]Torrance E. P. *Torrance tests of creative thinking: Normal - technical manual*. Princeton Publishers,1974.

## Constructing of the Scientific Creativity Test for Adolescent

SHEN Ji-liang<sup>1</sup> HU Wei-ping<sup>2</sup> LIN Chong-de<sup>1</sup>

(1. *Institute of Developmental Psychology, Beijing Normal University, Beijing, 100875*)

(2. *Curriculum and Instruction Institute, Shanxi Teachers University, Linfen, Shanxi, 041004*)

**Abstract:** On the basis of Scientific Creativity Structure Model and the Torrance Tests of Creative Thinking, the Scientific Creativity Test for Adolescent (SCTA), a 7-item scale, was developed. Results based on the data of 1307 English adolescents and 1087 Chinese adolescents showed that: (1) the SCTA has high reliability, including Coefficient alpha, rater reliability coefficient and test-retest reliability; (2) the SCTA has high construct validity. Result of exploring factor analysis showed that the test has only one factor.

**Key words:** adolescent; scientific creativity; test