



清华大学教育研究院
Institute of Education Tsinghua University



为新时代有中国特色的高校人才培养 制定评价标准 — 正在进行的探索

史静寰 清华大学教育研究院教授
清华大学全球学校与学生评价研究中心主任
shijhuan@tsinghua.edu.cn



引子:

最近拿到三个重要研究课题教育部哲社重大课题攻关项目“双一流建设背景下中国高校评价体系改革研究”；中国高教学会“中国高等教育改革发展重大理论实践问题研究”中的“中国特色双一流建设标准及评价体系研究”，上周开题会，教育部相关司局负责人全部参加，法规司、综改司、学生司司长直接做评委；清华大学文科双高计划支持课题“清华大学双一流建设评价体系研究”。

- 如何评价大学（顶尖大学、双一流大学、所有高校）是核心问题。



主要内容

一、高等教育评价之难：全球关注的理论、政策与实践问题

二、构建有中国特色的高教评价体系：中国面对的难而必为之事

三、以构建高校人才培养指数（EI）为例：持续进行的探索



高等教育：国家创新和综合国力竞争的动力来源

• 世界高等教育的发展

- **规模扩张：**全球高等教育已经进入大众化阶段。2013年，全球高等教育的毛入学率为32%，全球已有100多个国家高教毛入学率超过15%，有50多个国家超过50%，有14个国家超过75%。
- **各国的高水平大学建设计划：**韩国“韩脑计划”（Brain Korea 21, 1999）、一流大学项目（World Class Universities, 2009）、卓越计划（Center of Excellence, 2002）、德国“卓越大学计划”（Initiatives for Excellence, 2005）、日本“卓越计划”（Institutions of Eminence Strategy, 2007）

知识经济的竞争归根结底，是人才和人才培养能力的竞争。

• 中国的举措

- **规模扩张：**2015年，中国高等教育毛入学率为40%，超过全球平均水平。预计2020年毛入学率将超过50%，进入高教普及化阶段。
- **重点大学建设计划：**从985、211工程到“双一流”建设。





全球范围内存在学习危机

- **联合国教科文组织2014年提出:**

“当前无论是在发展中国家还是发达国家，虽然教育日见发达，但存在“**学习危机**（Learning crisis），很多国家和地区仍未能建立**有效支持学习**的教育系统。”

—UNESCO 2014

- **世界银行2017年报告谈“学习危机”：**

“Schooling without learning is a wasted opportunity.”

—World Bank 2017

Figure O.7 Why learning doesn't happen:
Four immediate factors that break down

教师教学技能不足
缺乏激励

学生学习准备不足



教育投入未
达及教学

学校教学管
理水平低下

Source: WDR 2018 team.



我国高校评价实践发展历程



清华大学教育研究院
Institute of Education Tsinghua University

学习、追赶和寻求认可阶段：211 和985工程的实施

对话、自主发展和凸显特色阶段：双一流建设启动

加强教学工作 提高教学质量



教育部本科教学工作水平评估



教育部学位与研究生教育发展中心 学科评估

服务发展大局、促进学校提升、回应公众需求

评价主体、对象的多样化

官方和半官方评价

科学计量学的引入

非官方（高校、商业机构）评价

1987



赵红洲
高校科研能力排行

预测、评价国家科研能力

1993

武书连
中国大学评价

1994

中国校友会网
公办民办分类排名
中国校友会网
专注大学与创富研究

2002

上海交通大学
世界大学学术排行



为学校发展寻找参照系

2003

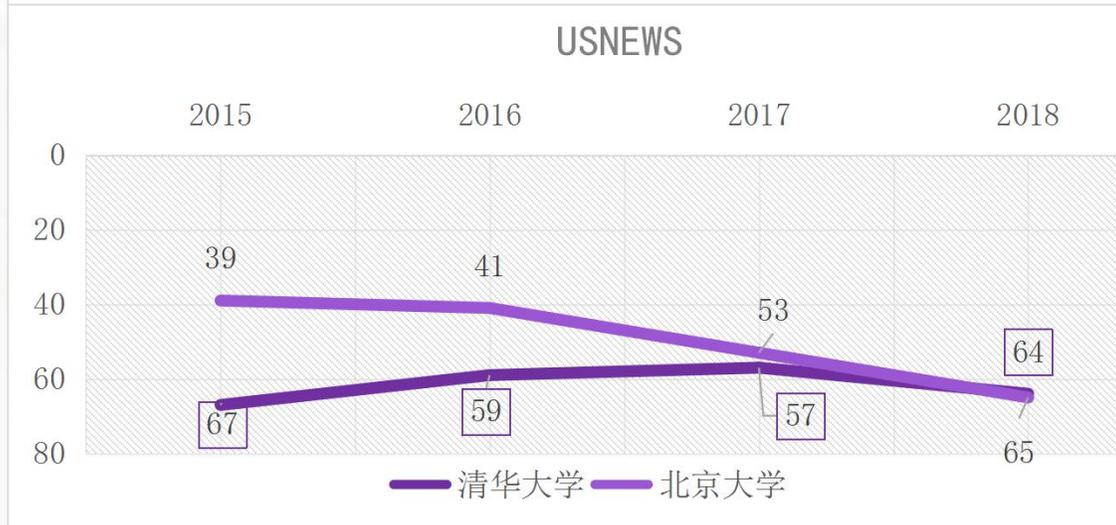


面向公众、商业驱动

2009



世界大学排名中的中国顶尖大学： 由“形似”到“神备”的转型



ARWU

学校	2017	2016	2015	2014	2013	2012	2011
清华	48	58	101-150	101-150	151-200	151-200	151-200
北大	71	71	101-150	101-150	151-200	151-200	201-300



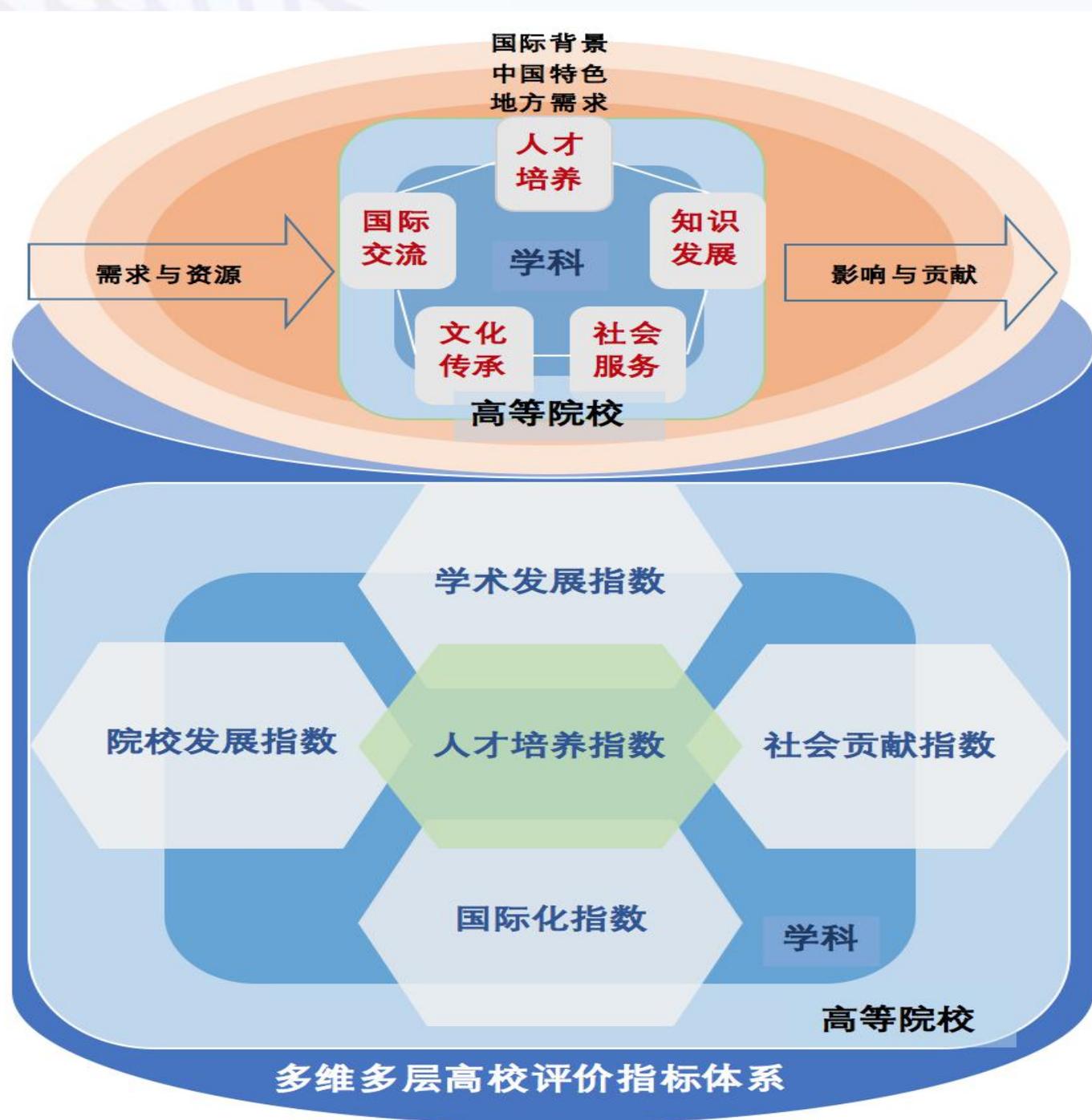
我国高等教育评价体系存在的问题

- 因过于注重效仿西方而**抽离了高校存在于其中的社会情境**，对中国独特国情下高等教育系统运作特点体现不足；
- 因理论研究不足，**评价标准固化，指标设计碎片**，未能整体审查高校发展的内外部质量，审查学校与学科、高校与学生、高校与区域发展等主体之间的复杂关联；
- 因创新探索不足，**未能基于实践探究**，提炼出我国高教发展进入新阶段的重点突破和改进方向，形成有中国特色的高教评价体系。

研究框架

双一流背景下的 多维多层高校评价指标体系

- 上层体现高校的主要职能及所处环境；
- 下层展示本课题将构建的高校综合评价指标体系中的五大模块；
 - **人才培养指数**
 - 学术发展指数
 - 院校发展指数
 - 社会贡献指数
 - 国际化指数

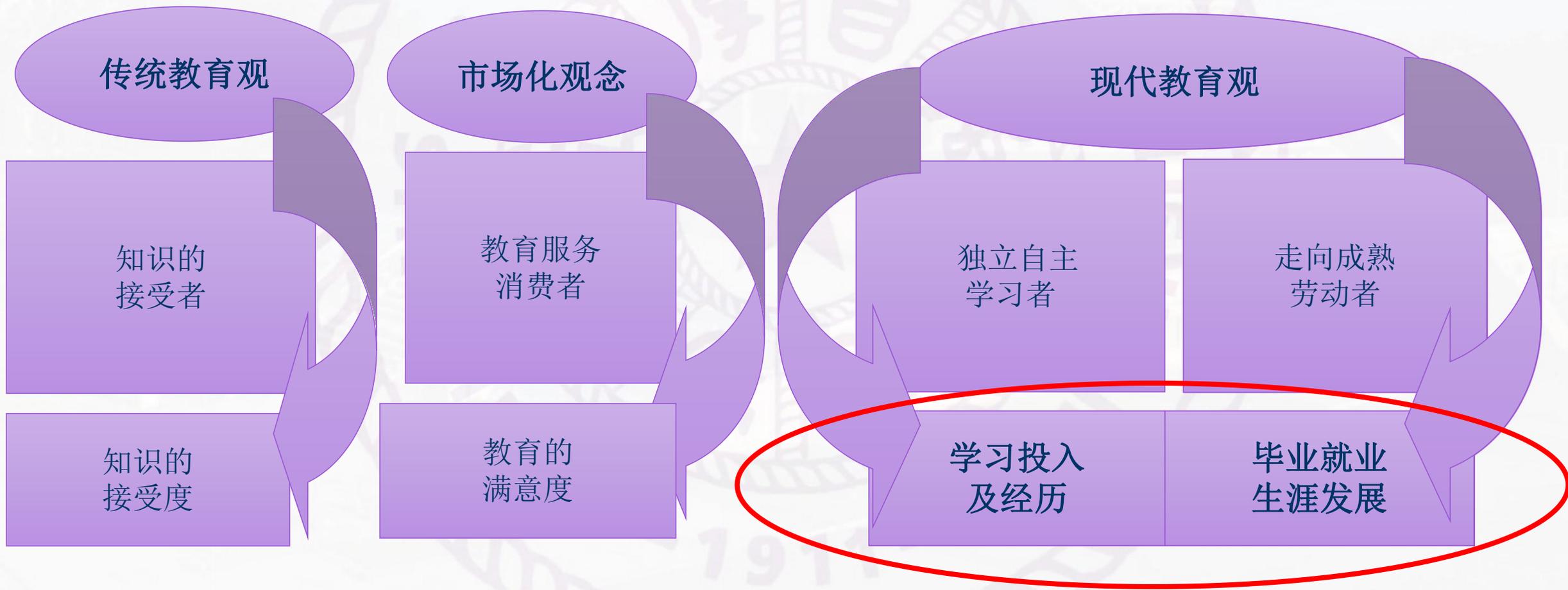


以人才培养指数的构建为例看

如何体现以学习者为中心的大学（教育）质量评价



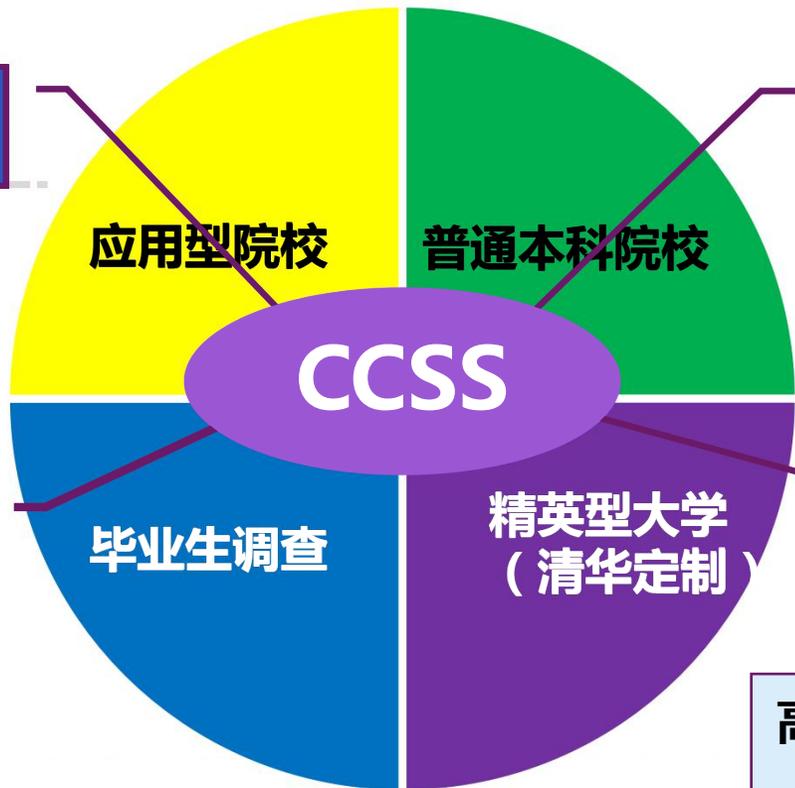
以学生为中心，以可显现的证据为基础，
以教育成果及其持续效用为核心的教育质量评价体系





CCSS

中国大学生学习与发​​展追踪研究
China College Student Survey



学生学习类型
和方式研究

不同类型院校的学
生发展研究

学生学习满意度
和毕业发展研究

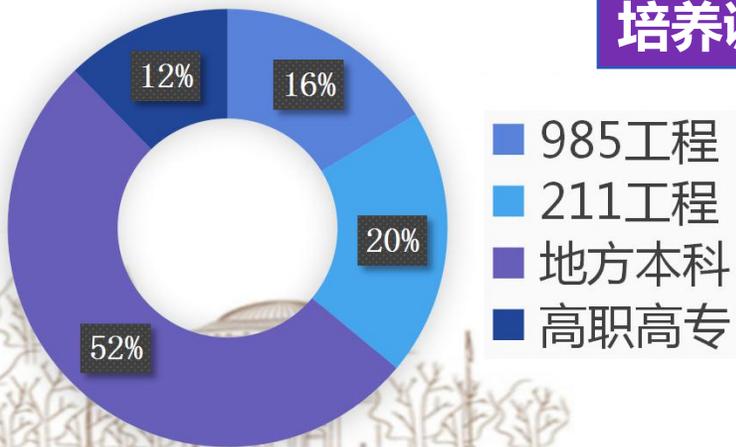
实验班/特殊人才培
养计划研究

清华大学博士生创新能力与
培养调查

高等工程教育质量国际比较项目
SUPERtest
(计算机科学、电子工程1、3年级的学生)
测试四科目：
审辩性思维、定理推理、数学、物理

突出以“学”（学生，学习）为中心核心理念；表现学生学习与发展的曲折过程。

参与院校类型分布



- 清华大学“中国大学生学习与发​​展追踪研究”（CCSS）团队开发了适合不同类型院校人才培养过程的调查工具。
- 2009-2016年在全国采集大学生学习样本近**60万**。
- 参与院校总数为147所。院校类型包括“985工程”院校24所，“211工程”院校29所，地方本科院校76所，高职高专院校18所。

CCSS普通本科学情调查指标体系



1 综合诊断指标

- 学业挑战度
- 主动合作学习水平
- 生师互动
- 教育经验的丰富程度
- 校园环境支持度

2 教育过程诊断指标

- 课程认识目标
- 课程要求严格程度
- 有效教学实践
- 课程学习行为
- 课程外拓展性学习行为
- 支持性环境
- 向学/厌学
- 自我报告的教育收获
- 在校满意度

- 课程教学
- 教师反馈
- 激发学习志趣
- 测评（多元评价）

3 学习诊断指标

- 高阶认知行为
- 学习策略
- 多元学习

- 高阶学习
- 信息分析
- 跨文化学习
- 在线学习

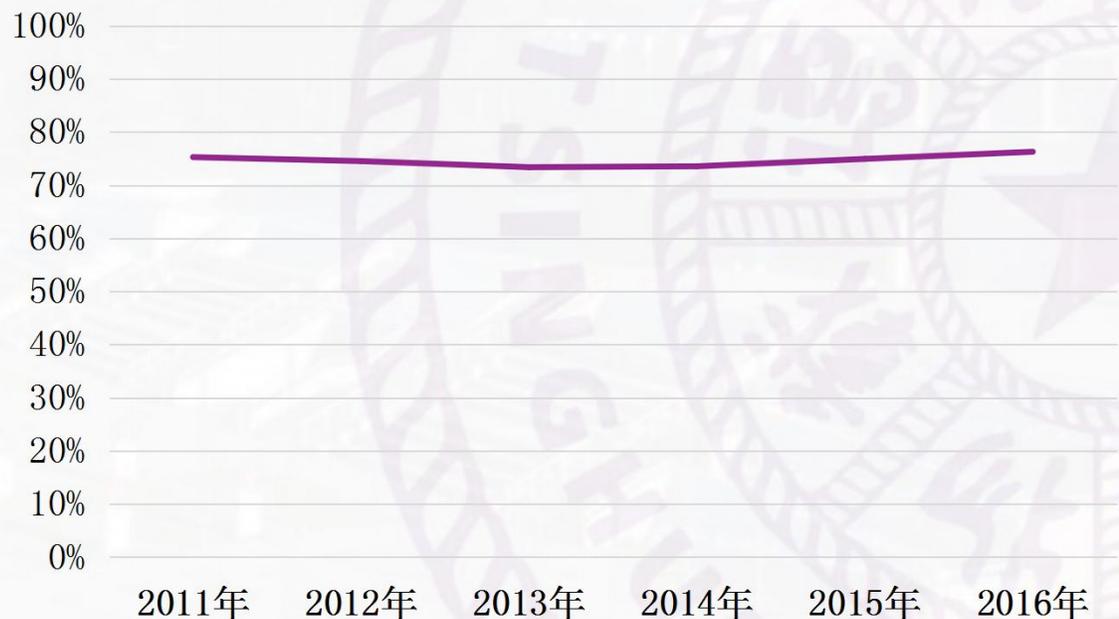
- 接受式学习
- 探究式学习
- 反思性学习
- 整合性学习
- 合作性学习





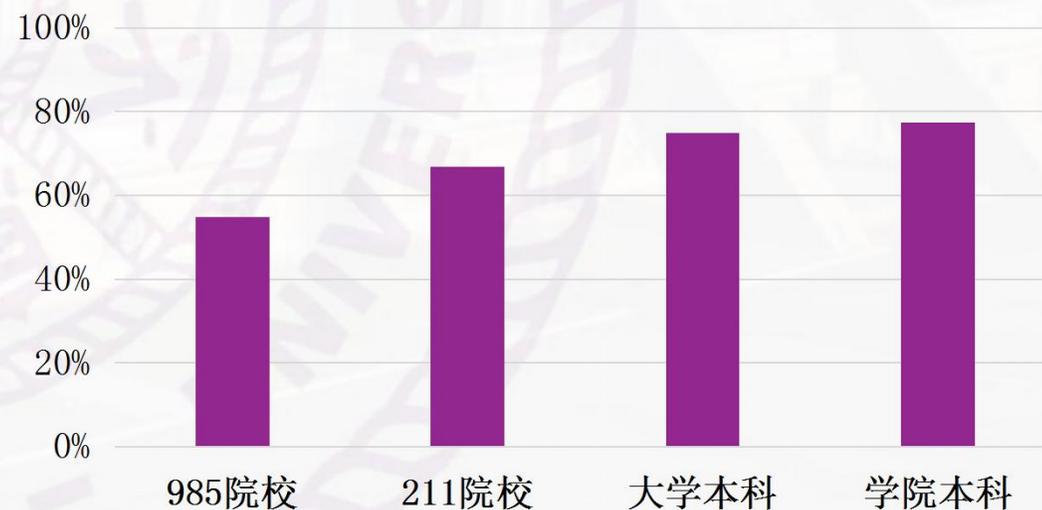
数据揭示的教育入口：教育公平

中国本科高校第一代大学生比例
(2011-2016各年)



中国高等教育的规模扩张使得原本少有机会接受高等教育的适龄青年进入大学。**中国70%以上的大学生是家庭第一代大学生**，体现出入学机会的提升。

各类型高校第一代大学生占比
(2011-2016平均)



第一代大学生：指其父母均只拥有高中或者高中以下学历的大学生



接受性学习尚可，主动性学习不足

不同类型学习行为	第一代大学生		非第一代大学生		差值	Cohen's d
	均值	标准差	均值	标准差		
接受式学习	57.55	19.19	59.23	21.03	-	-0.06
主动性学习	42.03	20.06	46.99	23.34	1.16***	-0.20
反思整合学习	57.00	17.14	60.05	18.92	4.27***	-0.14

主动性学习指的是在课程要求之外，学生主动参与课堂提问、讨论、作汇报等学习活动。

课内学术性人际互动表现不足：

学术性互动	第一代大学生		非第一代大学生		差值	Cohen's d
	均值	标准差	均值	标准差		
学术性生师互动	41.29	19.90	45.49	23.18	-3.52***	-0.17
学术性生生互动	58.50	18.54	61.84	20.45	-2.87***	-0.15

学术性互动指的是，与教师和同学就课程内容、作业、考试等方面进行讨论。

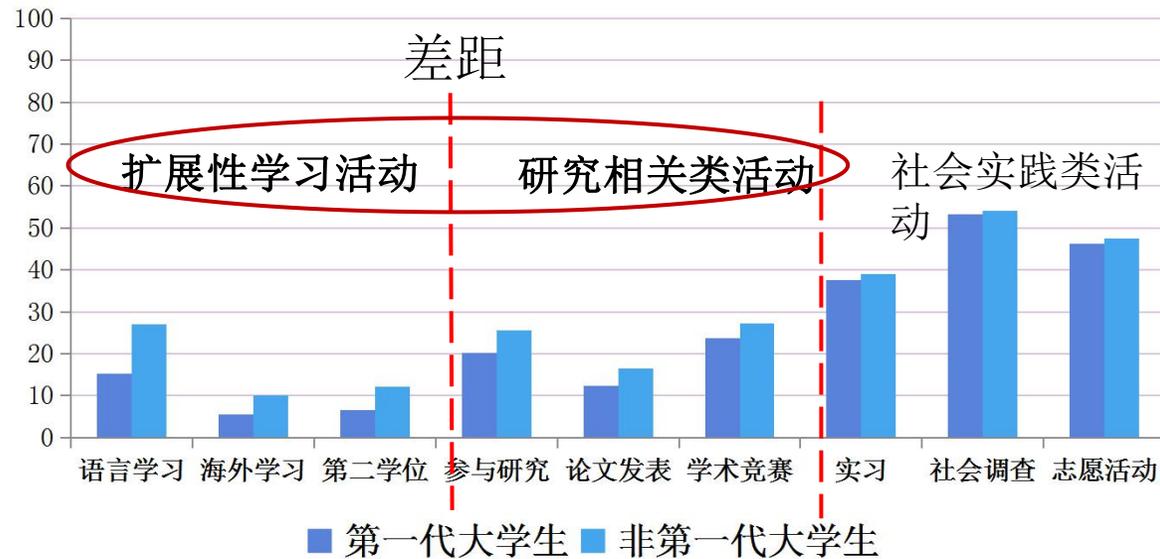
课外社交性人际互动方面表现欠缺

社交人际互动	第一代大学生		非第一代大学生		差值	Cohen's d
	均值	标准差	均值	标准差		
社交性生师互动	34.79	26.32	40.92	29.06	-5.14***	-0.19
社交性生生互动	45.55	19.70	51.42	21.72	-5.45***	-0.27

社交性互动指的是，以认识不同类型群体、丰富人际生活、明确发展目标为目的，和同学、老师进行的交流和交往

- 拓展性学习活动指的是在课程和专业常规要求之外的进一步或跨领域的学习。
- 研究相关性活动是指在深入学习基础上的研究探索，需要自身具有较强的研究意向、以及教师给予较多学术指导。
- 社会实践类活动指的是在校外现实情境中进行的、增加学生社会经验、促进综合认知、促进知识在现实情境中应用的活动。

参与扩展性和研究性高影响力活动不足



数据来源：CCSS绿色问卷2015年调查全国样本

“985”院校与美国研究型大学（RU/H）的比较（2011）

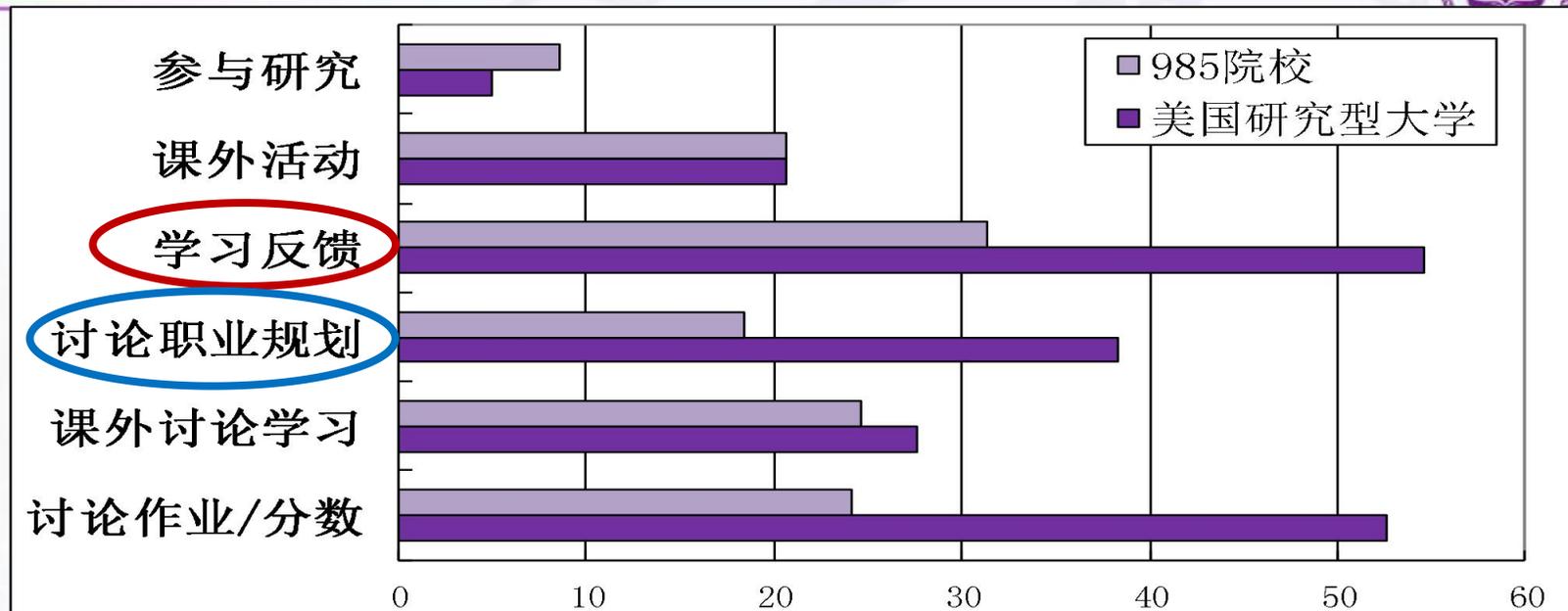


- **985院校**学生在教育经验丰富度表现不差，甚至更好（低年级）
- 校园环境支持度无实质性差异 ($es < 0.2$)，主动合作学习低年级无差异；
- 生师互动和学业挑战度，中美存在中度差异 ($es > 0.5$)。

生师互动的教育环节存在问题诊断

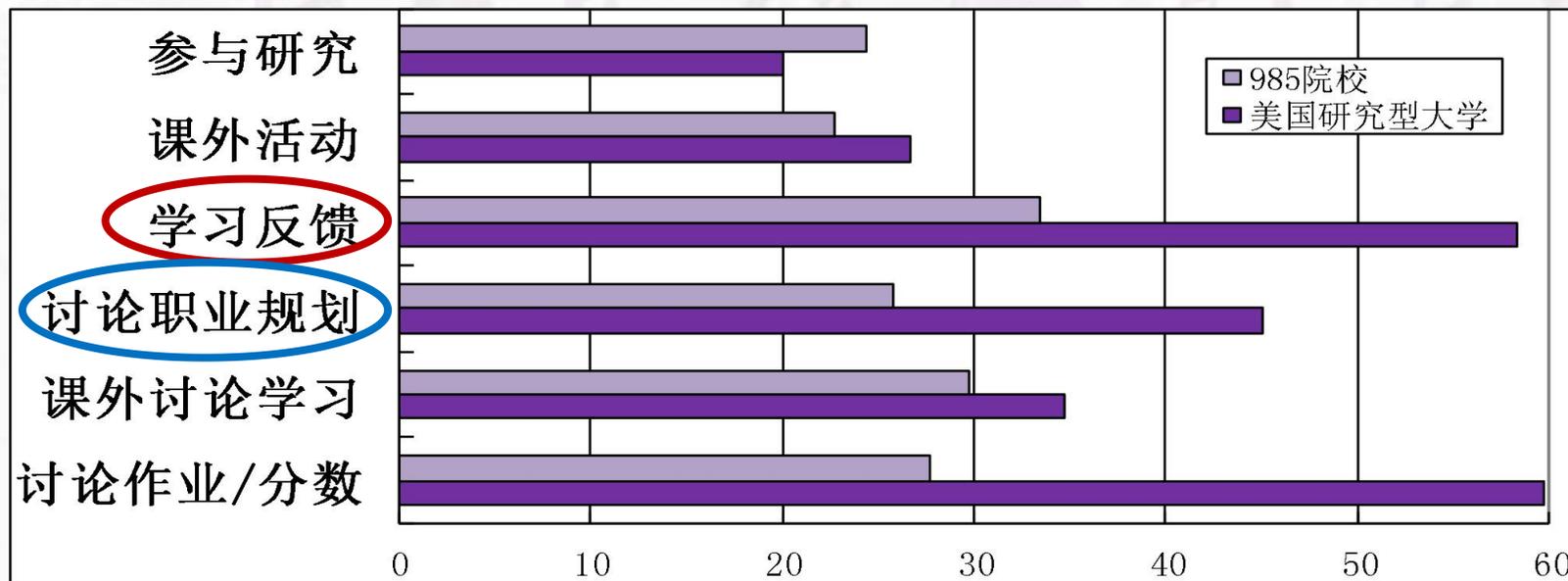


低年级



学术交往：老师对学生的学习反馈（口头、书面）不够。

高年级



社会交往：缺乏对学生职业规划的讨论与指导。

文化差异：生师讨论分数。



数据揭示的教育过程：影响因素探索

综合指标	自我汇报教育收获	GPA	学习指标	自我汇报教育收获	GPA
	beta/se	beta/se		beta/se	beta/se
学业挑战度	0.151*** (0.00793)	0.025* (0.000512)	向学厌学	0.275*** (0.00922)	0.091*** (0.000478)
主动合作学习	0.115*** (0.00495)	0.098*** (0.000359)	接受式学习	0.029*** (0.00745)	0.099*** (0.000248)
生师互动	0.077*** (0.00564)	0.056** (0.000537)	探索式学习	0.100*** (0.00444)	0.084*** (0.000483)
教育经验丰富度	0.221*** (0.00729)	0.012 (0.000411)	反思性学习	0.144*** (0.00537)	0.022* (0.000315)
校园环境支持度	0.263*** (0.00625)	0.023* (0.000410)	整合性学习	0.181*** (0.00496)	-0.002 (0.000268)
			合作性学习	0.081*** (0.00898)	0.000 (0.000332)
N	322552	207391	N	120845	79979
r ²	0.450	0.0909	r ²	0.471	0.152

学习投入综合指标、学习态度、学习策略均与学习结果显著相关，过程性指标总体的校标效度较好

学生的GPA和自我汇报的教育收获的影响因素略有不同

- 学生自我汇报的教育收获更多受学习态度、教育经验丰富度和校园环境支持度的影响
- GPA更多受学习态度、接受式学习（学习习惯）、探索式学习（主动学习）的影响

自我汇报的教育收获和GPA反映不同的学习结果

控制性别、年级、专业方向、院校类型、调研年；标准误聚集于学校层面；
* p<0.05, ** p<0.01, *** p<0.001



如何在教育质量评价中加入学生的声音？

- **基础：**以CCSS为代表的学情调查在中国高校中广泛展开
 - 2009-2017年，已累计有来自全国28个省 / 市 / 自治区的近150所普通本科、民办院校、高职高专等各类院校参与过CCSS项目调查
 - 绿色问卷参与院校98所，累计样本量近68万（有效随机样约45万）
- **积累：**学情调查结果如何用于人才培养质量评价？
 - 基于多年学情调查数据积累，找到具有稳定性、指征性的题项
 - 提炼、抽象出对高等教育人才培养质量影响较大、且能够区分院校人才培养水平的指标
 - 结合院校公开数据，形成有中国特色的高等教育评价的指标
- **举措：**设计提出“人才培养指数”。





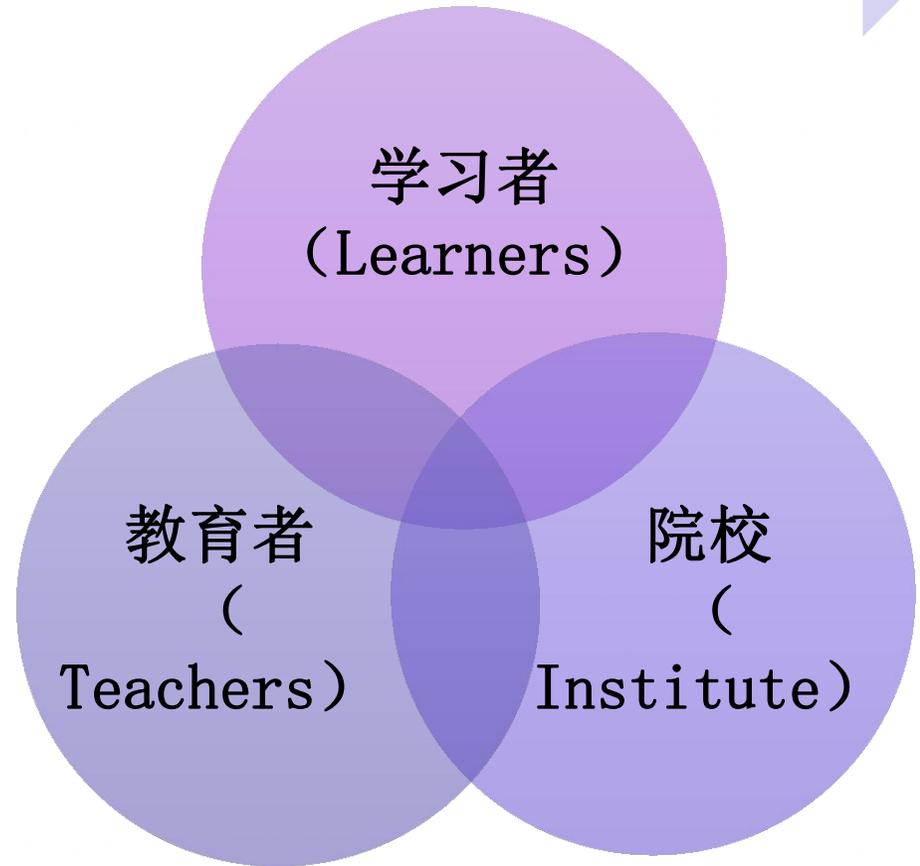
人才培养指数构建设想

- **一个理念**：学生在高等教育环境下的学习与发展，是学生与教师、他人和环境互动的结果
- **两类数据**：学生调查数据、院校公开数据
- **三方视角**：院校、教师、学生
- **三种质量**：投入水平、过程性质量和结果性质量
- **多种方法**：统计测量学分析、专家调查法（德尔菲Delphi法）、数据挖掘等。

投入

过程

产出



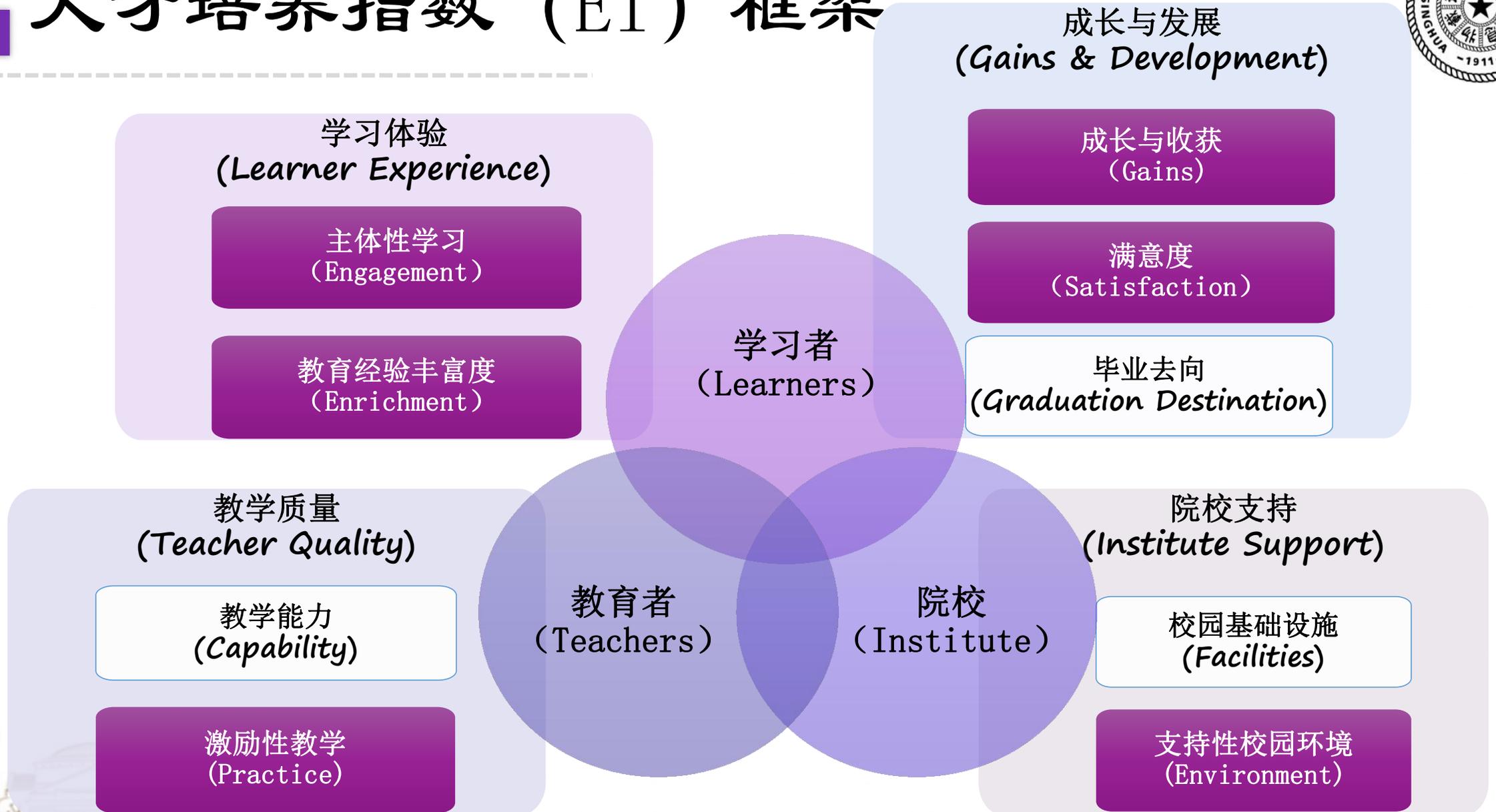


初步构建 (Process)

- 多轮数据挖掘+讨论会
- 结合统计测量学分析、理论及专家建议，初步选择了**50道**CCSS题项
- 参考各排行榜及现有评价体系的指标，综合数据的可得性，结合专家建议，初步选择了**20个**常用且较易得的客观指标
- 初步构建起**三层四维度**的人才培养指数



人才培养指数 (EI) 框架





二、三级指标及数据来源

		CCSS题项数目	公开数据数目
院校支持 (Institute Support)	校园基础设施 (Facilities)	2	3
	支持性校园环境 (Environment)	6	3
教学质量 (Teacher Quality)	教学能力 (Capability)	0	11
	激励性教学 (Practice)	7	0
学习体验 (Learner Experience)	主体性学习 (Engagement)	10	0
	教育经验丰富度 (Enrichment)	2	0
	成长与收获 (Growth)	17	0
成长与发展 (Growth & Development)	满意度 (Satisfaction)	2	0
	毕业去向 (Graduation Destination)	0	2



初步结果

- 2016年数据，35所院校 (35 universities)，4个学科领域 (4 major areas)
- 数据处理 (adjustment)：主客观所有题项均调整到0-100区间
- 权重处理 (weighting)：目前使用算术平均分 (equal weight)
- 分析方式 (reporting)：
 - 可在一级、二级、三级指标上分别分析 (analysis)
 - 可对学校整体分析，亦可分学科比较 (institutional major area)
 - 可使用得分分析 (score, 绝对值)，亦可根据值)
 - 分级方案：根据得分距院校均值的标准差分等级

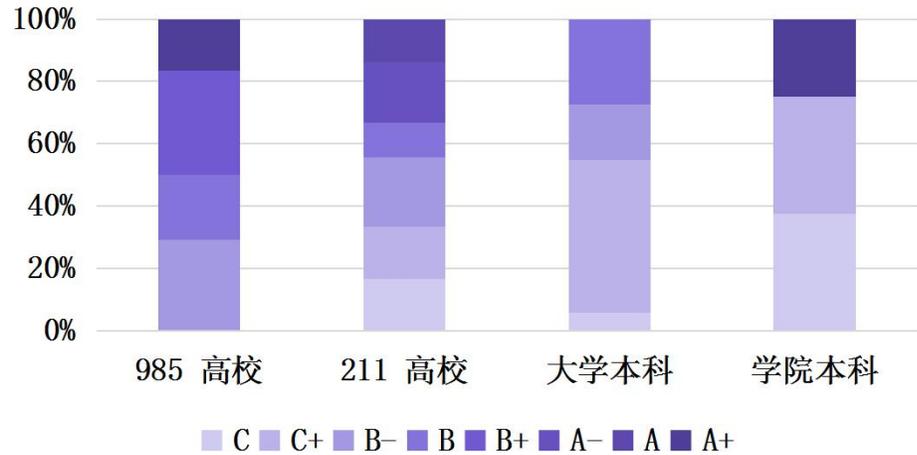
A+:	> 2	SD
A:	1.5 ~ 2	SD
A-:	1 ~ 1.5	SD
B+:	0.5 ~ 1	SD
B:	0 ~ 0.5	SD
B-:	-0.5 ~ 0	SD
C+:	-1 ~ -0.5	SD
C:	-1.5 ~ -1	SD
C-:	< -1.5	SD

二级指标 (2nd Level)

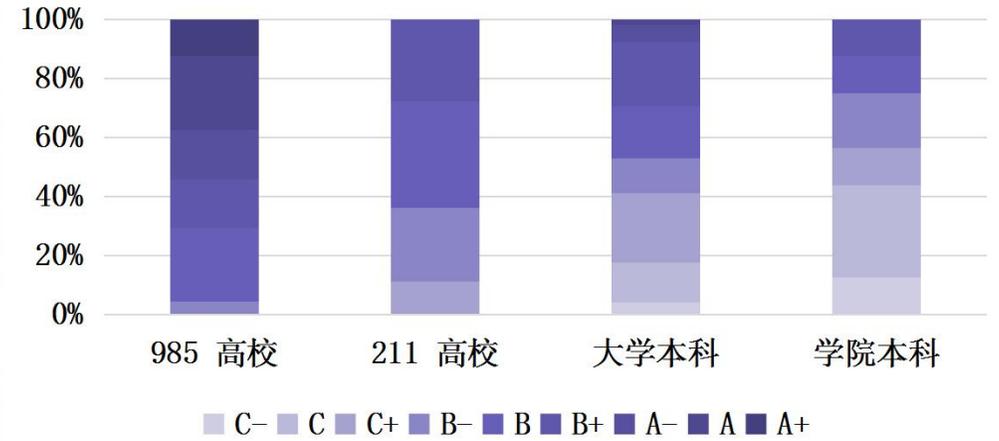


(以学科领域为分析对象, 共127个观测值)

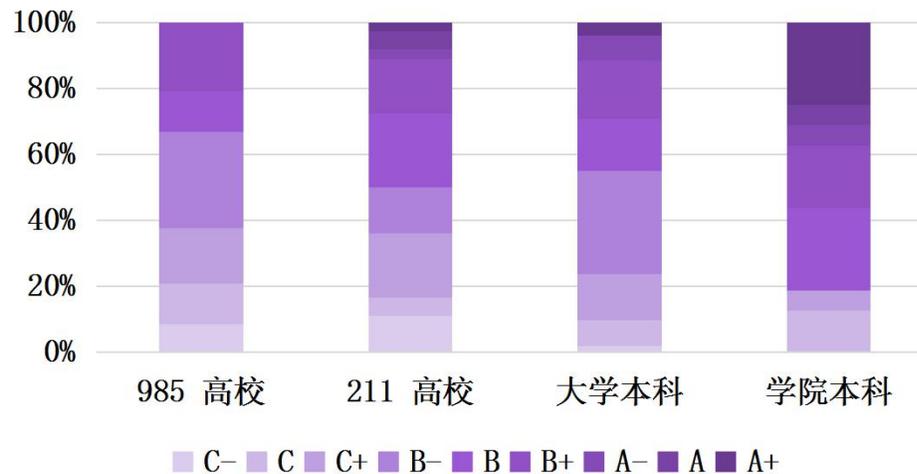
院校支持 (IS)



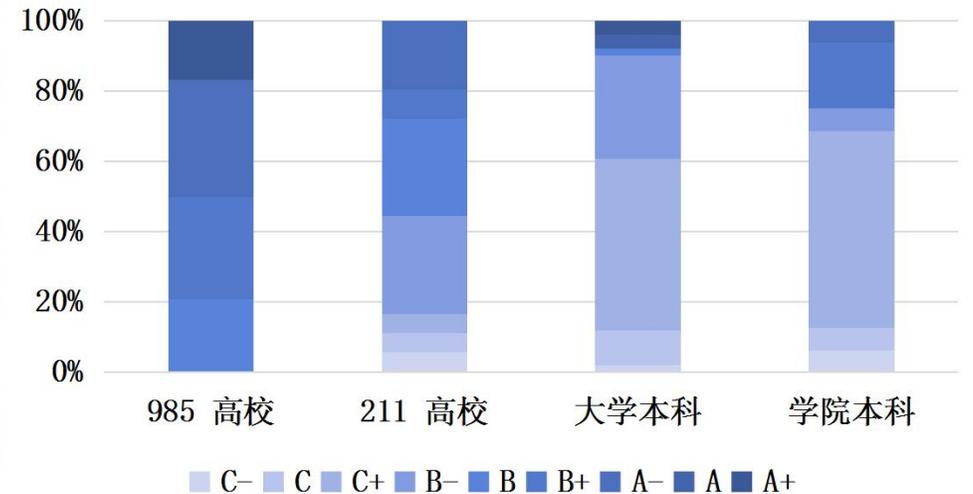
教学质量 (TQ)



学习体验 (LE)



成长与发展 (GD)





教师教学与学生学习结果之间具有复杂关系

对学生学习结果和影响因素的一项研究发现：

- 教师整体水平，课堂教学的有效性、讲解清晰度对学生的专业基础知识学习（数学、物理成绩）有显著正向影响，但对学生综合思维能力（批判性思维、量化抚养）的影响不明显。传统教学方式对学生的批判性思维甚至有负向影响。
- 教师的科研能力、成果和国际化程度对学生的专业基础能力有一定正向影响，对综合思维能力的影响不明显。教师整体在科研和国际活动上的时间投入，与学生的综合思维能力得分及其年级变化显著负相关。



教育：改变人的活动

- “教育具有一种道德上的目标，这就是不论学生背景如何，要使他们的一生有变化，并在充满活力且日趋复杂的社会中有助于造就出能够生存和有工作成果的公民。”
- “教育工作者必须把他们自己看做、也被别人看做是变革动力的专家。为了成为变革动力的专家，教育工作者、行政人员和教师必须成为熟练的变革力量。”
- “我们有一个从根本上说是保守的教育系统。教师的培训方式，学校的组织形式，教育层级的运作方式以及政治决策者对待教育的方式都容易导致维持现状和难以变革的制度。”

迈克尔·富兰《变革的力量—透视教育改革》



清华大学教育研究院
Institute of Education Tsinghua University



感谢
THANKS

