

公众教育集团评价技术优势

一、 专业性

- ✚ 基于国际先进的第五代教育评价理论；
- ✚ 高度匹配国家教育改革的方向、理念与模式；
- ✚ 研发时间最久(26年)，实测样本量最大(超过1200万人次)，评价模型验证度最高，数据积累最多；
- ✚ 深得国内外教育厅局、全日制学校、学生/家长的认可和赞许；
- ✚ 支撑多个国家级、省级教育科学“十二五”/“十三五”课题；
- ✚ 超过十年的教育大数据应用经验，包括评价报告的深度解读、学习/教学改进、学习/教学规划，以及历时性、共时性分析等。

二、 技术性

- ✚ 高度匹配国家教育改革的方向、理念与模式；
- ✚ 学业评价技术改写“双基”历史，独创“三维”模式，不仅能准确发现学生学习/教师教学中的长短版、优劣势，更能诊断成因，并给出科学的改进和规划建议；
- ✚ 发展潜能评价技术打破传统潜能测试“黑箱”，方法更科学、更透明、更精密，能做到更有效地诊断学生学习中的深层问题和预测学生未来发展，并给出科学稳定的改进方法和生涯发展规划；
- ✚ 综合素质评价技术不仅了解学生的认知、生活与社会行为，更去探索背后的偏好、习惯、遇到的问题、生活方式等，做出因果性分析，并给出学生素质成长方面的改进和规划建议。

三、 专业团队

- ✚ 拥有一支20位世界级顶级教育行业专家团队，其中16位拥有博士学位，并且都是行业内各自研究领域的带头人；
- ✚ 培训团队都是由行业内从事教育质量评价多年及有丰富一线教学经验的资深教授、研究员、教师组成，同时还拥有教育部、中国教育学会

的教育专家组成的高学历、高水平、高素质的一支专业专家咨询团队。培训团队运用公司行业内领先的综合评价技术，结合丰富的理论知识以及多年的实践经验，为沈河区的此次评价项目保驾护航。

四、 应用性

- ✚ 普及性推广和既有资源利用度极高；独创的将测评的“教育测量”、“数据处理”、“教育评价”进行功能分离，分别形成高度技术化，从而使整个评价工作可规模、高效实施；并可最大限度的利用既有试题试卷资源。实施容易：极大降低教管教研人员、教师队伍对评价技术的掌握和运用门槛；资源通用：不依赖专有试题试卷，支持学校自主命题，也提供丰富的成熟、优质题库卷库供调用；过程高效：高度满足实际应用场景的“短、频、快”需求；结果可靠：规模实施结果的等值化、标准化水平超过十年验证；
- ✚ 为 100 多个市/区(县)教育局提供教育质量综合评价及相关大数据服务，包括四川省成都市中小学教育质量综合评价改革实验、四川省成都市青羊区学业质量评价改革实验、黑龙江省哈尔滨市香坊区中小学教育质量综合评价改革实验、吉林省教科院义务教育质量综合评价改革实验、辽宁省沈阳市义务教育质量综合评价改革实验、河北省石家庄市桥西区学业质量评价改革实验、陕西省西安市中小学教育质量综合评价改革实验、江苏省镇江市扬中市中小学质量综合评价改革实验、广东省广州市越秀区、白沙区、天河区中小学教育质量综合评价改革实验；
- ✚ 为 10 余个省/直辖市教育厅局提供教育质量监测，如四川省教育质量监测（2014）、成都市义务教育质量监测（2014/2015/2016/2017）、成都市教育现代化发展水平监测项目（2017）、成都市教育校际均衡监测项目（2017）、河南省义务教育质量监测（2016/2017）；

五、 品牌影响力

- ✚ 行业公认的国内现代教育测评开创者和领导者；
- ✚ 多个国家级、省级教育科学“十二五”/“十三五”课题以公司核心教

育测评和大数据应用为专业基础与技术支撑；

- ✚ 教育评价技术获得中国教育学会唯一认可和推荐；
- ✚ 在国家中小学教育质量综合评价改革实验中的成果获得教育部唯一认可；
- ✚ IAP 中小学生综合竞赛成绩与奖项，获得多所大学自主招生报名资格认可；
- ✚ 主导研发的教育咨询师认证标准获得中国教育学会教育咨询师认证中心认可并颁布；
- ✚ 教育智能测评服务现为国内唯一超过 1,200 万份样本的成熟评价体系，覆盖 22 个省市、200 多个区县、20,000 多所学校；
- ✚ 公众教育的评价与考试技术与服务，得到了国内诸多教育厅局、全日制学校、教培机构及学生/家长的广泛认同和赞许；
- ✚ 公众教育的评价与考试技术与服务，得到了美国 64 个学区及大部分州立大学的认可，另得到加拿大多个学区、高中及大学的认可。

中小学教育质量综合评价项目介绍

一、中小学教育质量综合评价项目的背景

2013年6月教育部颁发《关于推进中小学教育质量综合评价改革的意见》(教基二[2013]2号)中提出:教育质量评价具有重要的导向作用,是教育综合改革的关键环节。推进中小学教育质量综合评价改革,是推动中小学全面贯彻党的教育方针、全面实施素质教育、落实立德树人根本任务的重要举措,是引导社会和家长树立科学的教育质量观、营造良好育人环境的迫切需要,是基本实现教育现代化、加强和改进教育宏观管理的必然要求。改革开放特别是新世纪以来,随着基础教育课程改革的实施,各地在改进中小学教育质量评价方面进行了积极探索,取得了一些进展。但总体上看,由于教育内外部多方面的原因,单纯以学生学业考试成绩和学校升学率评价中小学教育质量的倾向还没有得到根本扭转,突出表现为:在评价内容上重考试分数忽视学生综合素质和个性发展,在评价方式上重最终结果忽视学校进步和努力程度,在评价结果使用上重甄别证明忽视诊断和改进。这些问题严重影响了学生的全面发展、健康成长,制约了学生社会责任感、创新精神和实践能力的培养。要解决这些突出问题,适应经济社会和教育事业发展的新形势新要求,必须大力推进中小学教育质量综合评价改革。

二、测量工具介绍¹

(一) 学业发展水平测量工具

我国2001年的《基础教育课程改革纲要(试行)》和2002年《教育部关于积极推进中小学评价与考试制度改革的通知》及相关文件,给出了“三维”的学科教育目标,学科教育目标不再是过去的“双基”目标。依据学科教育三维目标与学科课程标准的要求,学业评价应解决学生、学校面向学科学习“三维目标”的测量与评价问题,而不再是“基础知识、基本技能”的两目标测量评价。

2013年教育部颁布的《关于推进中小学教育质量综合评价改革的意见》(教基二[2013]2号)中提出了学业发展水平作为评价学校教育质量的五大块内容之

¹测量工具介绍中信效度等所使用的数据是中国教育学会编制的测量工具在某市正式施测后的数据。

一,主要考查学生对各学科课程标准所要求内容的掌握情况,可以通过知识技能、学科思想方法、实践能力、创新意识等关键性指标进行评价,促进学生打好终身学习和发展的基础。

2014年教育部颁布《关于全面深化课程改革落实立德树人根本任务的意见》提出了核心素养和学业质量标准,官方解读为:核心素养是必备品格和关键能力,实质等于把三维目标的“知识、技能”及思想方法等提炼为关键能力,把“过程、方法”与“情感、价值”提炼为必备品格——也即在实质上学科学习三维目标与学科核心素养,是实质等同的。

由此,中小学教育质量综合评价改革中的学业发展水平评价,应根据学科学习三维目标(或学科核心素养)设立测量评价指标体系,并实施满足三维结构要求的学业测量、评价。

1. 测试科目

项目方可以根据教育局或者当地的需求选择所要测试的科目,可以选择当地教研员命制的试题,然后由中国教育学会进行试题指导,也可以交由公众教育集团来进行试题命制。

2. 测试年级

项目方可以根据教育局或者当地的需求选择所要测试的年级。

3. 小学5年级数学测量工具介绍

(1) 测量工具的指标编制依据

此次5年级学业发展水平(数学)测量工具的编制主要是在第四代教育评价理论框架下,以《教育部义务教育数学课程标准(2011版)》、学科三维目标(学科核心素养)为依据,参考国际学生评估项目(PISA)数学素养的框架来制定此次测试的指标和结构。

2011版的《义务教育数学课程标准》中对学段(4-6年级)的要求以及PISA数学素养框架与试卷各指标之间的对应关系见下表:

表1 2011版义务教育课程标准与某市5年级学业发展水平(数学)指标对应表

课标要求(4-6年级)	本次测试指标
理解分数、百分数的意义,了解负数;了解一些几何体和平面图形的基本特征;	认知技能
掌握必要的运算技能;理解估算的意义;	计算技能
了解确定物体位置的方法,掌握测量、识图和画图的基本方法	操作技能
经历数据的收集、整理和分析的过程,掌握一些简单的数据处理	应用技能

技能；	
能根据解决问题的需要，收集与表示数据，归纳出有用的信息	推理技能
能探索分析问题、解决问题的有效方法，了解解决问题方法的多样性	问题解决技能
能进行有条理的思考，能清楚地表达思考的过程与结果	数学表达技能
能够对生活中的数字信息作出合理的解释，会用数（合适的量纲）、字母和图表描述生活中的简单问题	语言理解能力
初步形成数感，发展符号意识	数字推理能力
体验事件发生的等可能性，掌握简单的计算等可能性的方法；在与他人交流过程中，能够进行简单的辩论	逻辑分析能力
在探索简单图形的性质、运动现象的过程中，初步形成空间观念	空间想象能力

表 2 某市 5 年级学业发展水平（数学）指标与 PISA 数学素养框架对应关系

PISA 数学素养框架		本次测试指标
表述	能够将具体情境中的内容转换成数学解释的方式，用数学结构加以表述，明确可变因素，做简单的假设来帮助解决问题或满足要求。	操作技能 数学表达技能 语言理解能力
运用	进行计算、列代数表达式、方程式以及其他数学模型，以数学的方式从数学图表中分析信息，进行数学描述和解释，运用数学工具来解决问题。	计算技能、应用技能、 操作技能、问题解决技能、 数字推理能力、逻辑分析能力
阐述	对数学解答进行评价，根据问题情境进行推理，判断在该情境下相应的数学解答是否合理，是否讲得通。	推理技能、数学表达技能、 逻辑分析能力

（2）工具编制流程

根据课程标准，某市所使用的教材以及截至测试前的教学进度，采用标准化命题方式，以中国教育学会采用的中小学教育质量综合评价改革实验专用技术——ACTS 学业（素质能力）评价技术系统作为处理、分析工具。试卷的编制遵照以下流程进行：

依据本次测试目的确定测量结构；

根据课程标准对本学年学生在该学科上应达到的要求确定测量内容的范围；

准备试卷多向细目表；

根据试卷多向细目表编题；

题目审查；

试卷定稿，编写试卷使用手册。

（3）测量内容的结构

5 年级学业发展水平（数学）测量内容的结构见下表：

表 3 某市 5 年级学业发展水平（数学）测量内容的结构

知识	技能	能力
----	----	----

指标/内容	满分	指标/内容	满分	指标/内容	满分
数与代数	43.0	认知	4.3	语言理解	40.1
图形与空间	22.0	计算	29.7	数字推理	19.8
统计与概率	22.0	应用	14.0	逻辑分析	21.3
实践与综合	13.0	操作	14.2	空间想象	18.8
		推理	15.3		
		问题解决	11.9		
		数学表达	10.6		
小计	100	小计	100	小计	100

(4) 测量工具

主要采用第四代教育评价理论的框架，依据认知诊断理论（Cognitive Diagnosis Theory）和实质性评价理论及其方法编制学业评价试卷——新一代学科学习测量工具。

(5) 测量工具的信度

用 SPSS 23.0 软件分析上述测量工具的内部一致性信度，分析结果见下表：

表 4 某市 5 年级学业发展水平（数学）测量工具可靠性检验结果

知识		技能		能力	
Cronbach's Alpha	基于标准化项的 Cronbach's Alpha	Cronbach's Alpha	基于标准化项的 Cronbach's Alpha	Cronbach's Alpha	基于标准化项的 Cronbach's Alpha
0.839	0.889	0.837	0.871	0.833	0.852
检验项目数	4	检验项目数	7	检验项目数	4

由上表可知，5 年级学业发展水平（数学）测量工具各个指标的信度系数均大于 0.8，符合《合同》以及《基础教育学业质量监测工具质量评价规范》（地方标准）中的要求，说明此次测量工具是可靠的。

(6) 测量工具的效度

学业发展水平（数学）测量工具的效度，见下表：

表 5 某市 5 年级学业发展水平（数学）测量工具的结构效度

一级指标	二级指标	因子载荷	CFI	TLI	RMSEA
知识	数与代数	0.870	1.000	1.000	0.000
	图形与空间	0.851			
	统计与概率	0.746			
	实践与综合	0.800			
技能	认知	0.438	0.986	0.979	0.057
	计算	0.810			
	应用	0.847			

	操作	0.517			
	推理	0.755			
	问题解决	0.762			
	数学表达	0.797			
能力	语言理解	0.887	0.998	0.993	0.045
	数字推理	0.598			
	逻辑分析	0.852			
	空间想象	0.741			

由上表可知,5 年级学业发展水平(数学)测量工具的拟合指标均符合《合同》中“测试工具结构效度拟合指数应至少在 0.8 以上以及 RMSEA 应至少小于 0.08”的要求;以及《基础教育学业质量监测工具质量评价规范》(地方标准 DB51)中“模型拟合指数应大于等于 0.9, RMSEA 应小于 0.1, 因子载荷大于 0.2”的要求,说明此次测量工具具有良好的结构效度。

(7) 工具的区分度

本次学业发展水平测试属于基于标准的测量,目的在于评估教学实践达成课程标准的程度,理论上,如果学生根据标准获得了他们需要掌握的东西,就达到了特定标准的要求。越多学生获得满分,意味着教学实践达成课程标准的程度越好,这对试卷区分度不做要求。因此,报告不对测量工具的区分度进行分析。

(二) 综合素质测量工具

2002 年《教育部关于积极推进中小学评价与考试制度改革的通知》提出了建立以促进学生发展为目标的评价体系,其中包括**基础性发展目标**,主要内容有:道德品质、公民素养、学习能力、交流与合作能力、运动与健康以及审美与表现。2013 年教育部颁布的《关于推进中小学教育质量综合评价改革的意见》(教基二[2013]2 号)提出了建立综合评价指标体系,其中将品德发展水平(包含行为习惯、公民素养、人格品质、理想信念等关键性指标)、学业发展水平(包含实践能力、创新意识等关键性指标)、身心发展水平(包含身体形态技能、审美修养以及情绪行为调控等指标)作为评价学校教学质量的内容。通过对学生全面发展状况的调查和分析,从而促进学生全面发展、健康成长。

基于基础性发展目标和核心素养的综合素质评价,属于质性研究领域中的实质性评价范畴,主要理论支柱为文化人类学、解释学、认知建构理论。综合素质评价目前主要有两类——真实性评价(写实记录)和表现性评价(活动、竞赛等展现)。在中小学教育质量综合评价改革实验中的综合素质评价,主要采用表现性评价。

1. 测试年级

项目方可以根据教育局或者当地的需求选择所要测试的年级。综合素质测量工具主要为量表,小学综合素质测量工具包括两份量表,其中量表(一)包括行为习惯、公民素养、人格品质三方面内容,量表(二)包括实践能力、创新意识和情绪行为调控三方面内容;中学综合素质测量工具包括两份量表,其中量表(一)包括行为习惯和公民素养两方面内容,量表(二)包括人格品质、实践能力、创新意识和情绪行为调控四个方面内容。量表(一)和量表(二)的题目均为单项选择题。

2. 小学综合素质测量工具介绍

(1) 测量工具的编制依据

综合素质测量工具的主要编制依据是基础性发展目标、核心素养体系、社会主义核心价值观、义务教育课程方案和相关学科课程标准、《中小学德育工作规程》、《中共中央国务院关于进一步加强和改进未成年人思想道德建设的若干意见》、《中学生日常行为规范(修订)》、《中小学文明礼仪教育指导纲要》、《中小学心理健康教育指导纲要(2012年(修订))》等。

(2) 测量内容的结构

综合素质测量内容的结构见下表:

表 6 小学综合素质测量内容的结构

内容 指标	行为 习惯	公民素养				人格品质				实践能力		创 新 意 识	情 绪 行 为 调 控
		规 则 意 识 和 行 为	权 益 责 任 意 识 和 行 为	学 校 社 区 活 动 参 与 度	尊 重 和 环 境 爱 护	乐 群 性	有 恒 性	独 立 性	自 律 性	学 习 动 力	学 习 方 法 与 技 能		
内容 指标 (细 分)	/											/	/

注:表中加底纹的指标为二级指标,未加底纹的为三级指标。

(3) 测量工具

主要采用第四代教育评价理论框架,依据多维项目反应理论及实质性评价理

论及其方法编制综合素质量表。

(4) 测量工具的信度

对综合素质测量工具各指标的内部一致性进行分析，得到如下结果：

表 7 某市 5 年级综合素质测量工具的可靠性检验结果

二级指标	三级指标	样本量	Cronbach's Alpha	基于标准化项的 Cronbach's Alpha
行为习惯	单一指标	14827	0.855	0.859
公民素养	规则意识与行为	14827	0.857	0.866
	权益责任意识和行为			
	学校社区活动参与度			
	尊重和爱护环境			
人格品质	乐群性	14827	0.660	0.701
	有恒性			
	独立性			
	自律性			
实践能力	学习动力	14885	0.898	0.904
	学习方法与技能			
创新意识	单一指标	14885	0.872	0.876
情绪行为调控	单一指标	14885	0.691	0.705

由上表可知，综合素质测量工具的内部一致性系数都大于 0.66，符合《合同》以及《基础教育学业质量监测工具质量评价规范》（地方标准）中的要求，说明此次测量工具是可靠的。

(5) 测量工具的效度

对综合素质测量工具各指标的结构效度进行分析，得到如下结果：

表 8 某市 5 年级综合素质测量工具的结构效度

二级指标	三级指标	因子载荷	CFI	TLI	RMSEA
行为习惯	单一指标	0.225-0.732	0.965	0.961	0.054
公民素养	规则意识和行为	0.461-0.666	0.900	0.892	0.057
	权益责任意识和行为	0.168-0.677			
	学校社区活动参与度	0.370-0.805			
	尊重和环境爱护	0.369-0.662			
人格品质	乐群性	0.115-0.730	0.922	0.844	0.064
	有恒性	0.487-0.758	0.983	0.978	0.036
	独立性	0.257-0.631	0.913	0.87	0.055
	自律性	0.380-0.684	0.972	0.963	0.041

实践能力	学习动力	0.415-0.715	0.921	0.917	0.046
	学习方法与技能	0.361-0.720			
创新意识	单一指标	0.394-0.687	0.963	0.96	0.036
情绪行为调控	单一指标	0.318-0.709	0.909	0.889	0.059

由上表可知，综合素质测量工具各指标的 CFI 和 TLI 均大于 0.8，RMSEA 均小于 0.08，也就是说所有拟合指标均符合《合同》中的要求，说明此次测量工具具有良好的结构效度。

（三）发展潜能测量工具

当今基础教育改革中，学生的多元、个性、全面综合发展是一个重要目标，而学生多元、个性、全面综合发展不仅包括已表现出来的可继续发展的部分，还包括尚未发现和挖掘的可发展的潜质部分，这就涉及了对学生发展潜能的开发。特别地，在新一轮高中课程改革中，“选课走班”“选科定向”成为新课改的一大特点，为适应社会对多样化人才的需求，满足不同学生的发展需要，在保证每个学生达到共同基础的前提下，各学科分类别、分层次设计了多样的、可供不同发展潜能学生选择的课程内容，以满足学生对课程的不同需求。这就要求学生对自己的兴趣特长（发展潜能）和优势有一定的了解，选择适合自己的课程，更好地计划自己未来的发展规划。

本次发展潜能评价使用的测量工具是由中国教育学会委托的公众教育科学研究院和公众考试研究院联合研发、采用多维项目反应理论及其方法编制的发展潜能量表，并采用 PCTS 发展潜能评价技术系统作为处理、分析数据的评价工具。

1. 测试年级

项目方可以根据教育局或者当地的需求选择所要测试的年级。

2. 小学发展潜能测量工具介绍

（1）测量工具的编制依据

发展潜能测量工具的主要编制依据是加德纳的多元智能理论，参考卡特尔的流体智力理论、吉尔福特的智力三维结构模型、瑟斯顿的群因素论等。

（2）测量内容的结构

发展潜能测量内容的结构见下表：

表 9 小学发展潜能测量内容的结构

认知能力倾向	认知过程	项目
--------	------	----

指标/内容	满分	指标/内容	满分	指标/内容	满分
词汇辨析	11.1	解释	8	词	20
语言理解	24.7	类比	18.4	句与段落	16
数字推理	26.3	推理	24.9	图形与空间	30
逻辑分析	29.5	操作	11.5	数与代数	20
空间想象	28.4	分析	18.8	实践与综合	14
		问题解决	18.4		
小计	120	小计	100	小计	100

(3) 测量工具

主要采用多维项目反应理论及其方法编制的发展潜能量表。

(4) 测量工具的信度

对发展潜能中认知能力倾向²的内部一致性进行分析,得到如下结果:

表 10 某市 5 年级发展潜能测量工具可靠性检验结果

认知能力倾向		
Cronbach's Alpha	基于标准化项的 Cronbach's Alpha	检验项目数
0.835	0.842	5

由上表可知,发展潜能测量工具中认知能力倾向的内部一致性系数和标准化后的内部一致性系数都大于 0.8,表明该测量工具中认知能力倾向的信度符合《合同》以及《基础教育学业质量监测工具质量评价规范》(地方标准)中的要求,说明此次测量工具是可靠的。

(5) 测量工具的效度

①效标效度

将量表与学科成绩进行相关分析,得到如下相关关系表:

表 11 某市 5 年级学业(数学)和发展潜能测量工具各指标间的相关关系表

	认知能力倾向	认知过程	项目
知识	.682**	.678**	.682**
技能	.680**	.676**	.680**
能力	.674**	.672**	.675**

**在.01水平(双侧)上显著相关。

发展潜能测量工具各指标与学科成绩各项指标的相关系数达到 0.6 以上,符合《合同》中“潜能与学业发展水平(数学)之间的相关系数应至少在 0.5 以上”的要求,说明量表具有较高的效标效度。

²发展潜能量表中,认知能力倾向是主要诊断、甄别指标,因此只介绍该指标的信度、效度和区分度。

②结构效度

对发展潜能测量工具各指标的结构效度进行分析，得到如下结果：

表 12 某市 5 年级发展潜能测量工具的结构效度

一级指标	二级指标	因子载荷	CFI	TLI	RMSEA
认知能力倾向	词汇辨析	0.561	0.985	0.971	0.077
	语言理解	0.770			
	数字推理	0.740			
	逻辑分析	0.823			
	空间想象	0.708			

由上表可知，5 年级发展潜能测量工具的拟合指标符合《合同》以及《基础教育学业质量监测工具质量评价规范》（地方标准）中的要求，说明此次测量工具具有良好的结构效度。

（6）工具的区分度

对发展潜能测量工具各分指标的区分度进行分析，得到如下结果：

表 13 某市 5 年级发展潜能测量工具（认知能力倾向）各指标的区分度

指标	区分度	区分度水平 ³
词汇辨析	0.38	好
语言理解	0.42	很好
数字推理	0.45	很好
逻辑分析	0.41	很好
空间想象	0.37	好
认知能力倾向	0.41	很好

注：表中加底纹的指标为一级指标。

通过上表可知，所有指标的区分度都能达到“好”以上，并且“语言理解”、“数字推理”、“逻辑分析”和“认知能力倾向”都能达到“很好”水平。由此可知 5 年级发展潜能测量工具整体的区分度较好，符合测量学要求。

三、评价工具介绍

（一）ACTS 学业评价技术系统

学业评价所采用的评价工具，依托于 ACTS 学业评价技术，符合义务教育学科课程标准中对各学科所要求掌握内容的；同时也满足《国家义务教育质量监测方案》中要“测查学生掌握知识、技能的程度和分析解决问题的能力”的相关要

³这里区分度水平的评定依据是《基础教育学业质量监测工具质量评价规范》（地方标准）中的区分度水平标准。

求，并能对采购方研制的试题试卷进行“知识、技能、能力”三大类指标的具体解析和评分标准的制定，评价数据报告能呈现出具体的知识、技能、能力的维度、指标数据，和明确的测量、评价标准；并能够提供多元、多维、多种比较的评估报告；评价指标之间的关联关系具有强大的诊断功能，在帮助学生发展、教师专业提升及教学改进、学校科学教研及教管方面取得了显著成效。

1. 数据处理系统

- 三维阅卷系统；
- 三维多元数据处理系统；
- 评价、甄别、诊断分析系统；
- 历时数据处理系统
- 历时评价、甄别、诊断分析系统；
- 决策性评价工具系统；
- 认定性评价工具系统；
- 服务性评价工具系统。

2. 评价报告生成系统

- 学生个体评价报告生成系统；
- 班级评价报告生成系统；
- 年级评价报告生成系统；
- 区域分级评价报告生成系统。
- 班级评价报告分析系统；
- 年级评价报告分析系统；
- 区域分级评价报告分析系统；
- 认定性评价报告分析系统；
- 决策性评价报告分析系统。

3. 评价报告系统

- 学生——评价、甄别、诊断、成长分析报告
- 班级——评价、甄别、诊断、教学特征报告
- 年级——评价、甄别、诊断、教研特征报告
- 学校——评价、甄别、诊断、教管特征报告

（二）ICTS 综合素质评价技术

综合素质评价所采用的评价工具，依托于 ICTS 综合素质评价技术，ICTS 是 Intelligence-Competence Test System 首字母缩写，是综合素质能力评价系统（也称综合素养能力评价、基础性发展评价系统等）的英语简称。ICTS 评价技术借鉴了国际上流行的“表现性评价”的方法，借助纸笔测试、行为测验、作品评估三大测量技术，并结合认知测验、行为核准、心理测试等评估手段，对学生的素养水平、基础能力等作出测量、评估和诊断，形成多元、多维、多指标的评价数据报告，从而为中小学生综合素养能力教育提供重要参考信息。

1. 数据处理系统

- 阅卷系统；
- 多元数据处理系统；
- 评价、甄别、诊断分析系统；
- 历时数据处理系统
- 历时评价、甄别、诊断分析系统；
- 决策性评价工具系统；
- 认定性评价工具系统；
- 服务性评价工具系统。

2. 评价报告生成系统

- 学生个体评价报告生成系统；
- 班级评价报告生成系统；
- 年级评价报告生成系统；
- 区域分级评价报告生成系统。
- 班级评价报告分析系统；
- 年级评价报告分析系统；
- 区域分级评价报告分析系统；
- 认定性评价报告分析系统；
- 决策性评价报告分析系统。

3. 评价报告系统

- 学生——评价、甄别、诊断、成长分析报告
- 班级——评价、甄别、诊断、教学特征报告
- 年级——评价、甄别、诊断、教研特征报告
- 学校——评价、甄别、诊断、教管特征报告

(三) PCTS 发展潜能评价技术

发展潜能评价所采用的评价工具，依托于 PCTS 发展潜能评价技术，PCTS 发展潜能评价技术，主要用于评估和预测学生的认知发展潜能情况，PCTS 是 Potential and Competence Test System 的首字母缩写，可用于评估和预测学生的学术潜能、技术潜能情况，该技术以多元智能理论为基础，借鉴了国内外认知、学术、技术潜能测试的研究成果，建立了多元、多维、多指标的潜能评价模型结构，为有效识别学生的发展潜能结构以及后续有针对性的培养提供参考。

1. 数据处理系统

- 阅卷系统；
- 多元数据处理系统；
- 评价、甄别、诊断分析系统；
- 历时数据处理系统
- 历时评价、甄别、诊断分析系统；
- 决策性评价工具系统；
- 认定性评价工具系统；
- 服务性评价工具系统。

2. 评价报告生成系统

- 学生个体评价报告生成系统；
- 班级评价报告生成系统；
- 年级评价报告生成系统；
- 区域分级评价报告生成系统。
- 班级评价报告分析系统；
- 年级评价报告分析系统；
- 区域分级评价报告分析系统；

- 认定性评价报告分析系统；
- 决策性评价报告分析系统。

3. 评价报告系统

- 学生——评价、甄别、诊断、成长分析报告
- 班级——评价、甄别、诊断、教学特征报告
- 年级——评价、甄别、诊断、教研特征报告
- 学校——评价、甄别、诊断、教管特征报告

四、中小学教育质量综合评价项目的报告及案例

中国教育学会可以为项目方提供数据版报告（包含学业评价、发展潜能和综合素质）和文字版报告，项目方可以根据自己的需求选择相应的报告，下面将逐一给出各个报告的样例，其中文字版的报告内容太多，我们将给出部分目录供大家了解。

（一）数据版报告⁴

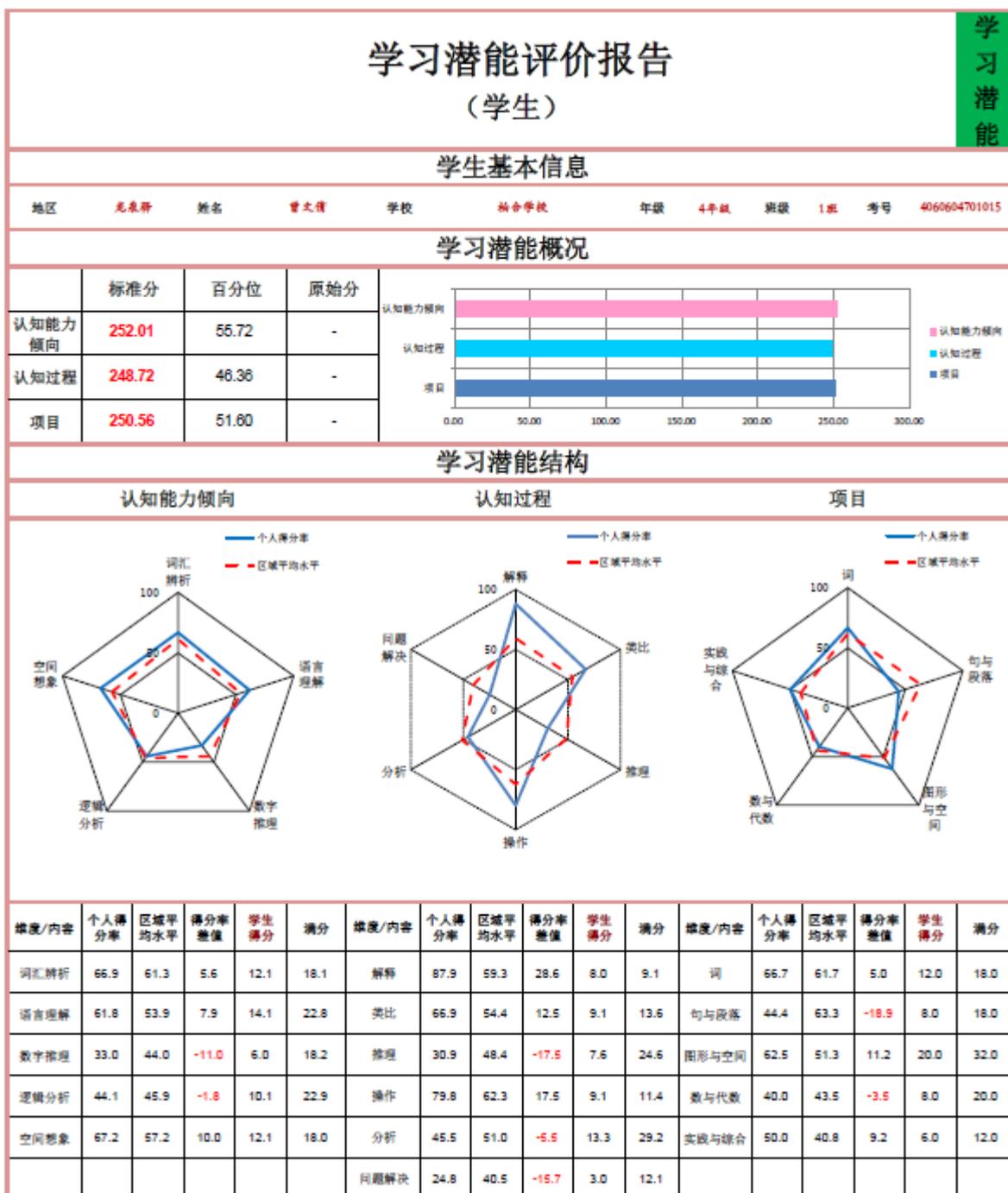
1. 学生评价报告单

⁴备注：由于篇幅原因，各评价报告单只能进行部分呈现，如果想要进一步进行了解，可以与相关负责人进行联系。

示例：学业评价（个人）报告单

ACTS		ACTS 学业评价 学生（语文）评价报告				语 文											
学生基本信息																	
地区	姓名	学校	年级	班级	学号												
总成绩																	
	标准分	百分位	原始分														
知识	240.89	25.76	74.3														
技能	240.76	25.46	73.0														
能力	238.72	21.02	71.0														
总成绩	240.72	25.37	73.5														
	标准分	百分位	原始分														
分项得分																	
知识		技能			能力												
分数构成																	
维度/内容	个人得分率	区域平均水平	得分率差值	学生得分	满分	维度/内容	个人得分率	区域平均水平	得分率差值	学生得分	满分	维度/内容	个人得分率	区域平均水平	得分率差值	学生得分	满分
字词	85.7	81.4	4.3	18.0	21.0	认知	88.8	86.3	2.5	7.1	8.0	词汇辨析	51.1	66.7	-15.6	2.3	4.5
课内句子	80.0	81.3	-1.3	6.0	7.5	书写	100.0	95.0	5.0	1.2	1.2	语言理解	74.7	78.3	-3.6	45.5	60.9
课外句子	100.0	92.5	7.5	8.0	8.0	理解	77.6	77.6	0.0	5.2	6.7	逻辑分析	84.8	79.3	5.5	7.8	9.2
课内语段	66.7	64.4	2.3	3.0	4.5	信息提取	57.7	61.5	-3.8	1.5	2.6	人际理解	50.0	78.0	-28.0	2.5	5.0
课外语段	44.4	64.4	-20.0	4.0	9.0	综合概括	76.9	64.1	12.8	3.0	3.9	自我认识	63.2	72.5	-9.3	12.9	20.4
课内文章	50.0	75.0	-25.0	4.0	8.0	推理	30.2	76.7	-46.5	1.3	4.3						
课外文章	62.5	75.0	-12.5	5.0	8.0	应用	91.3	85.6	5.7	9.5	10.4						
作文	76.0	75.3	0.7	22.8	30.0	分析	91.8	88.5	3.3	5.6	6.1						
文学常识	87.5	92.8	-5.3	3.5	4.0	表达	68.0	73.6	-5.6	38.6	56.8						
小计	74.3	77.6	-3.3	74.3	100.0	小计	73.0	76.7	-3.7	73.0	100.0	小计	71.0	76.6	-5.6	71.0	100.0

示例：学习潜能评价（个人）报告单



示例：综合素质评价（个人）报告单

综合素质评价报告 (学生)							综合素质
学生基本信息							
地区	姓名	学校	年级	班级	考号		
得分情况							
品德发展水平							
维度/内容	个人得分率	区域平均水平	得分率差值	标准9分	学生得分	满分	
行为习惯	98.7	87.1	11.6	7	74.0	75.0	
公民素养	99.0	89.6	9.4	7	99.0	100.0	
人格品质	88.6	86.4	2.2	6	62.0	70.0	
理想信念	86.7	86.9	0.2	6	66.0	75.0	
学业发展水平							
维度/内容	个人得分率	区域平均水平	得分率差值	标准9分	学生得分	满分	
实践能力	80.0	83.2	-3.2	6	40.0	50.0	
创新意识	78.0	76.1	1.9	6	78.0	100.0	
身心发展水平							
维度/内容	个人得分率	区域平均水平	得分率差值	标准9分	学生得分	满分	
健康生活方式	84.0	86.6	-2.6	6	42.0	50.0	
行为调控	76.0	82.3	-6.3	4	57.0	75.0	
人际沟通	80.0	77.1	2.9	6	60.0	75.0	
考察内容							
行为习惯	学生在文明礼貌、勤俭节约、热爱劳动、保护环境等方面的表现情况。						
公民素养	学生在珍爱生命、遵纪守法、诚实守信、团结友善、乐于助人等方面的表现情况。						
人格品质	学生在自尊自信、自律自强、尊重他人、乐观向上等方面的表现情况。						
理想信念	学生的爱国情感、民族认同、社会责任、集体意识等方面的情况。						
实践能力	学生关注现实生活、参加社会实践和志愿服务活动、解决实际问题的情况。						
创新意识	学生好奇心、求知欲和问题意识等方面的情况。						
健康生活方式	学生的生活与卫生习惯、参加课外文体体育活动等方面的情况。						
行为调控	学生对自己行为的自我约束情况，应对和克服学习、生活中遇到的困难的表现情况。						
人际沟通	师生关系、同伴关系等方面的情况。						

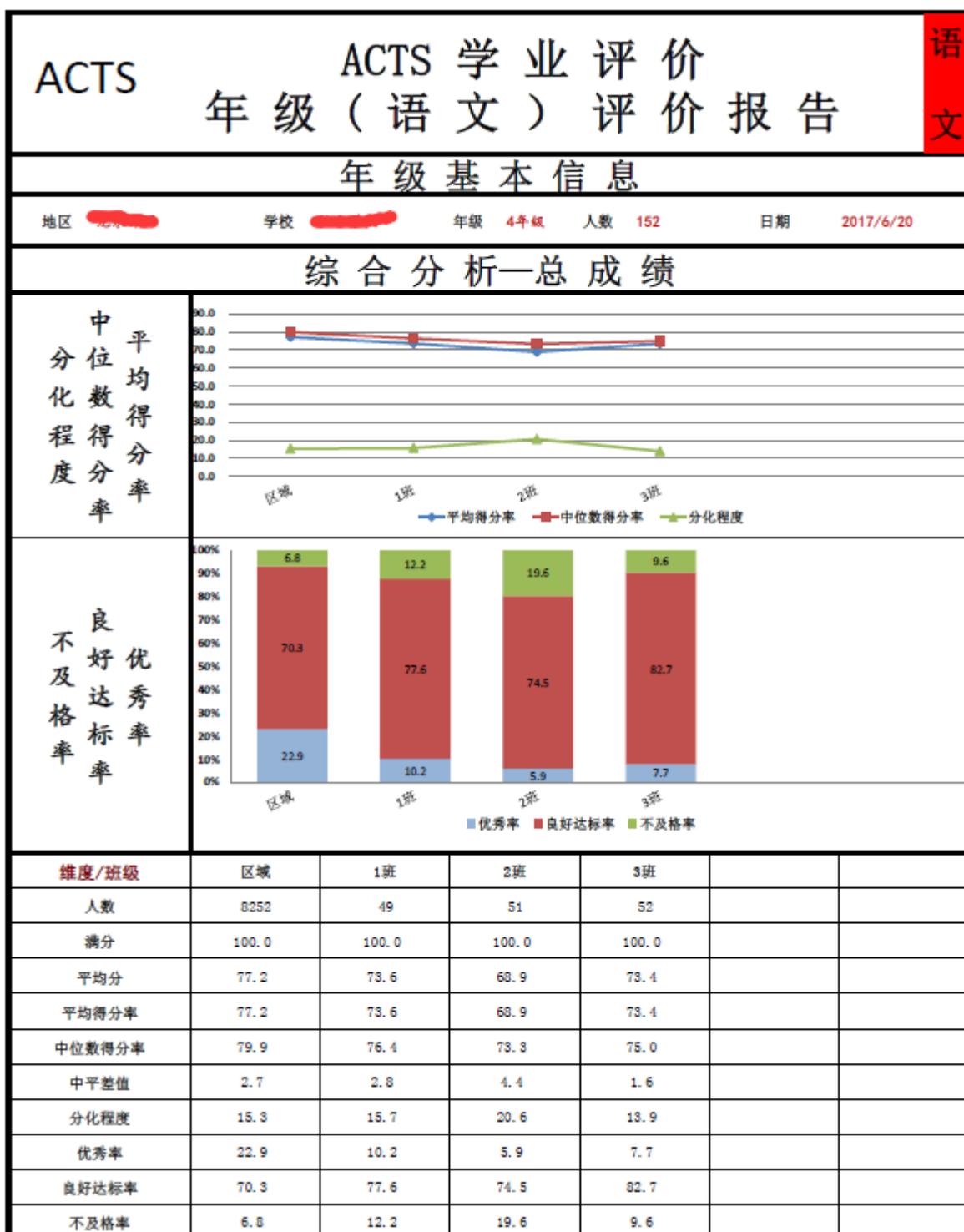
2. 班级评价报告单

示例：学业评价（班级）报告单

ACTS 学业评价						语文					
班级（语文）评价报告											
班级基本信息											
地区	龙东地区	学校	站前学校	年级	4年级	班级	1班	人数	49	考试日期	2017/6/20
知识分数构成											
维度（指标）/内容	班级平均分	区域平均分	差值	班级平均分	满分						
字词	73.8	81.4	-7.6	15.5	21.0						
课内句子	74.7	81.3	-6.6	5.6	7.5						
课外句子	90.0	92.5	-2.5	7.2	8.0						
课内语段	51.1	64.4	-13.3	2.3	4.5						
课外语段	58.9	64.4	-5.5	5.3	9.0						
课内文章	66.3	75.0	-8.7	5.3	8.0						
课外文章	70.0	75.0	-5.0	5.6	8.0						
作文	77.7	75.3	2.4	23.3	30.0						
文学常识	92.3	92.8	-0.5	3.69	4.0						
小计	73.8	77.8	-3.8	73.8	100.0						
维度评语											
1、“字词”成绩表明通过上一阶段的学习，本班学生对字词音、形、义的认识、理解、应用达到了						中等	水平。				
2、“课内句子”成绩表明通过上一阶段的学习，本班学生对课内学习过的句子的意义、结构、关系以及句中标点符号的认识、理解、分析达到了						中等	水平。				
3、“课外句子”成绩表明通过上一阶段的学习，本班学生对课外句子的意义、结构、关系以及句中标点符号的认识、理解、分析达到了						优秀	水平。				
4、“课内语段”成绩表明通过上一阶段的学习，本班学生对课内学习过的语段的意义、关系、结构的理解和分析达到了						未达标	水平。				
5、“课外语段”成绩表明通过上一阶段的学习，本班学生对课外语段的意义、关系、结构的理解和分析达到了						未达标	水平。				
6、“课内文章”成绩表明通过上一阶段的学习，本班学生对课内学习过的文章的体裁、意义、结构、写作手法、观点、所表达情感等的理解、分析、鉴赏达到了						达标	水平。				
7、“课外文章”成绩表明通过上一阶段的学习，本班学生对课外文章的体裁、意义、结构、写作手法、观点、所表达情感等的理解、分析、鉴赏达到了						中等	水平。				
8、“作文”成绩表明通过上一阶段的学习，本班学生根据题目要求进行创造性语言表达的能力和水平达到了						中等	水平。				
9、“文学常识”成绩表明通过上一阶段的学习，本班学生对著名作家、年代、作品，文学中的地理、历史典故，众所周知的文学基本知识等的认知达到了						优秀	水平。				
综合评语											
1、在知识中，本班学生得分率最高的项是 文学常识 ，最低的项是 课内语段											
表明在上一阶段的学习中，本班学生在 文学常识 取得了较好的学习效果，但在 课内语段 方面的学习却相对存在不足。											
造成这种不足的原因有两个：如果该项指标的差值为正，则表明这种不足是普遍性的，不是本班的特有问题；如果是差值为负，则表明不足是本班的特有问题，需要引起教师关注。											
2、在知识中，本班高于区域平均水平的项是 作文 ，低于区域平均水平的项是 字词 课内句子 课外句子 课内语段 课外语段 课内文章 课外文章 文学常识 表明在这次考试中，本班学生总体的优势是 作文 ，本班学生总体的劣势是 字词 课内句子 课外句子 课内语段 课外语段 课内文章 课外文章 文学常识 。											
优秀率、良好达标率、不及格率											
1、在知识方面，本班的平均分是 73.8 ，在本次考试中，达到了 中等 水平。与区域平均水平相比， 低 于区域平均水平。											
2、在知识方面，本班的优秀率是 14.3 ，表明在本次考试中，知识得分率在 85 以上的学生人数占本班全部考试人数的 14.3% 与区域优秀率相比， 低 于区域的优秀率。											
3、在知识方面，本班的良好达标率是 71.4 ，表明在本次考试中，知识得分率在 60-85 之间的学生人数占本班全部考试人数的 71.4% 与区域良好达标率相比， 高 于区域的良好达标率。											
4、在知识方面，本班的不及格率是 14.3 ，表明在本次考试中，知识得分率在 60 以下的学生人数占本班全部考试人数的 14.3% 与区域不及格率相比， 高 于区域的不及格率。											

3. 年级评价报告单

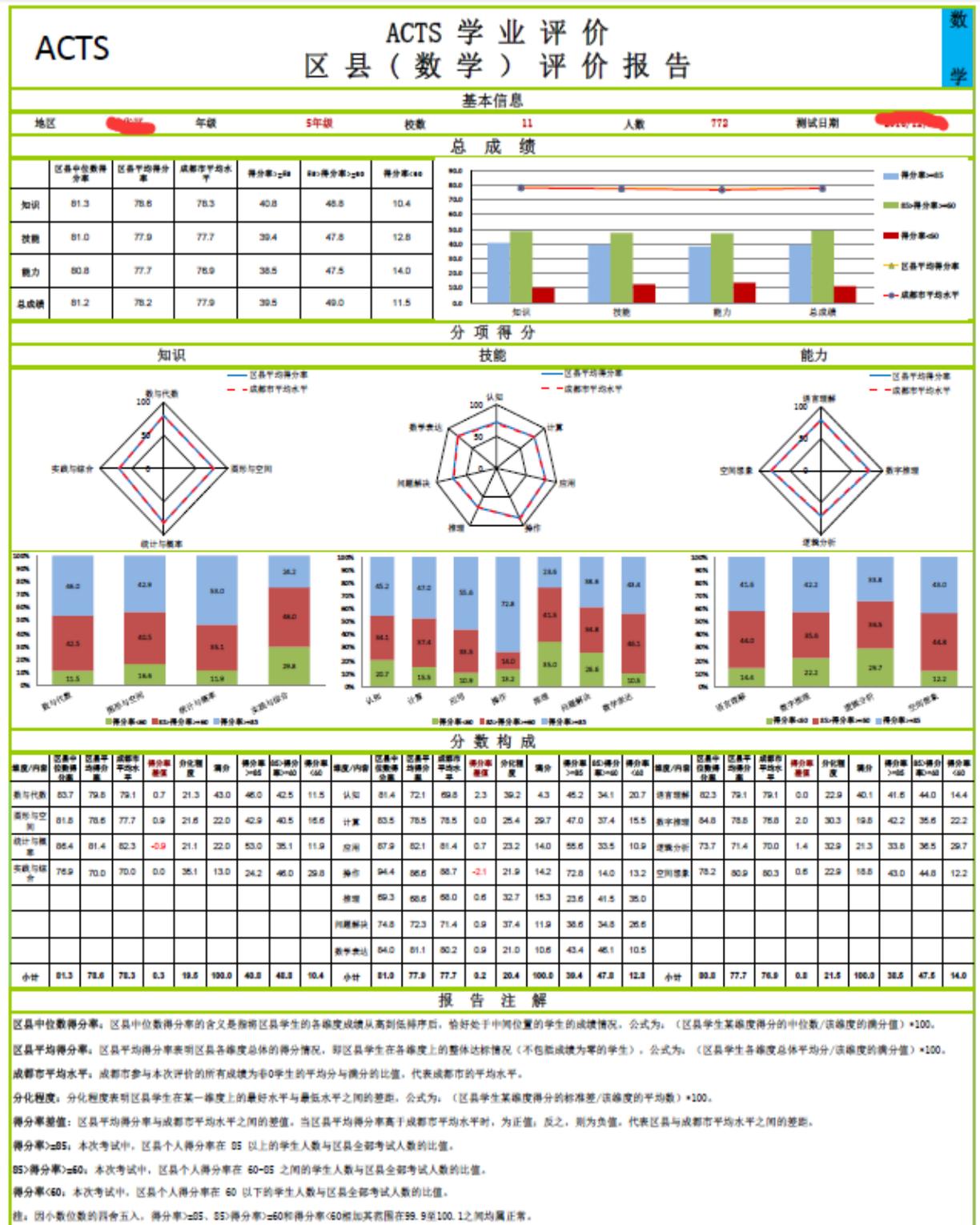
示例：学业评价（年级）报告单



4. 区域评价报告单

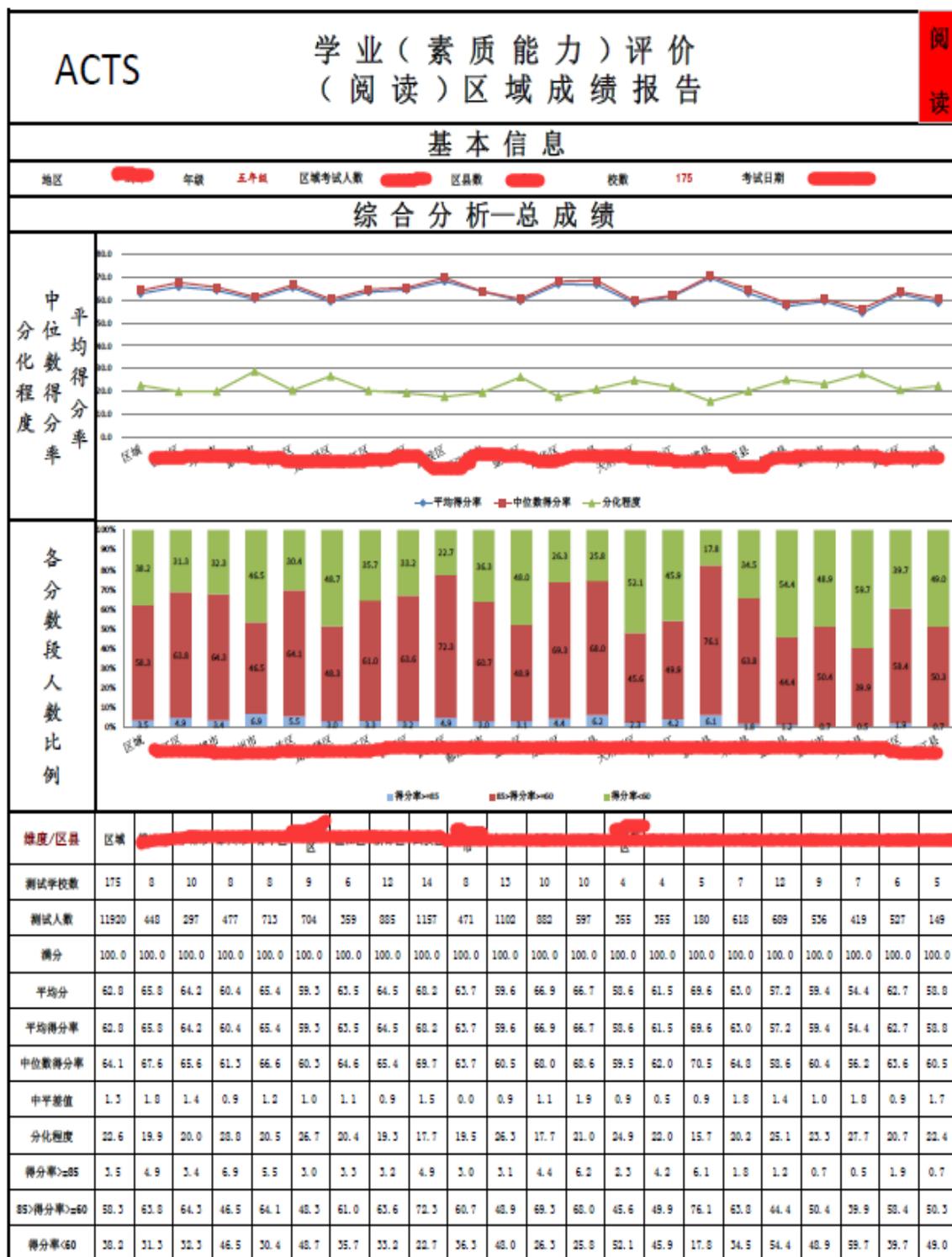
(1) 区县级

示例：学业评价（区域）报告单



(2) 市级

示例：学业评价（区域）报告单



(二) 文字版报告

1. 学业分析报告

第一篇 小学学业评价（数学）项目介绍

- 一、小学学业评价（数学）的背景及目的
- 二、小学学业评价（数学）采用的模式和指标
- 三、小学学业评价（数学）的工具说明
- 四、小学学业评价（数学）抽样说明

第二篇 小学学业评价（数学）结果分析

- 一、某市小学学业评价（数学）概况
- 二、某市各区（市）县小学学业评价（数学）各指标分布概况
- 三、某市各区（市）县小学学业评价（数学）分组分析
- 四、不同学校类型小学学业评价（数学）各指标概况及差异分析
- 五、不同社会经济背景学生小学学业评价（数学）各指标概况及差异分析
- 六、其他因素区分不同类别学生小学学业评价（数学）各指标概况及差异分析
- 七、学生按总成绩聚类分析结果

第三篇 教师、校长行为对学业成绩影响的结果分析

- 一、某市小学教师问卷指标概况
- 二、某市小学校长问卷指标概况

第四篇 小学学业评价（数学）与学校问卷各指标之间的相关分析及回归分析

- 一、小学学业评价（数学）指标与学校问卷指标的相关分析
- 二、小学学业评价（数学）指标与学校问卷指标的回归分析

第五篇 评价结果应用

- 一、发现与结论
- 二、建议

附表

2. 发展潜能报告

第一篇 项目说明

- 一、项目介绍
- 二、项目背景
- 三、项目开展
- （一）政策依据

(二) 科学依据

(三) 现实依据

四、项目目的

第二篇 发展潜能评价采用的指标

第三篇 发展潜能评价的测量工具和评价工具

一、测量工具说明

(一) 测量工具的编制流程

(二) 测量工具的内容结构

(三) 测量工具的效度

(四) 测量工具的信度

(五) 测量工具的区分度

二、评价工具说明

第四篇 抽样及实测说明

一、学校抽样

二、学生抽样

三、抽样结果

第五篇 测量评价结果分析

一、某市 5 年级发展潜能测量评价的总体特征描述

二、某市各区(市)县 5 年级发展潜能评价的特征描述

三、某市各区(市)县按 5 年级发展潜能评价分组分析

(一) 各区(市)县按发展潜能评价的聚类分析结果

(二) 不同组别学生发展潜能评价各指标的表现水平

四、不同学校类型学生发展潜能评价各指标的基本概况及差异分析

(一) 以城乡分布区分不同学校类型

(二) 以学校学段区分不同学校类型

五、不同社会经济背景学生发展潜能评价各指标的基本概况及差异分析

(一) 父、母亲不同受教育程度学生发展潜能评价各指标概况及差异分析

(二) 父、母亲不同职业类型学生发展潜能评价各指标的概况及差异分析

六、不同性别学生发展潜能评价各指标的概况及差异分析

七、其他因素不同类别学生发展潜能评价各指标的概况及差异分析

(一) 学生业余时间被占用不同情况学生发展潜能评价各指标的概况及差异分析

(二) 学校体育设置不同情况学生群体发展潜能评价各指标的概况及差异分析

(三) 不同学校班级规模学生群体发展潜能评价各指标的概况及差异分析

八、发现与结论

九、建议

附表

3. 综合素质报告

第一篇 综合素质评价项目介绍

一、某市综合素质评价的背景及目的

(一) 综合素质评价的背景

(二) 综合素质评价的目的

二、某市综合素质评价采用的指标

(一) 综合素质评价采用的统计指标

(二) 综合素质评价采用的内容指标

三、某市综合素质评价的测量工具说明

(一) 测量工具的编制依据

(二) 测量内容的结构

(三) 测量工具的信度

(四) 测量工具的效度

四、某市综合素质评价的评价工具说明

五、参测学校名单及人数说明

第二篇 某市综合素质评价结果分析

一、某市品德发展水平评价结果

(一) 品德发展水平特征概况

(二) 学校聚类分析结果及各组别品德发展水平概况

(三) 班级聚类分析结果及各组别品德发展水平概况

(四) 不同类型学校学生品德发展水平各指标概况及差异分析

(五) 不同性别学生品德发展水平各指标概况及差异分析

二、某市实践能力、创新意识评价结果

(一) 实践能力、创新意识特征概况

(二) 学校聚类分析结果及各组别实践能力、创新意识概况

(三) 班级聚类分析结果及各组别实践能力、创新意识概况

(四) 不同类型学校学生实践能力、创新意识指标概况及差异分析

(五) 不同性别学生实践能力、创新意识指标概况及差异分析

三、某市身心发展水平评价结果

(一) 身心发展水平总体概况

(二) 学校聚类分析结果及各组别身心发展水平概况

- (三) 班级聚类分析结果及各组别身心发展水平概况
- (四) 不同类型学校学生身心发展水平指标概况及差异分析
- (五) 不同性别学生身心发展水平指标概况及差异分析

第三篇 发现与建议

一、发现与结论

二、建议

(一) 决策管理建议

(二) 教学育人建议

附表