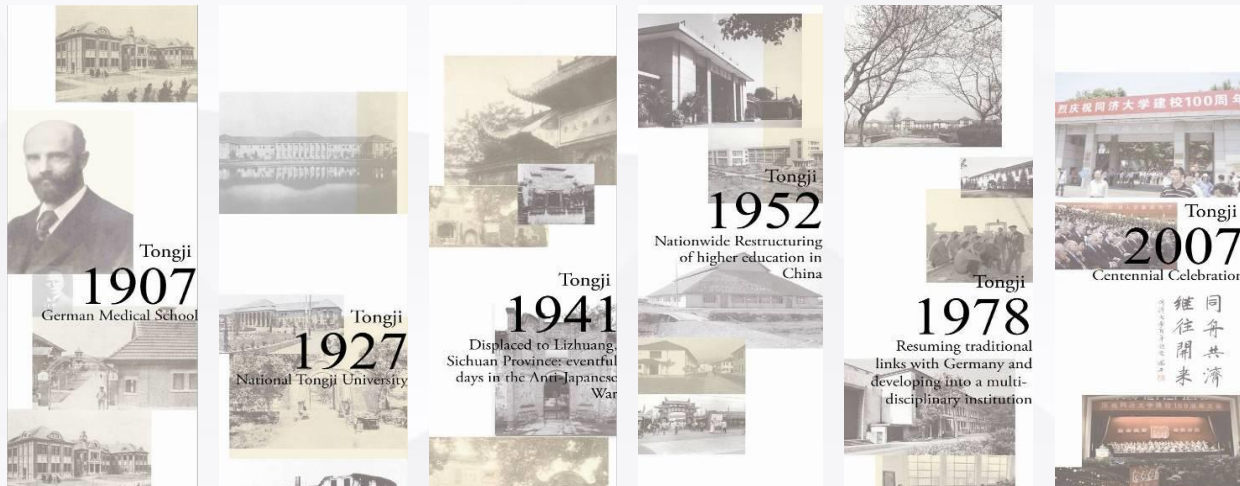




TONGJI
UNIVERSITY



第九届全国高校物理实验教学研讨会

美国物理教育评估方法简介与应用

同济大学

物理科学与工程学院

方 恺

2016/7/17

目录

Contents

01 物理教育研究 (PER) 简介

02 物理教育评估

03 对学习内容的调查

04 对学习态度的调查

05 对逻辑推理能力的调查

06 教学应用

**第九届全国高校物理实验教学研讨会——
美国物理教育评估方法简介与应用**

物理教育研究（PER）简介

物理教育研究

(Physics Education Research, PER)

- 在上世纪90年代美国逐渐发展成为一个新的物理研究方向；
- 成为物理学与教育学两个领域之间的关联学科；
- 基于对学生思维和学习的模式的研究，提出一些新的物理教学方法和技巧，应用于教学以提高教学质量、发现并解决物理教学中存在的问题和难题，激发学生学习物理的积极性，加深对物理概念的理解和掌握，并提高他们的科学分析能力。

APS Statement on the Position of PER

99.2 RESEARCH IN PHYSICS EDUCATION

(Adopted by the Council, 21 May 1999)

- In recent years, physics education research has emerged as a topic of research within physics departments.

This type of research is pursued in physics departments at several leading graduate and research institutions, it has attracted funding from major governmental agencies, it is both objective and experimental, it is developing and has developed publication and dissemination mechanisms, and Ph.D. students trained in the area are recruited to establish new programs.

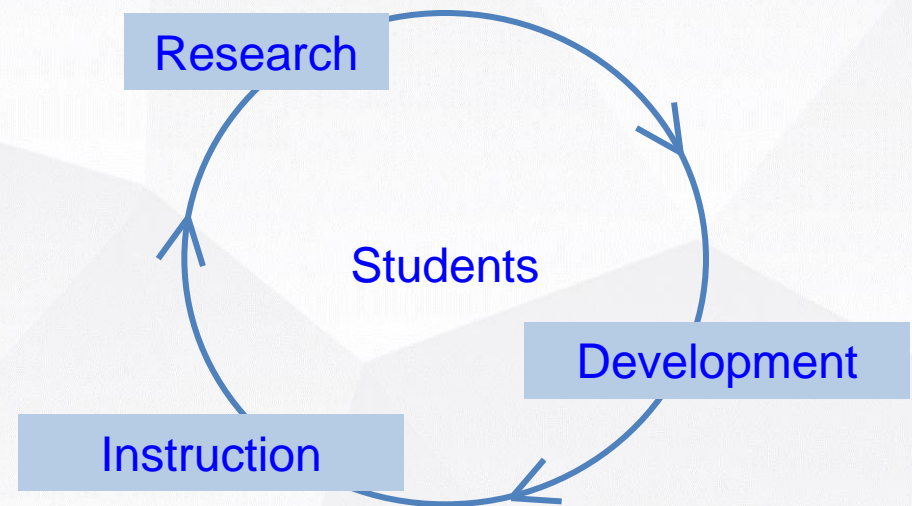
Physics education research can and should be subject to the same criteria for evaluation (papers published, grants, etc.) as research in other fields of physics. The outcome of this research will improve the methodology of teaching and teaching evaluation.

- The APS applauds and supports the acceptance in physics departments of research in physics education.

Much of the work done in this field is very specific to the teaching of physics and deals with the unique needs and demands of particular physics courses and the appropriate use of technology in those courses. The successful adaptation of physics education research to improve the state of teaching in any physics department requires close contact between the physics education researchers and the more traditional researchers who are also teachers. The APS recognizes that the success and usefulness of physics education research is greatly enhanced by its presence in the physics department.

学生为中心的新的教学方法

- **Physics by Inquiry (Pbi)**交互式物理教学
- **Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP)**以学生为中心的大班本科物理课程教学计划
- **Voting Machine (Clicker) Method**基于Voting machines的课堂教学方法
- **Peer Instruction**同伴教学法
- **Workshop Physics**
- **Real Time Physics**



(McDermott, UW, 1990s)

Inquiry-based Teaching Strategy

Problem Statement

Determine what is to be investigated and formulate a question or hypothesis.

Data Collection

Gather as much information about the topic from appropriate sources.

Analysis

Examine and discuss the findings and provide explanations or clarity.

Conclusions

Based on analysis determine solutions related to the original problem statement.

研究方法

- **Qualitative**
 - Interviews, open ended questions, controlled training and testing,
- **Quantitative**
 - Standardized large scale assessment (theoretical models and instruments)
- **Basic Learning**
 - controlled conditioning/training and testing,
 - technology
 - Eye tracking,
 - brain wave EEG (Electroencephalography),
 - FMRI (Functional magnetic resonance imaging).

研究方法

- **Qualitative**
 - Interviews, open ended questions, controlled training and testing,
- **Quantitative**
 - Standardized large scale assessment (theoretical models and instruments)
- **Basic Learning**
 - controlled conditioning/training and testing,
 - technology
 - Eye tracking,
 - brain wave EEG (Electroencephalography),
 - FMRI (Functional magnetic resonance imaging).

Typical Process in PER

- **Observations in education practice and assessment**
- **Formation of hypothesis**
- **Experiments to collect evidence**
 - **Detailed investigations**
(recorded interviews, open questionnaires, surveys, ...)
 - **Controlled Studies**
(randomized treatment and control groups, pre-post testing)
- **Designing/revising treatment**
- **Implementations**
- **Documentation of findings**
- **Further cycles of research, development and implementation**

**第九届全国高校物理实验教学研讨会——
美国物理教育评估方法简介与应用**

物 理 教 育 评 估

物理教学评估

- **物理教学效果的评估的意义：**

从教师的角度而言，需要了解学生的学习进程和学习状态；

从学生的角度而言，需要知道自己是否达到了学习的要求及是否需要改进自己的学习方法。

- **传统的教学效果评估方法：**通过作业完成情况和考试等来评定的。
- **PER常用的方法：**以研究为基础的量化的物理教学评估，涉及的内容包括力学、电磁学等物理学的各个方面，还拓展到逻辑推理能力等的评估。

物理教育评估的类型

(1) 对学习内容调查

(2) 对学习态度调查和科学分析

(3) 逻辑推理能力调查

以教学研究为基础的调查的特点

- **完成时间**：10至30分钟。
- **题目类型**：选择题、判断题和简答题（以选择题为主）。
- **答卷方式**：笔试，或在计算机、手持设备等上答题。
- **测试人数**：通常学生数量较大，一般进行课前和课后两次测试。
- **结果处理**：教育统计学方法等。

有效调查测试的特点

- 1、根据学生在测试中的表现，可以定性地研究学生的学习状态模型。
- 2、设计理论框架以模拟学生对一定专题的反应。
- 3、对测试结果的分析，也包括对学生选择的错误答案的分析，促进了新的教学模式和诊断评估方法的发展。
- 4、以测试结果为导向，构建新的定性的研究，设计新的、更为有效的调查。
- 5、调查测试的有效性和可靠性。

Assessment Results of PER Based Instruction

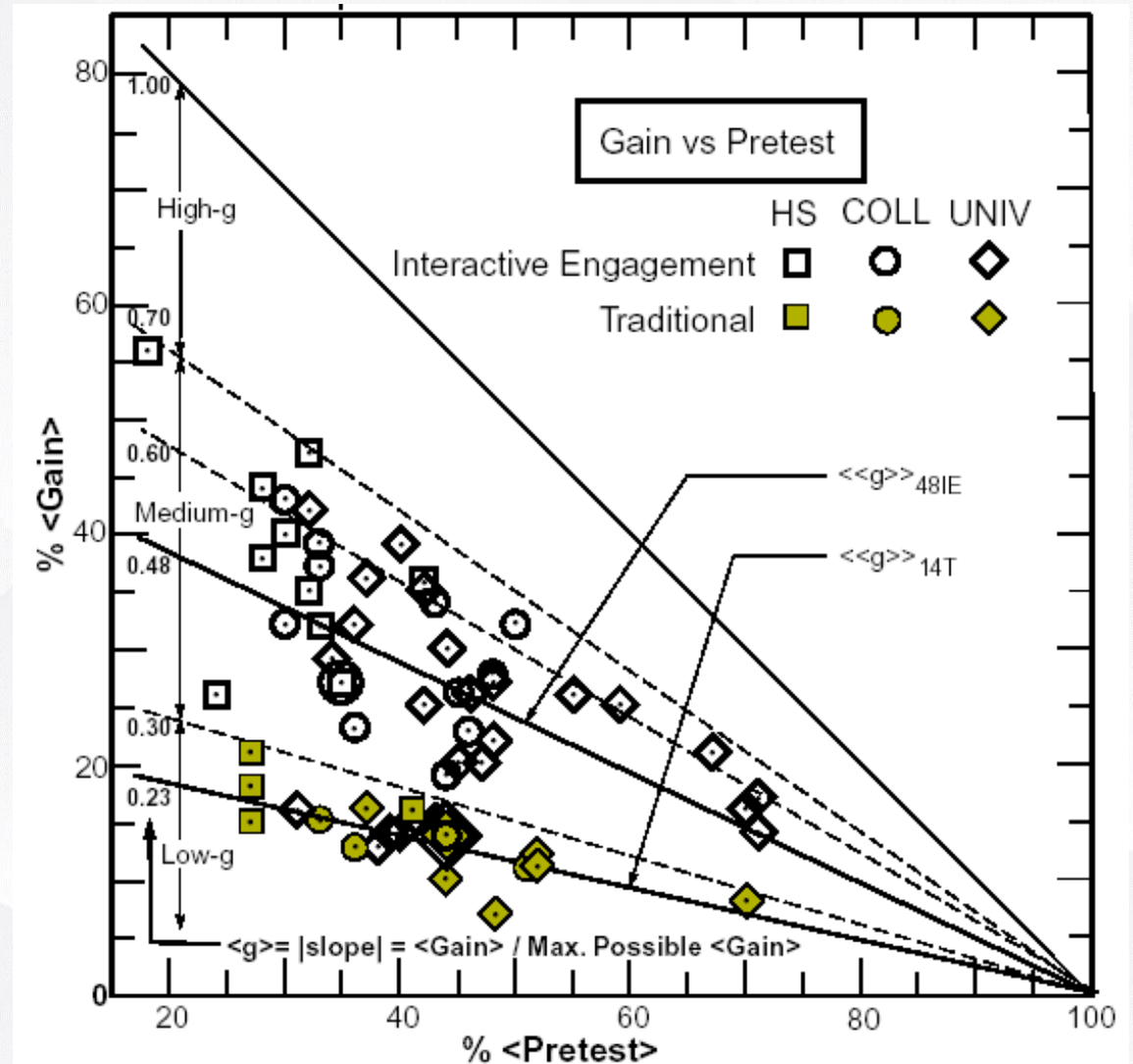
Hake 1998, 6000 students, 100+ schools.

Assessment tool:
Normalized Gain (g)

$$g(x, y) = \begin{cases} \frac{y-x}{1-x} > 0 & y \geq x \\ \frac{y-x}{x} < 0 & y < x \end{cases}$$

Main Results

- Traditional Instruction is not effective in helping students develop correct understanding of physics concepts.



第九届全国高校物理实验教学研讨会—— 美国物理教育评估方法简介与应用

对学习内容的调查

1. 力学概念汇总 (Force Concept Inventory, FCI)

力学概念汇总 (Force Concept Inventory, FCI) 通常称之为FCI测试, 由美国亚利桑那州立大学教授David Hestense等设计, 并于1992年发表, 而后在1995年进行了修改, 其前身是1985年发表的力学诊断测试 (The Mechanics Diagnostic Test, MDT)。

FCI测试 (1995年版) 由30道选择题构成, 用于测试学生对力学和动力学概念的理解和掌握情况预计完成时间为15到30分钟。

FCI设计理念

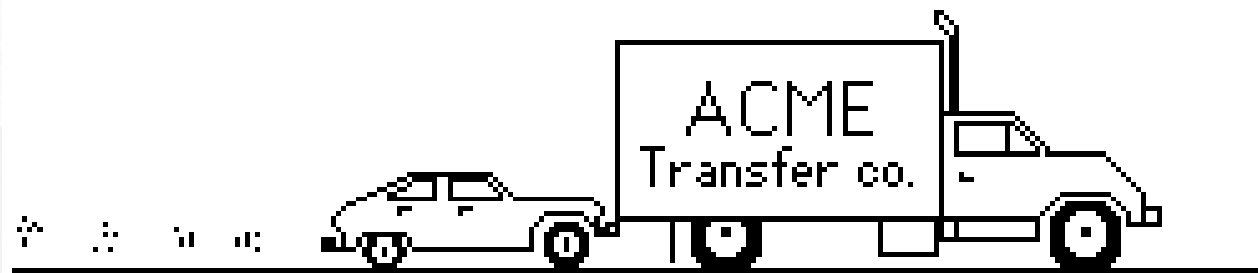
主要内容： 是对于牛顿力学概念的理解，知识点涉及牛顿三大定律、运动学、矢量合成等。

出发点：

- 学生在学习物理之前已经通过生活常识和个人经验，形成了一系列对客观世界的认识。
- 从事物理教学研究的教师们发现学生的这些认识在物理学习的过程中常占有优势地位，于是在理解物理问题时会产生误导，使学生在解答测试题时，其结果出现偏差或错误。
- 通过FCI测试，学生必须在牛顿力学概念和常识两者之间做出选择，因此测试可以发现由于上述原因而造成的对物理问题理解的错误的详细情况。

FCI测试（1995年版）中的15题：

一辆大卡车在路上抛锚了，现它被一辆小型汽车从后面推它向城里开去，如图所示。



在小车推大车的速度慢慢增大到一个稳定速度的过程中：

- A) 小车推大车的作用力和大车对小车向后的作用力相等；
- B) 小车推大车的作用力比大车对小车向后的作用力小；
- C) 小车推大车的作用力比大车对小车向后的作用力大；
- D) 因为小汽车的发动机工作着，所以它对大卡车有推动作用，但是大卡车的发动机没有工作，所以大卡车对小车没有向后的作用力。大卡车只受到向前的作用力，使它和小车一样运动；
- E) 两车之间没有相互作用力。

学生遇到上题时可能会错误地认为两个物体之间相互作用时，胜利属于较强的一方，即他们认为的优先原理，于是觉得牛顿第三定律不正确，结果选则了错误的答案。

FCI应用情况

- FCI测试是迄今为止被应用得最为广泛的测试。自1992年发表后，立刻受到物理教育研究领域的广泛关注。参与测试的学生的范围从中学生到大学生，并由此收集到大量的数据（1992年以收集了超过60个班级的测试结果），可用于评估和对比各个不同层次的物理教学水平。
- 美国大学生在FCI测试中的平均正确率：

	有微积分基础的大学生	有代数基础的大学生
课前测试（Pre-test）	40%-50%	30%-45%
课后测试（Post-test）	60%	50%

2. 有关牛顿力学和运动学的其他测试

- 力学和运动概念的评估

(The Force and Motion Conceptual Evaluation, FMCE)

由Ron Thornton教授等设计，并发表于1998年。与FCI 测试相同之处在于都是测试学生对于牛顿力学和运动学概念的理解，题目类型都是选择题，不同之处是加强了学生对于动力学问题的理解，及用文字描述和用物理图像描述之间的转换。

经比较发现，FCI测试和FMCE测试的结果具有很强的相关性，两者间的差异在于对于成绩较低的学生FMCE测试显得更难一些。

为测试学生运用物理概念解决物理问题的能力，Hestenes教授等设计了力学基本测试 (Mechanics Baseline Test, MBT)。MBT测试可作为进行FCI测试之后的第二步测试。

MBT测试中的错误答案不再是常识性错误选项，而是学生由于对概念理解出现的典型错误。

3. 电磁学测试

(The Brief Electricity and Magnetism Assessment, BEMA)

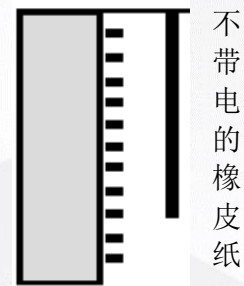
电磁学测试 (The Brief Electricity and Magnetism Assessment, BEMA) 由北卡罗来纳州立大学的Ruth Chabay教授和Bruce Sherwood教授设计，用于测试学生对基本的电磁学概念的理解，适用于有微积分基础的大学生。

设某一不导电的墙上带有负电荷。此时，有一柔软的不带电的橡皮纸悬挂在该带电墙的附近。如图所示。

题7 该橡皮纸

- a) 不会受到墙的作用，因为橡皮纸是绝缘的；
- b) 不会受到墙的作用，因为橡皮纸不带电；
- c) 将会离开墙向外弯曲，因为橡皮纸中的电子与带电墙之间有排斥力；
- d) 将会离开墙向外弯曲，因为橡皮纸中的分子与带电墙作用而产生的极化作用；
- e) 将会靠近墙向内弯曲，因为橡皮纸中的分子与带电墙作用而产生的极化作用；
- f) 以上都不对。

带负电的墙



4. 对学习内容和掌握情况的测试汇总

主题	测试名称
力学	Force Concept Inventory (FCI)
	The Force and Motion Conceptual Evaluation (FMCE)
	The Mechanics Diagnostic Test (MDT)
能量	Energy and Momentum Conceptual Survey
	Energy Assessment
电磁学	The Brief Electricity and Magnetism Assessment (BEMA)
	Conceptual Survey in Electricity and Magnetism(CSEM)
	Determining and Interpreting Resistive Electric Circuits Concepts Test
	The Electric Circuits Concept Evaluation (ECCE)
热力学	Question on Heat and Temperature
波动	Wave Diagnostic Test (University of Maryland)

第九届全国高校物理实验教学研讨会—— 美国物理教育评估方法简介与应用

对学习态度的调查

对学习态度调查

- 学生的学习态度对于教学效果，及能否达到教学目标起着至关重要的作用，学生的学习态度的变化情况也与教学情况有着直接的关系。
- 教师希望了解学生的学习态度，有很多方法如**面谈**、完成**实验**和**测试**等，其中最为直接的方法也是通过调查测试。

用于调查学生学习物理学等科学课程的学习态度的测试

- 物理期望调查（The Maryland Physics Expectations Survey, The MPEX）,
 - 对物理科学的认识信心的调查
（The Epistemological Beliefs Assessment for Physics Science, EBAPS）,
 - 科学观点调查（Views About Sciences Survey, VASS）
 - 科罗拉多大学设计的CLASS
-
- 题目类型为选择题,
 - 适用于对于一个班级等具有一定数量的学生进行测试, 对结构进行分析后, 可以获得对这批学生的学习态度的总体描述, 从而对教学改进和优化起导向作用。

物理期望调查 (The Maryland Physics Expectations Survey, The MPEX)

- MPEX测试于90年代中期由马里兰州立大学的E. F. Redish教授设计，用于测试学生对于科学本质和科学知识的认识和观点。
- 测试总共有34道题目，每题的候选答案是一样的，有五条，即对于这一陈述的态度。测试结果多方面反映了学生的学习态度。
- 这五条候选答案为：
 - 1、完全不同意 (Strongly Disagree) ;
 - 2、不同意 (Disagree) ;
 - 3、不确定 (Neutral) ;
 - 4、同意 (Agree) ;
 - 5、完全同意 (Strongly Agree) 。



Student Expectations in University Physics: *The Maryland Physics Expectations Survey*

Here are 34 statements which may or may not describe your beliefs about this course. You are asked to rate each statement by circling a number between 1 and 5 where the numbers mean the following:

1: Strongly Disagree	2: Disagree	3: Neutral	4: Agree	5: Strongly Agree
----------------------	-------------	------------	----------	-------------------

Answer the questions by circling the number that best expresses your feeling. Work quickly. Don't over-elaborate the meaning of each statement. They are meant to be taken as straightforward and simple. If you don't understand a statement, leave it blank. If you understand, but have no strong opinion, circle 3. If an item combines two statements and you disagree with either one, choose 1 or 2.

1	All I need to do to understand most of the basic ideas in this course is just read the text, work most of the problems, and/or pay close attention in class.	1	2	3	4	5
2	All I learn from a derivation or proof of a formula is that the formula obtained is valid and that it is OK to use it in problems.	1	2	3	4	5
3	I go over my class notes carefully to prepare for tests in this course.	1	2	3	4	5
4	"Problem solving" in physics basically means matching problems with facts or equations and then substituting values to get a number.	1	2	3	4	5

马里兰州立大学的物理期望调查

(The Maryland Physics Expectations Survey, The MPEX) 描述的五个方面的学生学习物理的态度

	赞成	不赞成
独立性	学习是独立的，应形成自己的理解。	接受权威（教师或教材）的观点，不予评价。
一致性	相信物理学具有相互关联的一致的框架。	相信物理学是由分裂的片段组成的。
概念	在掌握概念的基础上加强理解。	注重记忆和使用公式。
与实际的联系	相信概念与大量的实际情况是相关的，并是有用的。	相信学到的物理概念与课堂外的经验是无关的。
与数学的联系	认为数学是用于描述物理现象的有效的方法。	认为物理和数学是独立、不相关的。

**第九届全国高校物理实验教学研讨会——
美国物理教育评估方法简介与应用**

对逻辑推理能力的调查

What is “Scientific Reasoning”?

Domain-general cognitive and meta-cognitive skills needed to conduct open-ended scientific explorations such as the abilities to systematically explore a problem, formulate and test hypotheses, manipulate and isolate variables, and observe and evaluate the consequences.

Scientific Reasoning skill dimensions for assessment:

- Control of Variables
- Proportional Reasoning
- Deductive Reasoning
- Inductive Reasoning
- Abductive Reasoning
(Hypothesis Forming)
- Multi-representational Data Analysis
- Hypothesis-Evidence Evaluation
- Probabilistic Reasoning
- Correlational Reasoning
- Causal Reasoning
- Categorization
- Combination
- Basic Logic
- Inference
- Analogy

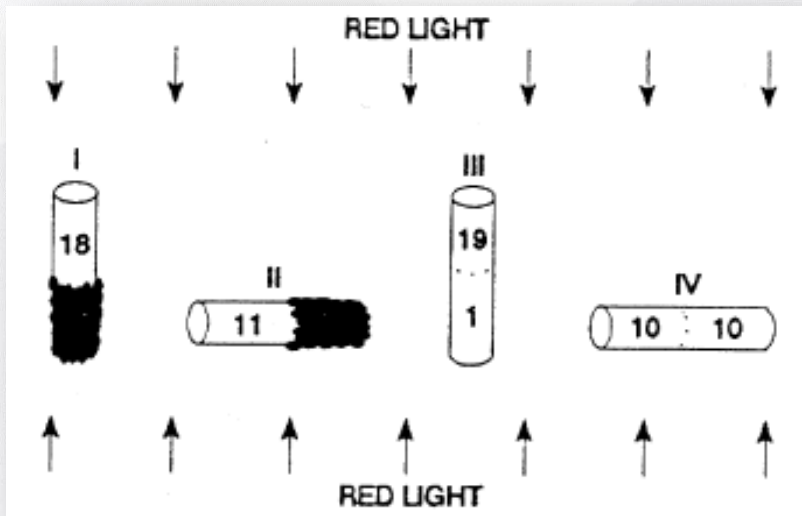
对逻辑推理能力调查

- 学生在学习物理学等科学课程的过程中，不但学到了知识（包括概念、定理、公式和分析问题、解决问题的方法等），他们的逻辑推理能力和思维能力是否也得到了提高，这是也物理教育领域感兴趣的问题。
- 用于测试学生逻辑推理能力的测试：
 - Lawson test,
 - Tools for Scientific Thinking: Mathematical Modeling Conceptual Evaluation,
 - Tools for Scientific Thinking: Vector Evaluation等
- 其中，Lawson test是由加利福尼亚大学的Anton E. Lawson教授于1978年设计的，目前常用的是2000年修订后的版本。

逻辑推理能力调查特点

- 在测试中，会要求学生根据提供的条件进行推理分析，得出结果。
- 考察学生孤立和分析变量、组合推理、盖然论推理和均衡推理的能力。
- Lawson Test在实际应用中的**优点**在于其题目的内容是常识性的和生物学类的，测试的核心是学生的逻辑分析能力，而大多数参加测试的学生是非生物专业的，所掌握的生物学知识都有限，因此测试的结果可以屏蔽专业背景对测试结果的影响。

- **Lawson Test 的第11题：**有20只果蝇分布在四只密封的试管内。试管1和试管2部分地包裹了黑色的纸。试管3和试管4未被包裹。试管如图中位置放置。再用红色的光照射5分钟。果蝇在每根试管上未被覆盖的部分分布的数量如图所示。



- 这个实验显示果蝇会对什么情况作出反应（即移动或离开）？
 - 红光照射，而非重力产生影响。
 - 重力产生影响，而非红光照射。
 - 红光照射和重力都产生影响。
 - 红光和重力都不产生影响。
- **Lawson Test 的第12题：**请分析原因：
 - 试管3中的大多数果蝇位于顶端，而试管2中的果蝇均匀分布。
 - 试管1和试管3中的大多数果蝇不在试管的底部。
 - 果蝇需要光线才能看见，并必须逆着重力方向飞行。
 - 大部分果蝇在试管的上端和光线亮的一端。
 - 一些果蝇位于每根试管的两端。

第九届全国高校物理实验教学研讨会—— 美国物理教育评估方法简介与应用

教 学 应 用

1. Learning and Scientific Reasoning

Science, Vol. 323. no. 5914, pp. 586 – 587 (2009).

Nature Physics: Science education: Lessons to be learned, Vol.6, 6, 2010.

NPR Science Friday: (Interview) Learning Facts vs Learning to Reason. 1/30, 2009

EDUCATIONFORUM

PHYSICS

Learning and Scientific Reasoning

Lei Bao,^{1*} Tianfan Cai,² Kathy Koenig,³ Kai Fang,⁴ Jing Han,¹ Jing Wang,¹ Qing Liu,¹ Lin Ding,¹ Lili Cui,⁵ Ying Luo,⁶ Yufeng Wang,² Lieming Li,⁷ Nianle Wu⁷

The development of general scientific abilities is critical to enable students of science, technology, engineering, and mathematics (STEM) to successfully handle open-ended real-world tasks in future careers (1–6). Teaching goals in STEM education include fostering content knowledge and developing general scientific abilities. One such ability, scientific reasoning (7–9), is related to cognitive abilities such as critical thinking and reasoning (10–14). Scientific-reasoning skills can be developed through training and can be transferred (7, 13). Training in scientific reasoning may also have a long-term impact on student academic achievement (7). The STEM education community considers that transferable general abilities are at least as important for students to learn as is the STEM content knowledge (1–4). Parents consider science and mathematics to be important in developing reasoning skills (15).

We therefore asked whether learning STEM content knowledge does in fact have an impact on the development of scientific-reasoning ability. The scientific-reasoning ability studied in this paper focuses on domain-general reasoning skills such as the abilities to systematically explore a problem, to formulate and test hypotheses, to manipulate and isolate variables, and to observe and evaluate the consequences.

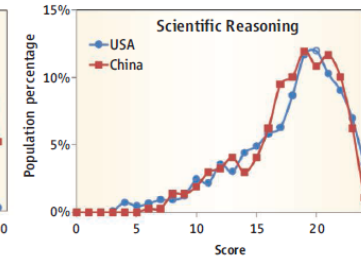
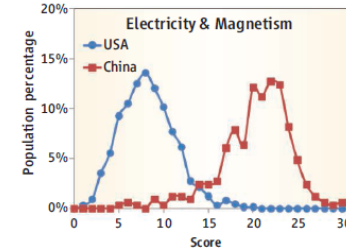
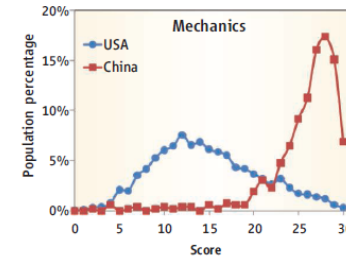
learning has any impact on the development of scientific-reasoning ability. Scientific reasoning is not explicitly taught in schools in either country.

In China, K–12 education is dominated by the nationwide college admission exam given at the end of grade 12. To comply with the requirements of this exam, all Chinese

Comparisons of Chinese and U.S. students show that content knowledge and reasoning skills diverge.

understanding and problem-solving skills are very different in the two countries. Similar curriculum differences between the United States and China are reflected in other STEM areas such as chemistry, biology, and mathematics (16).

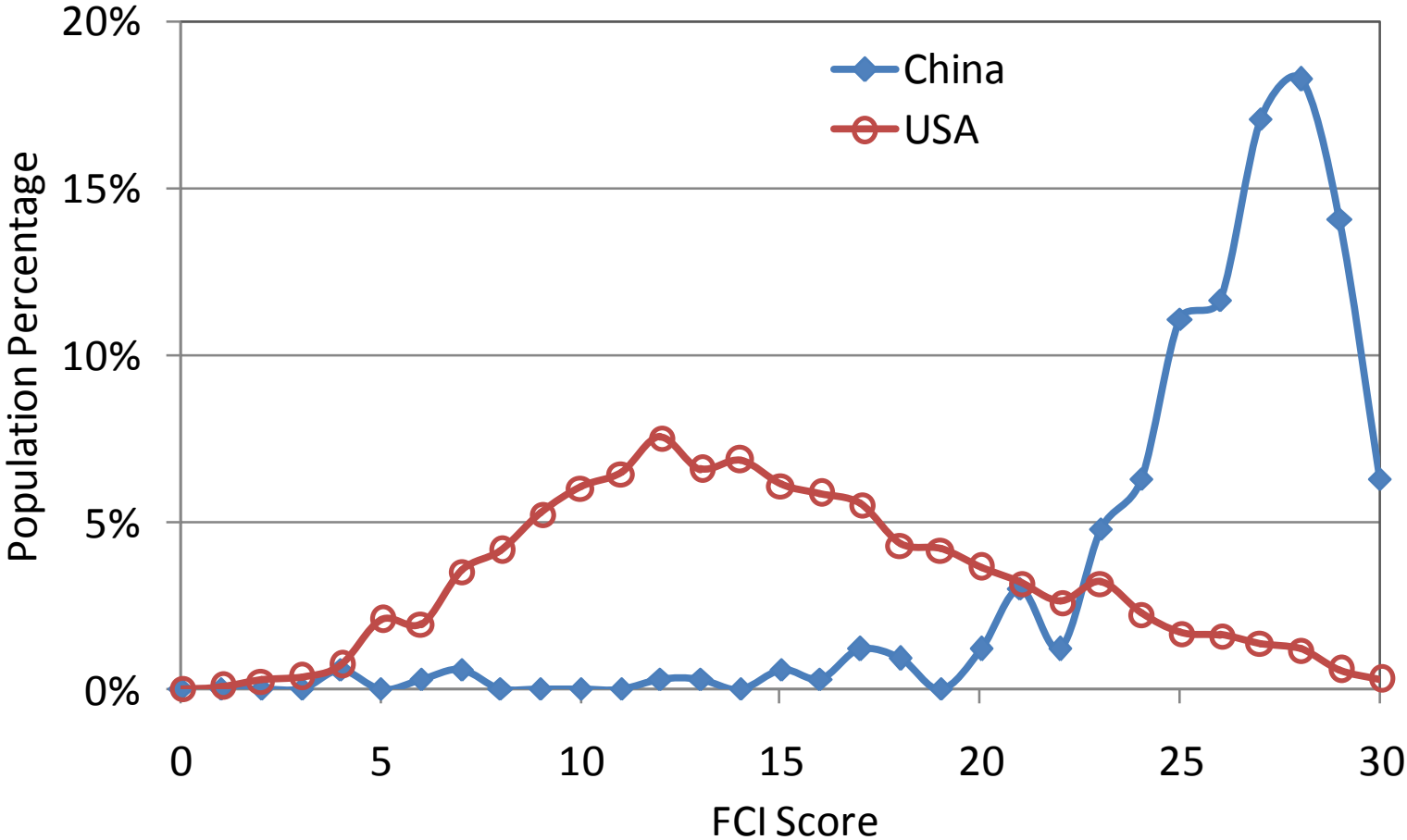
Chinese students go through rigorous problem-solving instruction in all STEM



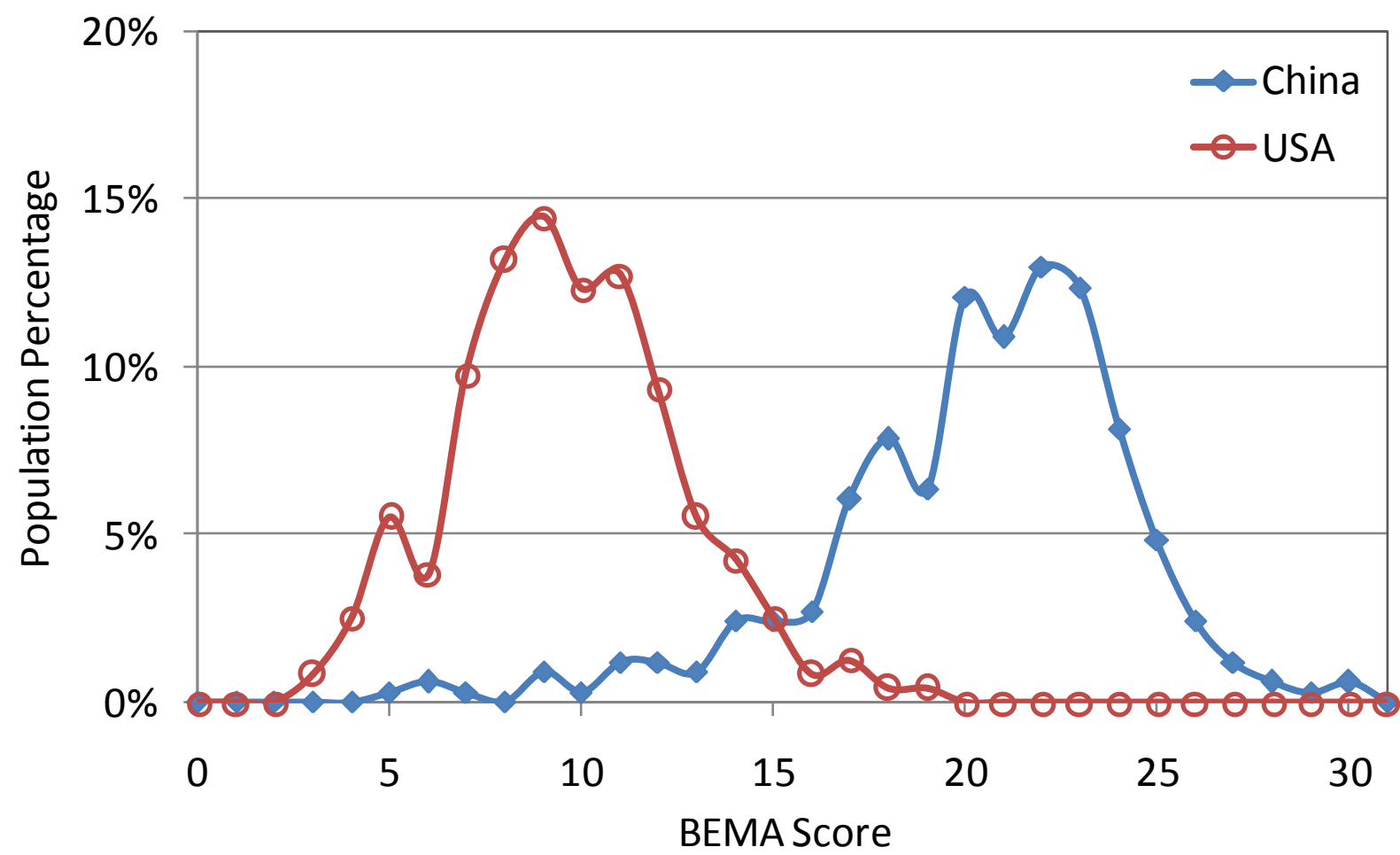
Test	TEST SCORES (%)		
	China (n)	USA (n)	Effect size
FCI	85.9 ± 13.9 (523)	49.3 ± 19.3 (2681)	1.98
BEMA	65.6 ± 12.8 (331)	26.6 ± 10.0 (650)	3.53
LCTSR	74.7 ± 15.8 (370)	74.2 ± 18.0 (1061)	0.03

Content knowledge and reasoning skills diverge. Comparisons of U.S. and Chinese freshmen college students show differences on tests of physics content knowledge but not on tests of scientific reasoning.

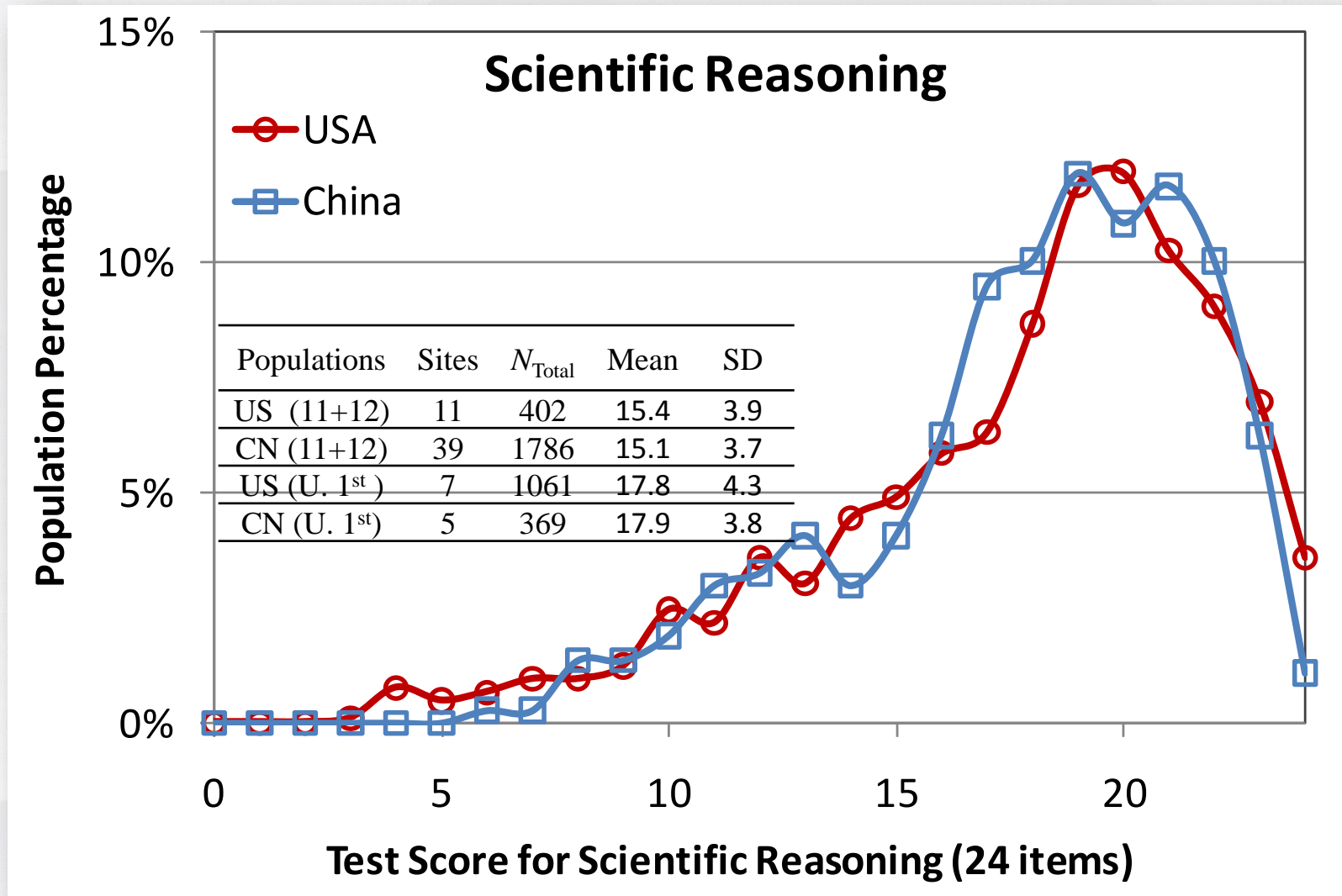
First Year College Students' FCI Results



First Year College Students' BEMA Results



First Year College Students' Lawson Results



2. 微课程混合式教学效果的调查与分析

调查对象

同济大学学习普通物理实验的非物理专业大学一年级学生

调查内容

学生情况、学习态度、学习习惯、教学效果...

问卷发放与回收

物理实验微课程学习系统了解情况调查：

共发放问卷200份，收回有效问卷197份，回收率为98.5%。

物理实验微课程学习系统使用情况调查：

共发放问卷200份，收回有效问卷196份，回收率为98.0%。

➤ 问卷调查结果及分析

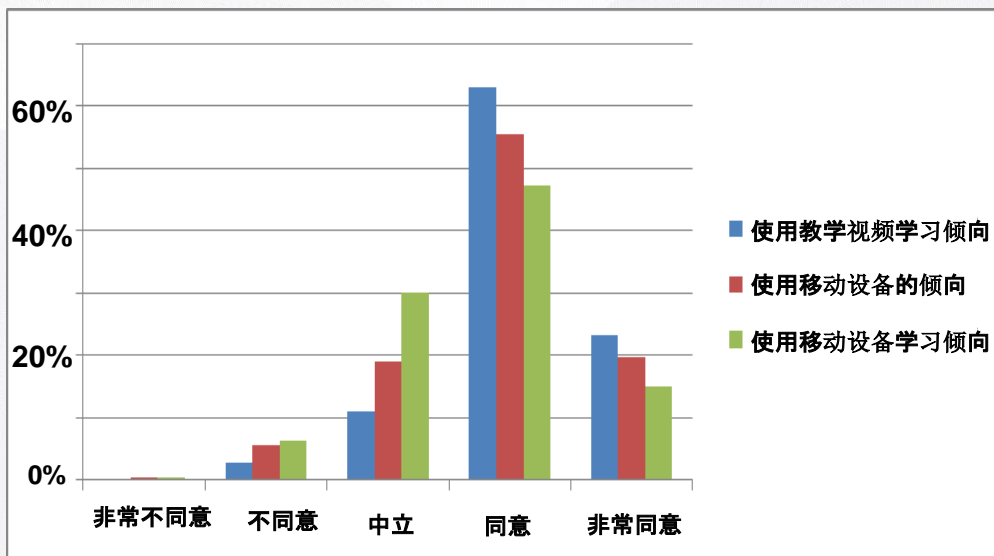


图1 学生使用移动设备的情况及用于实验学习的态度统计

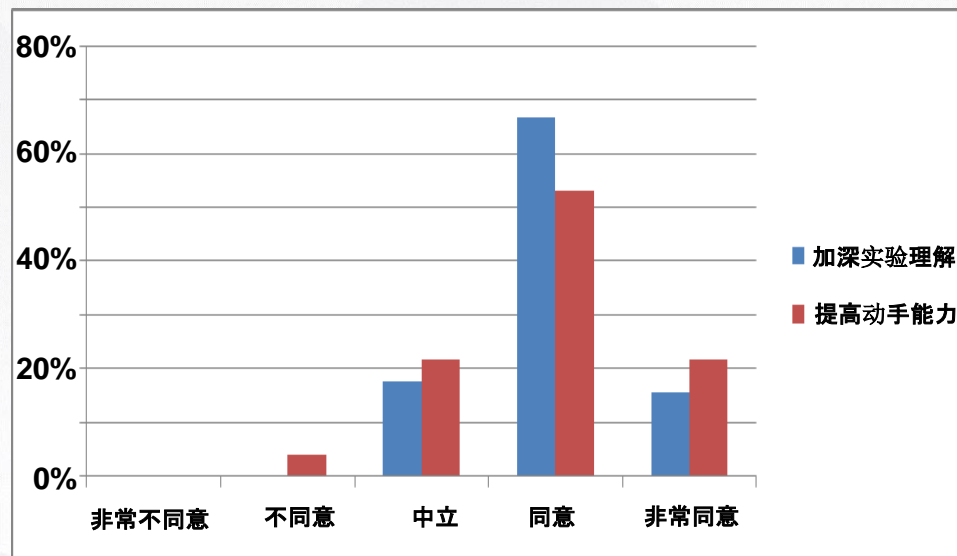


图2 实验微课程对实验的理解和动手能力的影响

➤ 微课程混合式教学效果分析 —— 以扭摆法测定物体转动惯量为例

独立样本t检验

是否在用扭摆法测定物体转动惯量的实验中用微课程		N	均值	标准差	均值的标准误
用扭摆法测定物体转动惯量成绩	否	2521	84.27	7.12	.14
	是	67	86.22	5.84	.71

	方差方程的 Levene 检验		均值方程的 t 检验						
	F	Sig.	t	df	Sig. (双侧)	均值差值	标准误差	差分的 95% 置信区间	
								下限	上限
用扭摆法测定物体转动惯量成绩	3.06	.08	-2.22	2586	.03	-1.95	.88	-3.67	-.23
假设方差相等			-2.68	71.32	.01	-1.95	.73	-3.40	-.50

非常感谢 ！