



# 《中国英语能力等级量表》 及其在语言测评中的应用

何莲珍

2019年7月28日 苏州



# 《中国英语能力等级量表》 解读

《实施意见》提出加强“外语能力测评体系建设”，第一次从国家层面对我国外语教学和考试综合改革提出了明确要求。

加强统筹谋划，积极稳妥推进。整体设计从基础教育到高等教育考试招生制度改革，促进普通教育、职业教育、继续教育之间衔接沟通，统筹实施考试、招生和管理制度综合改革，试点先行，稳步推进。

4. 深化高考考试内容改革。依据高校人才选拔要求和国家课程标准，科学设计命题内容，增强基础性、综合性，着重考查学生独立思考和运用所学知识分析问题、解决问题的能力。改进评分方式，加强评卷管理，完善成绩报告。加强国家教育考试机构、国家题库和外语能力测评体系建设。2015年起增加使用全国统一命题试卷的省份。

# 国家外语能力测评体系



## 外语能力测评体系的目标定位

**3**个覆盖

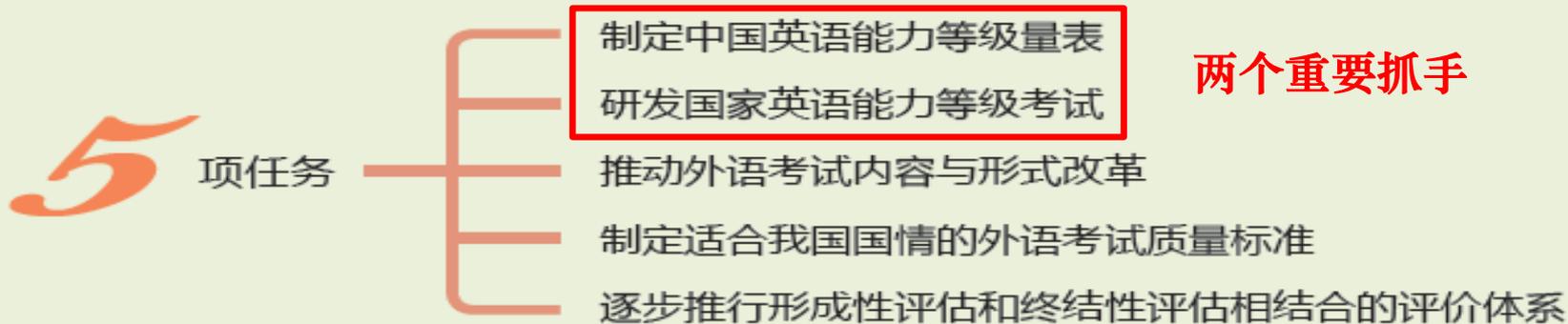
- 覆盖大中小学各教育阶段
- 覆盖听说读写译综合能力
- 覆盖外语学习及教学与测评的评价系统



# 外语能力测评体系建设目标及任务



**建设目标：**根据《实施意见》总体目标要求，2020年建立标准统一、功能多元的外语能力测评体系，促进各种学习成果的沟通互认，促进各级各类外语教育教学与考试的有序衔接和协调发展。



# 语 言 文 字 规 范

(GF 0018—2018)

## 中国英语能力等级量表

China's Standards of English Language Ability

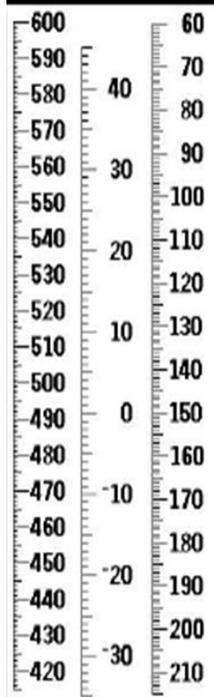
(2018-01 发布)

(2018-06-01 实施)

中华人民共和国教育部  
国家语言文字工作委员会

发 布

## China's Standards of English Language Ability

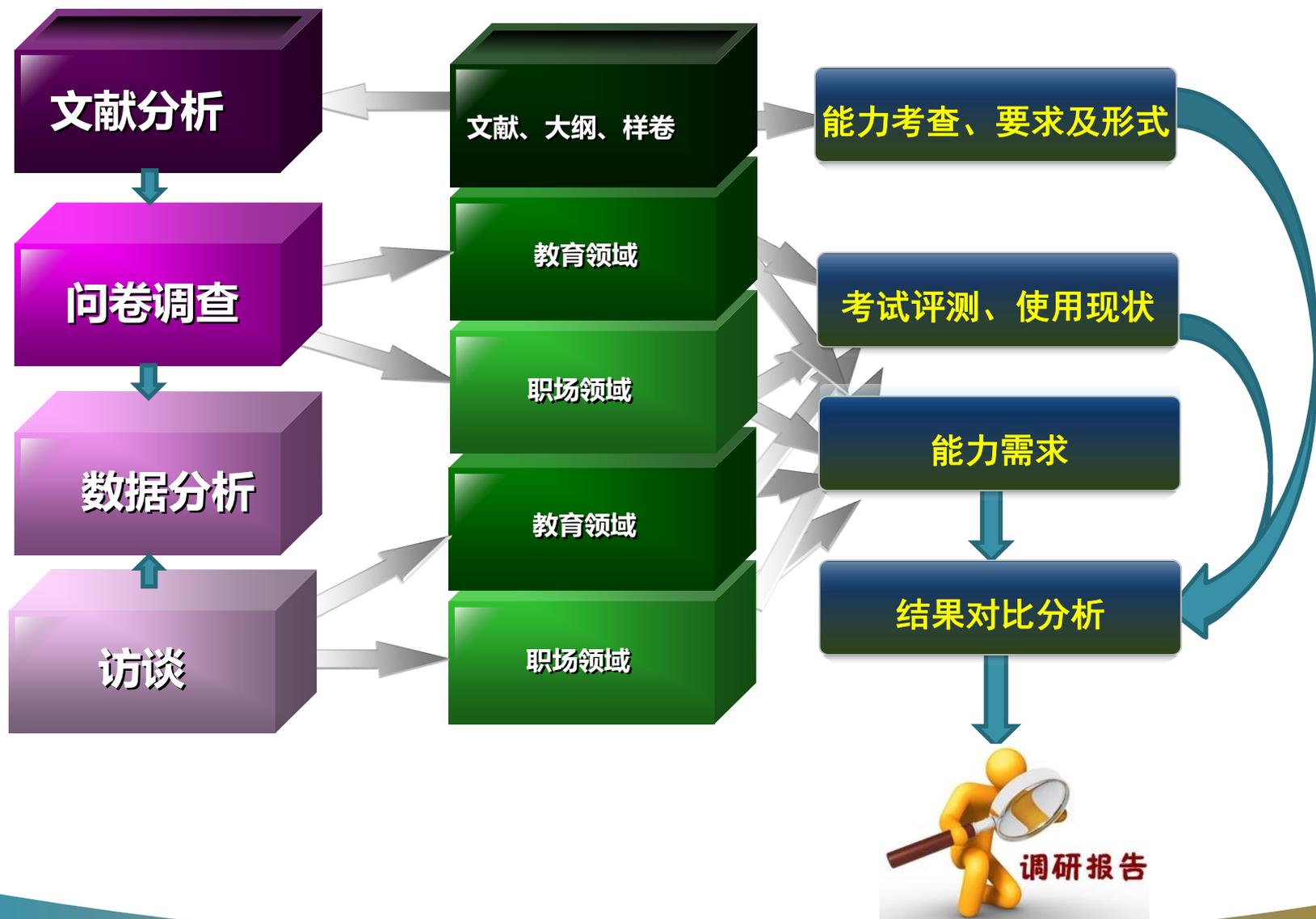


# 制定中国英语能力等级量表的基础

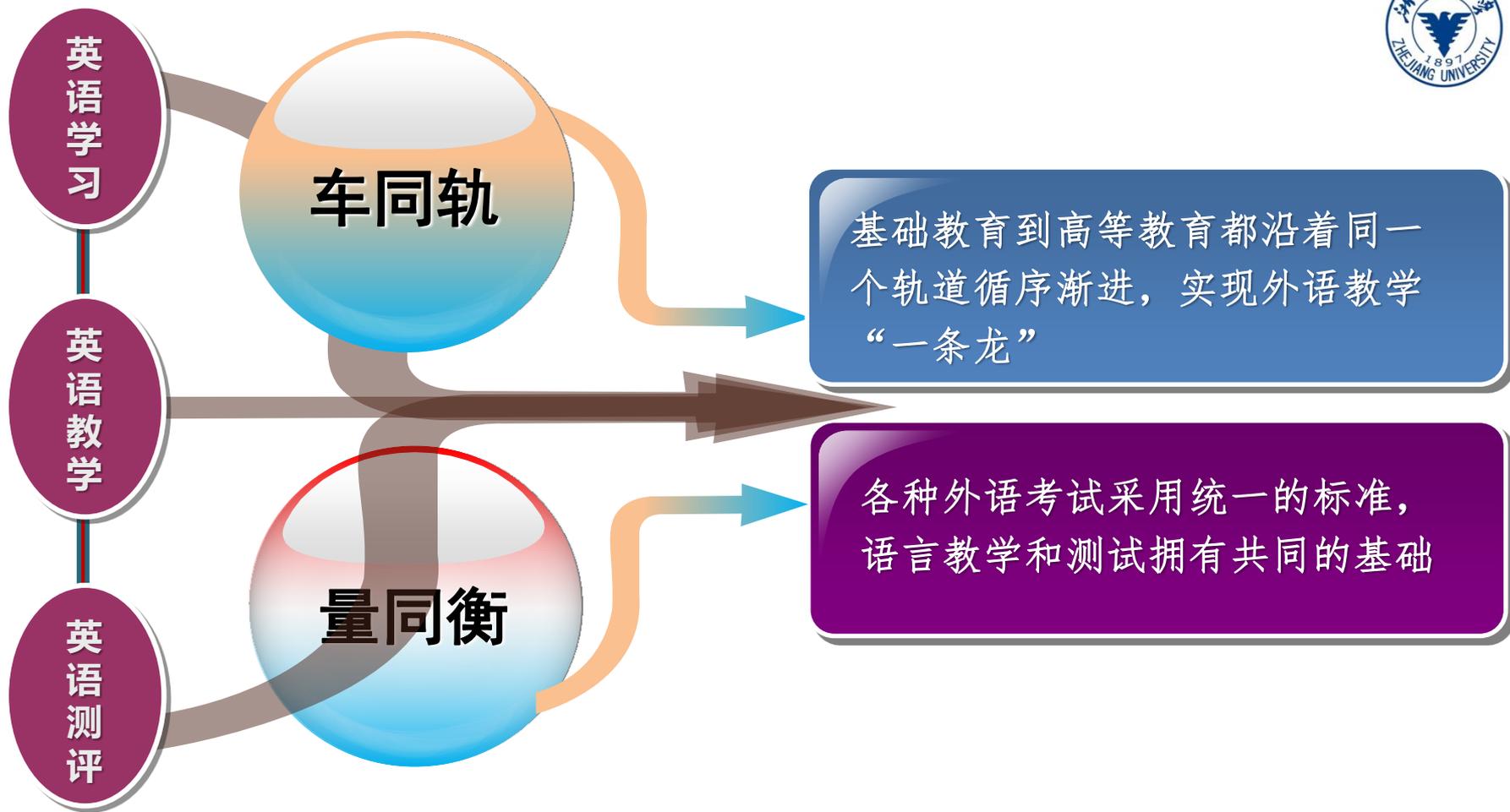


- 理论支撑
  - 美国FSI(Foreign Service Institute)量表
  - 美国ILRS( Interagency Language Roundtable Scale)量表
  - 美国ACTFL( the American Council on the Teaching of Foreign Languages)量表
  - 欧洲的ALTE(Association of Language Testers in Europe)量表
  - 欧洲CEFR (Common European Framework of Reference)量表
  - Pearson公司开发的GSE (Global Scale of English)量表
  - 加拿大CLB( the Canadian Language Benchmarks)量表
  - 澳大利亚的ISLPR( International Second Language Proficiency Ratings)量表
- 这些国外的量表不能照搬到我国的量表制定中，但他们采用的理论依据、制定原则和方法值得借鉴。

# 外语能力、教学、测评现状及社会需求调研报告

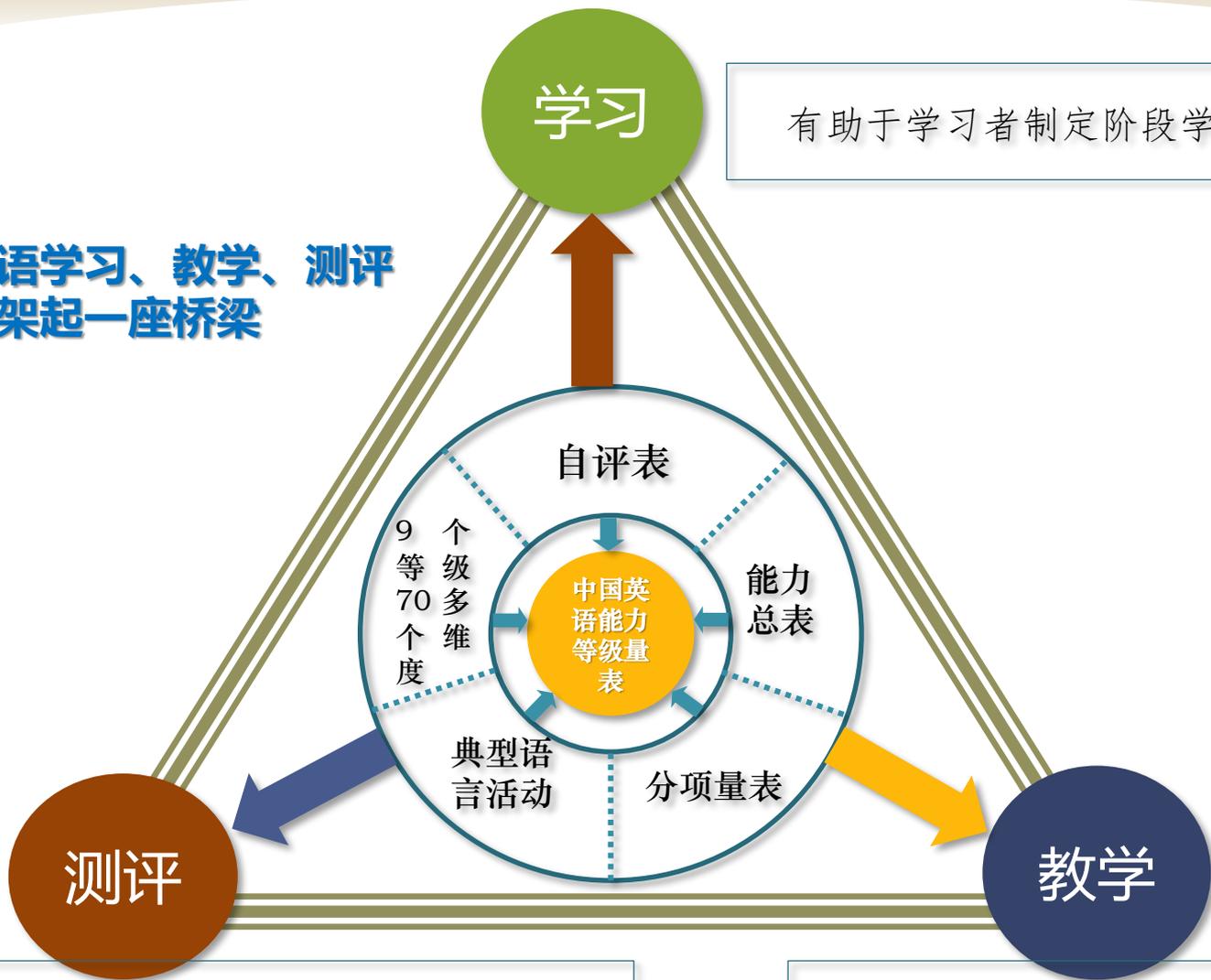


# 中国英语能力等级量表的作用



# 中国英语能力等级量表的应用

在英语学习、教学、测评之间架起一座桥梁



有助于学习者制定阶段学习目标

有利于学习者阶段性的语言能力得到承认

有助于加强综合能力培养、因材施教



# 《中国英语能力等级量表》 在语言测评中的应用

# CSE v.s. CEFR

理论框架

听力理解  
能力

识别/提取

在收听、观看语速较慢、话题熟悉的广播影视节目时，能识别其主题，获取主要信息。

概括/分析

能听懂语速正常、有关社会问题的演讲，分清观点与事实。

批判/评价

能听懂有关政治、经济、历史、文化等抽象话题的论述，评价说话者的观点与立场。



# CSE v.s. CEFR

描述语  
成分

“能做” (Can Do) 描述语的三个组成部分:



描述语 = 行为 + 标准 + 条件

Descriptor = Performance + Criteria + Condition

描述语 = 能做什么事 + 做到什么程度 + 在什么情况下做

描述语	行为	标准	条件
能听懂语速正常、有关社会问题的演讲，分清观点与事实。	分清观点与事实	语速正常、有关社会问题	\

# CSE 典型特征

描述语的  
系统性

级别	语速(Criteria)
9	语速较快 (140-180 词/分)
8	语速较快 (140-180 词/分)
7	语速较快 (140-180 词/分)
6	语速正常 (100-140词/分)
5	语速正常 (100-140词/分)
4	语速正常 (100-140词/分)
3	语速较慢 (80-100词/分)
2	语速缓慢 (60-80词/分)
1	语速缓慢 (60-80词/分)



# CSE 典型特征



级别	语音特征(Criteria)
9	带有口音
8	带有口音
7	\
6	语音标准
5	语音标准; 发音清晰
4	语音标准; 发音清晰
3	语音标准; 发音清晰
2	语音标准; 发音清晰
1	发音清晰

# CSE 典型特征



级别	语言与内容(Criteria)
9	用词生僻; 含有不常见俗语或行话; 用词复杂
8	含有多种英语变体; 俚语较多; 含有部分术语; 长篇
7	包含俚语或习语; 含有双关语、隐喻等语言现象; 带有专业术语
6	信息量大; 细微; 复杂
5	复杂
4	简单; 简短
3	简单; 简短
2	词汇常见; 简短; 简单
1	用词简单; 简单; 常见

# CSE 典型特征



级别	话题 (Criteria)
9	各种话题
8	广泛话题; 相近领域的专业话语
7	与自己专业领域相关的; 抽象话题; 有关公共政策与社会问题
6	与自己专业领域相关的; 实事
5	一般性话题; 话题熟悉; 有关社会问题的; 有关学习及工作; 科普类
4	一般性话题; 与个人兴趣相关; 话题熟悉
3	话题熟悉
2	有关日常生活
1	有关日常生活

# CSE 典型特征



级别	认知处理 (performance)
9	进行分析、推断与评价; 理解其中的典故; 理解其涉及的社会历史文化内涵
8	评价各方观点的合理性与逻辑性
7	评价演讲者的观点与立场; 概括主要内容; 理解说话者话语的隐含意义; 理解其涉及的社会文化内涵; 评判其中的主要观点
6	概括主要内容; 判断报道之间的异同; 判断说话者用语的得体性
5	概括大意; 获取要点和细节
4	理解大意; 获取主要的事实性信息; 理解说话者的观点和意图; 理解说话者的言外之意
3	获取关键信息; 识别其主题
2	获取人物、地点、事件等信息
1	识别有关自己、家庭、学校的词汇和短语

# CSE 典型特征



级别	条件 (Condition)
9	
8	
7	基本无需对方重复或解释
6	环境嘈杂; 面对面
5	
4	
3	
2	在图片、图像、手势等的帮助下
1	借助语音、语调或手势、表情等

# 认知诊断测试



认知诊断  
测试

可观察的  
表现、反应模  
式

不可观察的  
认知结构和加  
工技能

教、学、评

粒化认知属性  
(fine-grained attribute)

认知诊断  
测试相关  
研究

- 技术：模型开发与比较
- 应用：  
“翻新法”：分析非诊断语言测试  
“归纳法”：设计诊断语言测试

# 研究目的



以一项校本考试——**浙江大学英语水平考试**为例，综合使用**考试对接量表**方法和**认知诊断模型**分析，探讨量表作为对英语考试**衔接定位**、对英语能力进行**诊断和个性化反馈**的共同标尺，如何实现真正意义上的“量同衡”。

# 研究工具



- 考卷: 8-10套/年
- 锚题率: 约30%



# 研究方法



## 个性化能力描述语

### 1) 总体描述（能听懂什么样的材料）

主要参考CSE5描述语：考试与量表对接

### 2) 个性化的听力技能掌握情况描述

听力技能：认知诊断分析

# 研究方法



学生口头报告

专家判断

认知诊断分析

内容分析

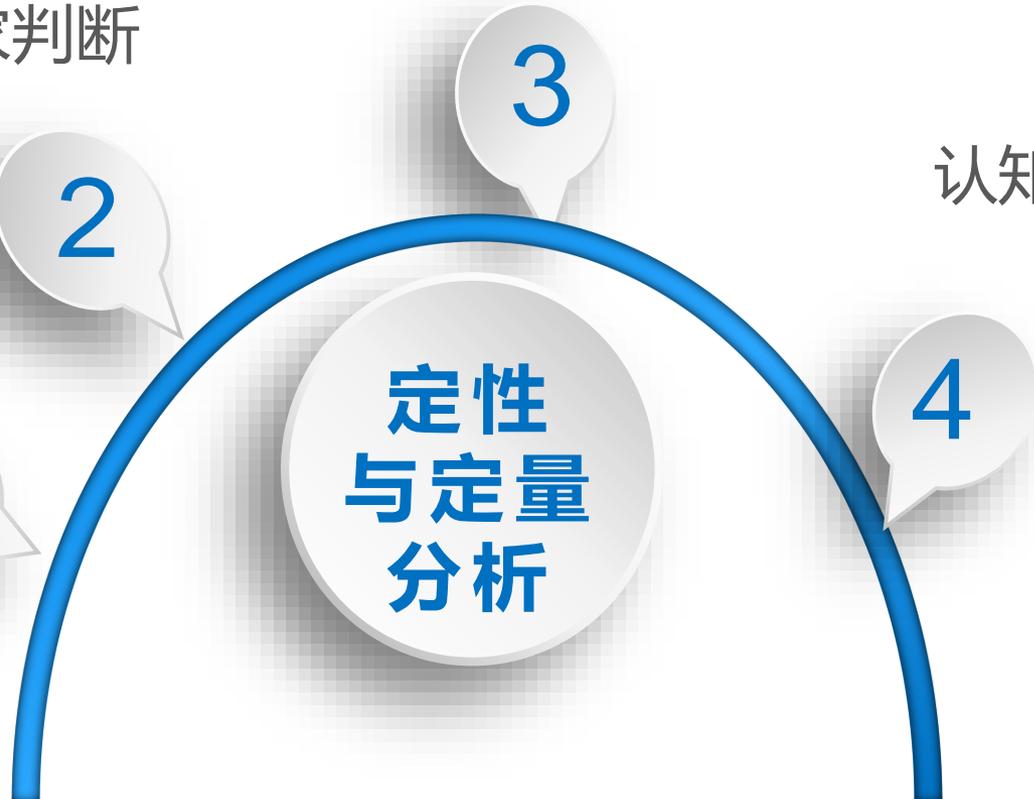
定性  
与定量  
分析

1

2

3

4



# 研究步骤

对接



熟悉阶段

01

02

考试说明阶段

标准设定阶段

03

➤ 方法一：basket method

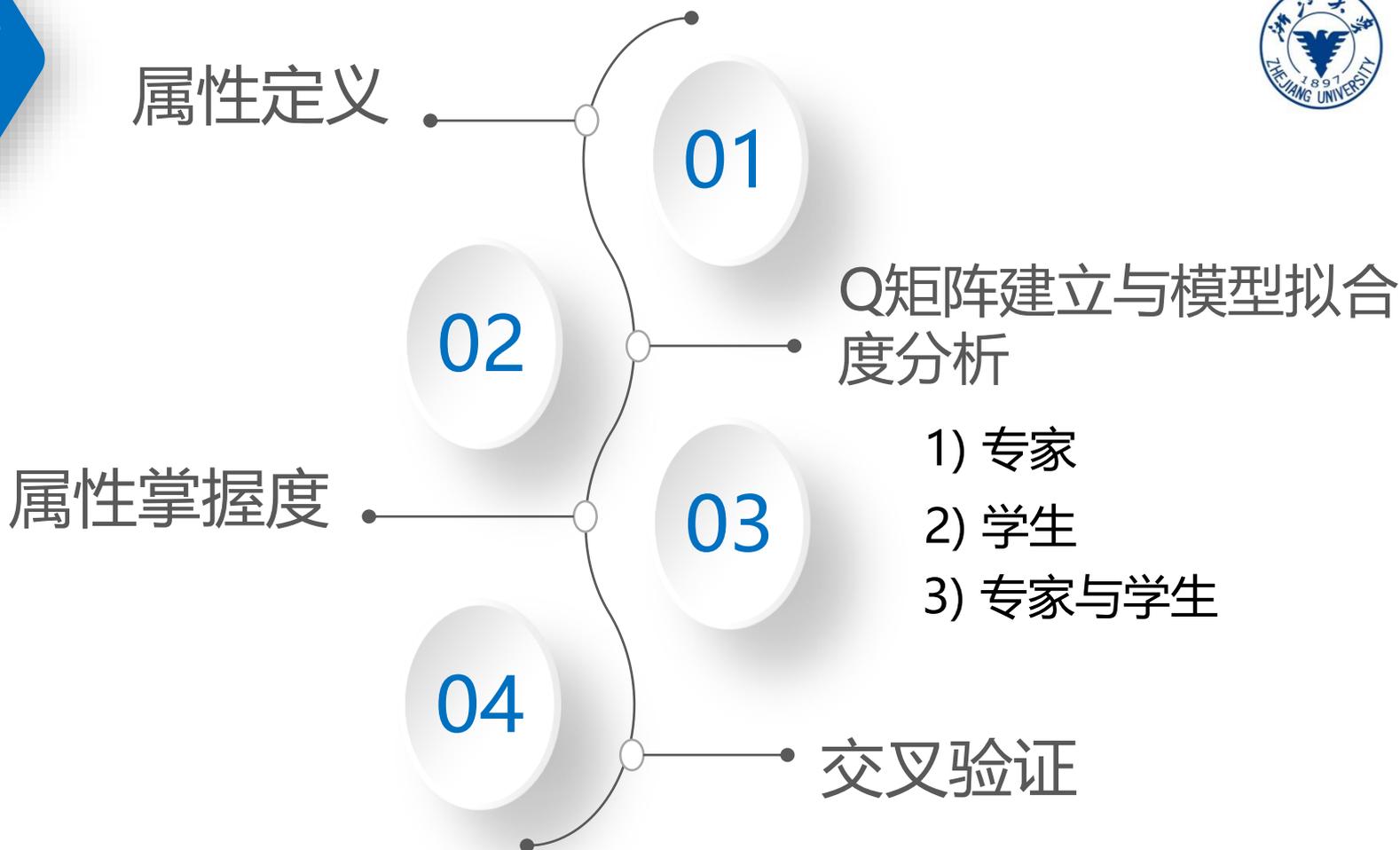
➤ 方法二：modified Angoff method

04

效度验证阶段

# 研究步骤

认知诊断  
分析



# 研究步骤



## 属性定义

### 1) 文献梳理

- 研究者（如Buck, Carroll, Oakeshott-Taylor, Richards, Munby等）对听力技能的划分

### 2) 内容分析

- CSE3~7级听力描述语
- 浙江大学英语水平考试大纲

### 3) 专家与学生判断

- 考试试题中涉及的听力技能

**(A1) 细节信息提取**

**(A2) 主旨大意归纳**

**(A3) 推断和引申**

**(A4) 信息综合处理**

# 研究步骤

## Q矩阵建立



題目	A1细节	A2主旨	A3推断	A4综合
1	1	0	0	1
2	0	0	0	1
3	0	1	0	1
4	1	0	1	
...	...	...	...	...
27	0	0	1	1
28	0	1	0	0
29	0	0	0	1
30	0	1	0	1

对接

## 1. 总体描述（能听懂什么样的材料）

① 分级： 0-7分： 无法听懂

8-14分： CSE-3 听~~~时有一定困难

15-19分： CSE-4 能基本听懂

20-24分： CSE-5 能听懂

25-30分： CSE-6 能完全听懂

② 描述语： 主要参考CSE5级

对接

## 1. 总体描述（能听懂什么样的材料）

非专业讲座、  
演讲、对话  
等

- 0-7分：无法听懂语速正常、有关一般性话题的口头表达。
- 8-14分：在听语速正常、有关一般性话题的口头表达时有一定困难。
- 15-19分：（达到了CSE4级），能基本听懂语速正常、有关一般性话题的口头表达。
- 20-24分：（达到了CSE5级），能听懂语速正常、有关一般性话题的口头表达。
- 25-30分：（达到了CSE6级），能完全听懂语速正常、有关一般性话题的口头表达。



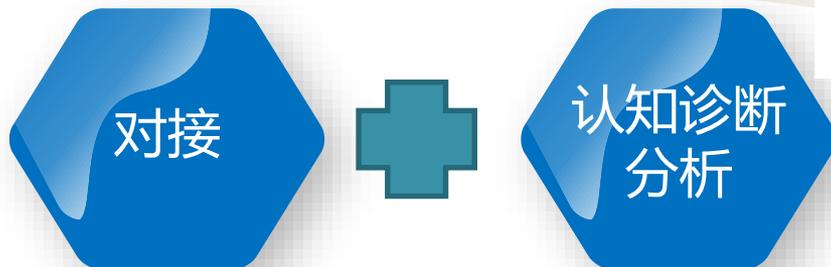
## 2. 个性化的听力技能掌握情况描述

技能掌握模式 (共16种)	描述语
0000 (9.5%)	<ul style="list-style-type: none"><li>① 在细节信息提取上有缺陷, 未能理解并提取出特定的细节信息, 如人物、时间、地点、事件的原因、过程、结果等;</li><li>② 在主旨大意归纳方面表现欠佳, 未能获取说话者或文段中心思想;</li><li>③ 推断能力有限, 未能根据已知信息推断出相关内容 (如: 事件的原因、过程、结果或说话者的意图、态度等);</li><li>④ 信息综合处理能力不足, 未能综合分析上下文或相关信息, 把握说话者的信息组织方式, 明确文段逻辑关系。</li></ul>
1000	<ul style="list-style-type: none"><li>① 细节信息提取能力良好, 能理解并提取出特定的细节信息, 如人物、时间、地点、事件的原因、过程、结果等;</li><li>② 在主旨大意归纳方面表现欠佳, 未能获取说话者或文段中心思想;</li><li>③ 推断能力有限, 未能根据已知信息推断出相关内容 (如: 事件的原因、过程、结果或说话者的意图、态度等);</li><li>④ 信息综合处理能力不足, 未能综合分析上下文或相关信息, 把握说话者的信息组织方式, 明确文段逻辑关系。</li></ul>

## 2. 个性化的听力技能掌握情况描述

技能掌握模式 (共16种)	描述语
0111	<ul style="list-style-type: none"><li>① 在主旨大意归纳方面表现良好，能理解中心思想；</li><li>② 推断能力掌握较好，能根据已知信息推断出相关内容（如：事件的原因、过程、结果或说话者的意图、态度等）；</li><li>③ 信息综合处理能力良好，能综合分析上下文或相关信息，把握说话者的信息组织方式，明确文段逻辑关系；</li><li>④ 在细节信息提取上有缺陷，未能理解并提取出特定的细节信息，如人物、时间、地点、事件的原因、过程、结果等。</li></ul>
1111 (26.6%)	<ul style="list-style-type: none"><li>① 细节信息提取能力良好，能理解并提取出特定的细节信息，如人物、时间、地点、事件的原因、过程、结果等；</li><li>② 在主旨大意归纳方面表现良好，能理解中心思想；</li><li>③ 推断能力掌握较好，能根据已知信息推断出相关内容（如：事件的原因、过程、结果或说话者的意图、态度等）；</li><li>④ 信息综合处理能力良好，能综合分析上下文或相关信息，把握说话者的信息组织方式，明确文段逻辑关系。</li></ul>

## 分数报告示例



在听力部分得分为22分，技能掌握模式为1011的考生将会得到如下报告：

能听懂语速正常、有关一般性话题的口头表达。具体听力技能掌握情况如下：

- ① 细节信息提取能力良好，能理解并提取出特定的细节信息，如人物、时间、地点、事件的原因、过程、结果等；
- ② 推断能力掌握较好，能根据已知信息推断出相关内容（如：事件的原因、过程、结果或说话者的意图、态度等）；
- ③ 信息综合处理能力良好，能综合分析上下文或相关信息，把握说话者的信息组织方式，明确所听语篇逻辑关系；
- ④ 在主旨大意归纳方面表现欠佳，未能理解中心思想。

英语学习

英语教学

英语测评

车同轨

量同衡

提供统一的能力标准，使基础教育到高等教育沿着同一个轨道循序渐进，实现英语教学“一条龙”

对语言能力进行全面、清晰的界定和描述，在我国英语学习、教学和测评之间架起一座桥梁

提供对各种英语考试**衔接定位**、对英语能力自我诊断的共同标尺，实现多种学习成果的沟通互认

提供**个性化的定性反馈报告**，有利于教师因材施教，学生自主学习，促进教、学、评的有机融合

## 量表对考试的指导作用

- 有利于制定科学合理的考试目标及能力要求
- 使各级各类考试有统一的标准，促进外语考试的科学化、标准化
- 有利于梳理和整合现有的各类英语考试
- 促进学习成果沟通互认和各阶段外语教学和考试的有序衔接
- 有助于我国现有英语考试的国际化



# 建设英语能力等级考试是改革的重要抓手

## 建设目标

1 中国标准，国家考试

2 等级连贯，能力阶梯

3 有效测评，以评促学

4 一考多用，提质增效



*Thank You!*