

# 大学学習効果とその測定

## --測定方法の分類と概括--

木村 拓也

(長崎大学アドミッションセンター

同志社大学高等教育・学生研究センター)

# 話題構成

- ラーニングアウトカムズ(LO)の測定に関する各種の取り組みの概要と分類
- ラーニングアウトカムズ(LO)の測定方法が有する利点と限界
- ラーニングアウトカムズ(LO)の測定に関する取り組みと、学術会議で策定しようとする参照基準との関係

# 共通テストによる取り組みの概要I

	GRE	GRE Subject	法学 検定	経済学 検定	医療系共用 試験CBT	医療系共用 試験OSCE
実施機関	Educational Testing Service	Educational Testing Service	財団法人 日弁連 法務研究財団 / 社団法人 商事法務研究会	日本経済学 教育協会	医療系大学間共用試験 実施評価機構(CATO)	
設問数	analytical writing:2題, verbal:30問 Quantitative:28問	生化学、細胞・分子生物学180 問、生物学200問、化学130 問、コンピュータサイエンス 70問、英文学230問、数学66 問、物理学100問、心理学205 問	2級：55問、3級：75 問、4級：60問、既修 者：125～175問	ERE：90問 EREミクロマクロ ：50問	320問 (うち80問は採点除外)	標準6 ステーション
回答時間	analytical writing:45/30 分, verbal:30, Quantitative:45分	170分	2級：150分、3級： 150分、4級：120分、 既修者：150分から 210分	ERE180分 EREミクロマクロ ：100分	計6時間	1ステーション 約5分～10分
設問形式	verbal, quantitative: CBT 多肢選択	マークシート 多肢選択	正誤問題 択一問題	多肢選択	CBT 多肢選択	パフォーマンス テスト
実施回数	毎日実施 (但し、年5回まで)	年3回	年1回	年2回	年2回(但し、臨床実習開始時期に合わせたも のであって、複数回受験ではない)	
受験者数	analytical writing:130万, verbal:130万, Quantitative: 130万(2004-7)	生化学、細胞・分子生物学6 千、生物学12千、化学8千、コ ンピュータサイエンス6千、英 文学11千、数学10千、物理学 13千、心理学26千(2004-7)	2級：1647、3級： 4674、4級：4045、既 修者：3099(2008年)	1418(14回) 1626(15回)	7445(2007医学) 2144(2007歯学)	7428(2007医学) 2545(2006歯学)
評価方法	スコアのみ 通知(IRT)	スコアのみ 通知	スコアのみ 通知	ランク判定 (偏差値)	6段階評価 (IRT)	6段階評価
受験料	\$140(米国内)	\$130(米国内)	2級：¥12600、3級： ¥6300、4級： ¥4200、既修者： ¥12600	ERE: ¥5250 EREミクロマクロ ：¥3150	¥28,000	

# 共通テストによる取り組みの概要2

	GRE	GRE Subject	法学 検定	経済学 検定	医療系共用 試験CBT	医療系共用 試験OSCE
内容	verbal reasoning quantitative reasoning analytical writing	生化学・ 細胞・分子生物学、 生物学、化学、 コンピュータサイエ ンス、英文学、 数学、物理学、 心理学	2級：法学基礎論・民法・ 憲法・刑法・民事訴訟法・ 国際取引法・倒産・執行・ 独占禁止法・租税法・労働 法・環境法、3級：法学一 般・民法・憲法・刑法・民 事訴訟法・行政法・商法・ 労働法、4級：法学入門・ 憲法・民法・刑法、既修 者：必須 憲法・民法・刑 法・民事訴訟法・刑事訴訟 法、任意 商法・行政法	ミクロ経済学 マクロ経済学 財政学 金融論 国際経済 統計学	基礎・臨床の知識の総合的 理解と問題解決能力をコン ピュータを用いた客観試験 (CBT)で評価 例)医学系 基本事項5%、医学一般 20%、人体各器官の正常構 造と機能、病態、診断、 治療40%、全身におよぶ生 理的变化、病体診断治療 15%、診療の基本10%、医 学・医療と社会10%	態度・臨床技能を客観的 臨床能力試験 (OSCE) で評価 例) 医学系 医療面接(10分)、頭頸部 診察(5分)、胸部診察(5 分)、腹部診察(5分)、神 経診察(5分)、救急(5分)
目的	特定の研究分野に関係 しない、言語的推論、 数量的推論、批判的思 考、分析的記述力を測 ることが目的。	上記の8分野における 学部時代の学業到達度 を測ることが目的。当 該分野を専攻している 人か、或いは、様々な 分野に心得のある人が 対象。	法学に関する学力水準 を客観的に評価する、 わが国唯一の全国規模 の検定試験。	主として経済学部および 社会科学系学部の学生を 対象に、全国規模で経済 学の数理的・理論的な基 礎知識の習得程度と実体 経済での応用能力のレベ ルを判定する試験。	学生の能力と適性について の一定水準を確保 するために実施する全国共通 の標準評価試験。臨床実習開始 前の共用試験により、態 度・基本的臨床技能・知識の 総合的理解力と 問題解決能力の評価を行い、 社会の求める優 れた医師・歯科医師の要請に 応える。	
利用	大学院、ビジネススクール、 プロフェッショナルスクールなど の進学時にスコアを提出		法学に関する学力を客観的 に評価する唯一の試験とし て、大学での単位認定、企 業の入社・配属時等の参考 資料「法学既修者試験(法 科大学院既修者試験)」 は、各法科大学院の「既修 者コース(2年コース)」 を目指す受験者のために実 施しているもので、法学既 修者であることを証明す るための資料として、多く の法科大学院で提出	大学院入試への活用(代 用)、就職活動において 成績証明書等の提出、単 位認定・カリキュラム導 入、過去問を利用したE RE対策講座・公務員試 験対策、向上心に富む学 生を後押しするバック アップ体制(受験料全額 負担、学長表彰、学部表 彰、報奨金制度など)	標準評価試験としての共用試験(CBT、 OSCE)の成績を、各大学固有の成績と併 せて総合的な臨床実習開始前の成績評価・進級 判定等の指標として利用する。共用試験成績 の評価基準は、各大学の臨床実習の在り方に 基づき、各大学が責任をもって設定する。	

- 註)各実施機関のHPより抜粋

# 学習モニタリングによる取り組みの概要

	<b>CSS</b> College Senior Survey	<b>TFS</b> The Freshman Survey	<b>NSSE</b> National Survey of Student Engagement	<b>JCSS</b> Japan College Student Survey	<b>JFS</b> Japan Freshman Survey	全国 大学生調査
実施主体	<b>HERI</b> カリフォルニア大学 ロサンゼルス校 高等教育研究所	<b>HERI</b> カリフォルニア大学 ロサンゼルス校 高等教育研究所	<b>CPR</b> インディアナ大学 ポスト中等教育 研究センター	同志社大学 高等教育・ 学生研究センター	同志社大学 高等教育・ 学生研究センター	<b>CRUMP</b> 東京大学 大学経営・ 政策研究センター
目的・内容	4年生の出口調査として、大学のアウトカムと卒業後の目標や計画など広範囲の質問を行う。例) 学業成績と大学への関与、認知的・情緒的発達、学生の価値観・態度・目標、大学経験の満足度、卒業後の計画など	大学入学前に新入生がどんな経験をしてきたのかについての情報を収集し、2つの追跡調査CSSとYFCFに連携させて、大学時代の認知的・情緒的成長の経年変化を吟味することが目的である。	学業面でのチャレンジ、学習における積極性と協調性、学生と教員の交流、教育経験の向上、大学の支援環境、という5つのベンチマークについての全国規模の調査	自大学の学生の特徴、学習行動、生活行動など学び習慣を把握することに加えて、学び習慣の根底にもなる学生自身の自己評価、価値観や満足度が分かる。ベンチマークとしての利用も可能。	学生の成長を把握する出発点として、新入生の高校時代の学習行動や生活行動を把握し、大学での適応について知る	大学でどのように生活・学習しておられるのかを明らかにし、日本の大学教育のあり方を考える基礎とするとともに、皆さんが在学する大学での教育改善にも役立てることを目的
設問数	約270	約370	約50	約400	約250	約140
受験者数	400,000 (毎年)	240,580 (2008)	56,545 (2000)	1,400(2004試行) 3,961(2005) 6,228(2007) 実施予定(2009)	19,661(2008) 実施予定(2009)	44,905 (2006)
調査対象 実施時期	4年次生 11月～	1年次生 3月～	1年次と4年次 5～6月	1年～4年次生 11～12月	主として1年次 6～7月	1年～4年次生 1月

- 註)各実施機関のHPより抜粋



分類基準	例
(1)測定内容のちがいによる分類	知能検査, 学力検査, 適性検査, 感覚運動検査, 興味検査, 性格検査, 価値・態度尺度
(2)測定目的のちがいによる分類	記述, 選抜, 診断, 分類, 予測, 動機づけ, 教育訓練のための検査
(3)被測定体のちがいによる分類	自己検査(学力検査, 作業検査, 自記式性格検査等) 他者評定(行動評定, 性格評定, ゲス・フー・テスト等)
(4)測定人数のちがいによる分類	個人検査(1検査者対1被検者) 集団検査(1検査者対多数被検者)
(5)測定用具のちがいによる分類	紙筆検査(紙と鉛筆による質問形式の検査) 器具検査(器具を使つての作業検査, 感覚運動検査)
(6)測定行動パターンのちがいによる分類	最高値検査(自己の持てる最高の力を発揮して答える. 学力・作業検査) 典型値検査(最もふつうにみられる典型的行動を答える. 性格評定検査)
(7)回答形式のちがいによる分類	客観テスト(評価が客観的にできる. 機械採点も可能) 論文・作品テスト(論文・絵・作曲等の作品の評価) 投映法テスト(具体性をもたない刺激に対する反応・動作の評価)
(8)回答の量・質のちがいによる分類	速度検査(作業のスピードが問題, どの位早く計算ができるか) 力量検査(回答のレベルが問題, どの位むずかしい計算ができるか)
(9)検査作成手続きのちがいによる分類	標準検査(規準集団, 検査手続が規格化されたテスト)(池田1971『行動科学の方法』) 個別作成検査(必要に応じて新しく作られたテスト) 東京大学出版会、p.37)

# 大学学習成果の測定方法の分類

outputの種類	GRE	GRE Subject	法学検定	経済学検定	医療系CBT	医療系OSCE	CSS/TFS	NSSE	JCSS/JFS	CRUMP調査	卒論	GPA	国家試験合格率	院進学率	授業評価
結果利用	個人 [直接評価]	個人 [直接評価]	個人 [直接評価]	個人 [直接評価]	個人 [直接評価]	個人 [直接評価]	機関 [間接評価]	機関 [間接評価]	機関 [間接評価]	機関 [間接評価]	個人 [直接評価]	個人 [直接評価]	機関 [間接評価]	機関 [間接評価]	個人 (教師)
測定目的	選抜 診断	選抜 診断	選抜 診断	選抜 診断	選抜 診断	選抜 診断	評価 把握	評価 把握	評価 把握	評価 把握	選抜 診断	選抜 診断	選抜 診断	選抜 診断	診断 (教師)
被測定体	自己 検査	自己 検査	自己 検査	自己 検査	自己 検査	自己 検査	自己 検査	自己 検査	自己 検査	自己 検査	自己検査/ 他者評価	自己検査/ 他者評価	自己 検査	自己 検査	他者 評価
測定内容	適性 検査	学力 検査	学力 検査	学力 検査	学力 検査	適性検査 態度尺度	興味検査、 価値尺度、 態度尺度	興味検査、 価値尺度、 態度尺度	興味検査、 価値尺度、 態度尺度	興味検査、 価値尺度、 態度尺度	学力 検査	学力 検査	学力 検査	学力 検査	適性 検査
回答量・質	speed /power	speed	speed	speed	speed	speed	speed	speed	speed	speed	power	speed /power	speed /power	speed /power	speed
回答形式	論文/ 客観	客観	客観	客観	客観	パフォーマンス	客観	客観	客観	客観	論文	論文/ 客観	論文/ 客観	論文	客観
評価形式	到達度 評価	到達度 評価	到達度 評価	相対 評価	到達度 評価	到達度 評価	総括的 診断的 評価	総括的 診断的 評価	総括的 診断的 評価	総括的 診断的 評価	絶対 評価	相対 評価	相対 評価	相対 評価	総括的 評価

- 組み合わせは、多種多様。各分野・各大学の状況に合わせてが妥当な線か？

# 話題構成

- ラーニングアウトカムズ(LO)の測定に関する各種の取り組みの概要と分類
- ラーニングアウトカムズ(LO)の測定方法が有する利点と限界
- ラーニングアウトカムズ(LO)の測定に関する取り組みと、学術会議で策定しようとする参照基準との関係



# 共通テストと学習モニタリングの利点と限界

- 相互補完的な関係

## 利点

カリキュラムベース  
目標標準型学習の促進  
動機付けの容易さ

## 利点

学生プロフィールの把握

JCSS 5群(大学エンジョイ群、学業充実群、不本意入学奮起群、  
入学後失望群、不本意入学諦め群)

JFS 6群(無目的型、探求学習経験型A[無理解入学]、高校指導従順  
型、受験勉強型A[他律型]、探求学習経験型B[本命入学]、受験勉強型B  
[自律型])

自大学 学生の得意/不得意分野の把握  
(大学生の認知発達順序の把握)

## 限界

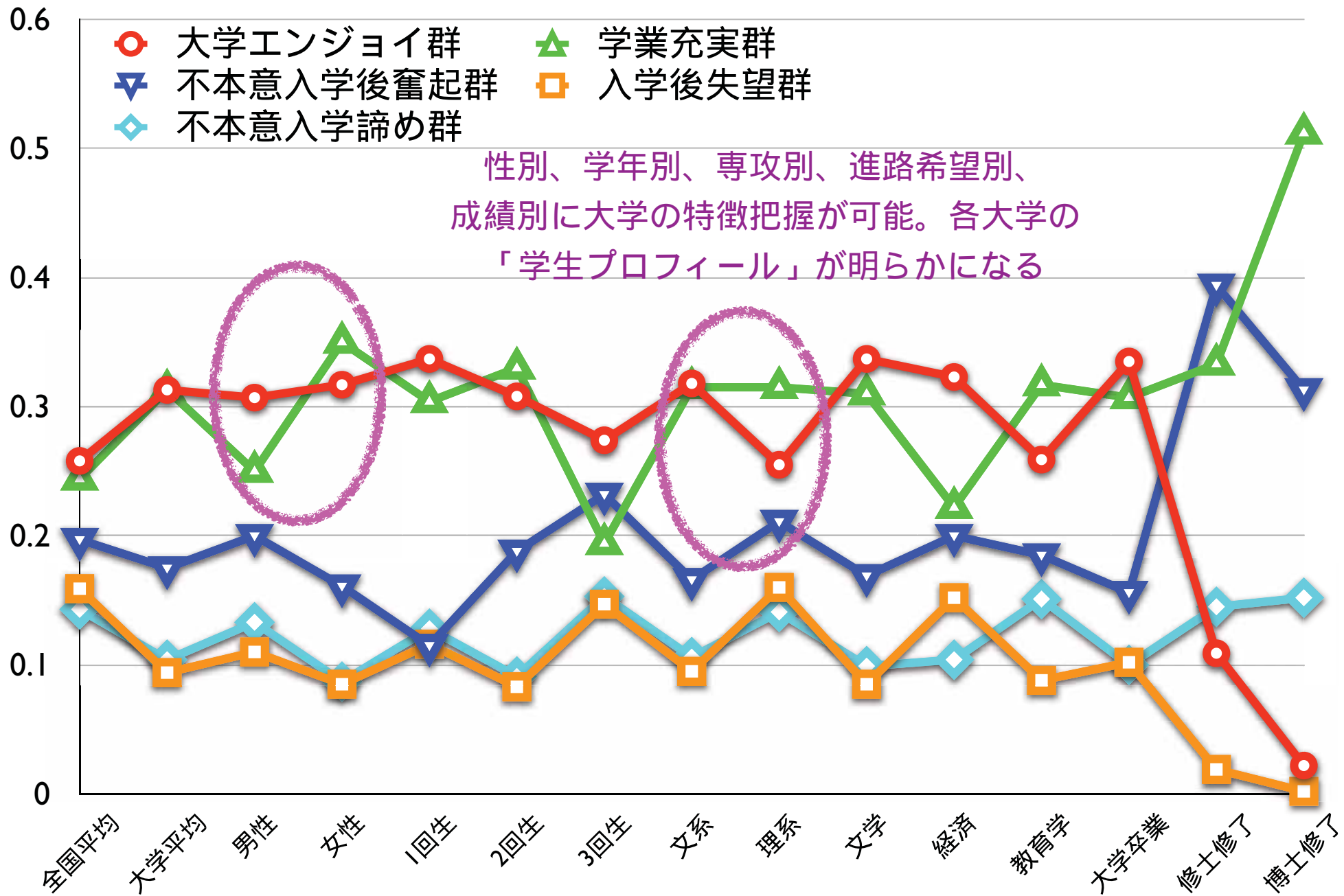
プロセスの評価ではない  
結果の評価でしかない  
スピードテストになりやすい(卒論:重要)  
大人数になればなるほど労力「大」

## 限界

個人の測定ではない  
機関の評価でしかない、悉皆は無理  
あくまで自己申告、外部規準の必要性  
high-stakesではない、データの信用性

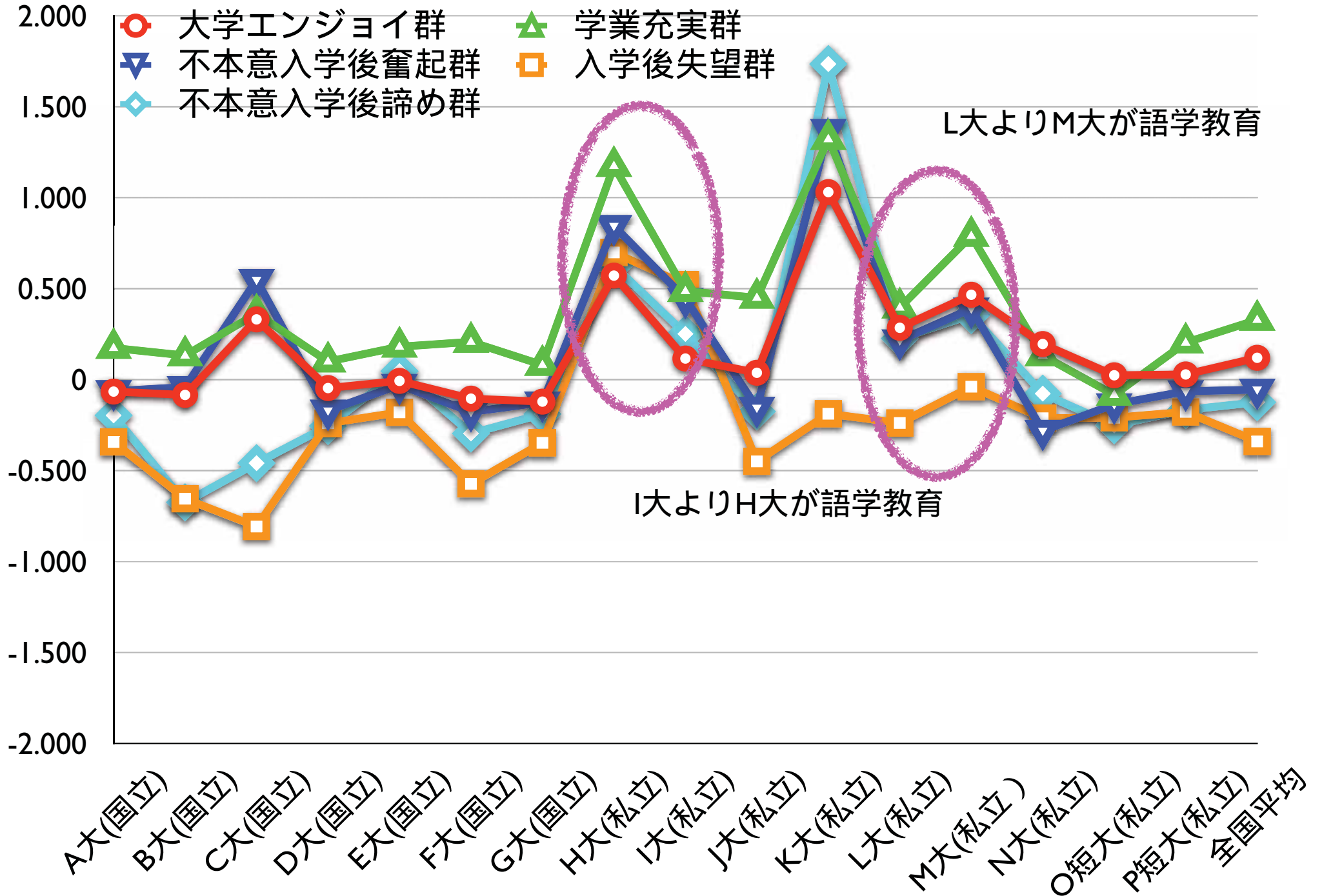
割合

# M大学の学生プロフィール



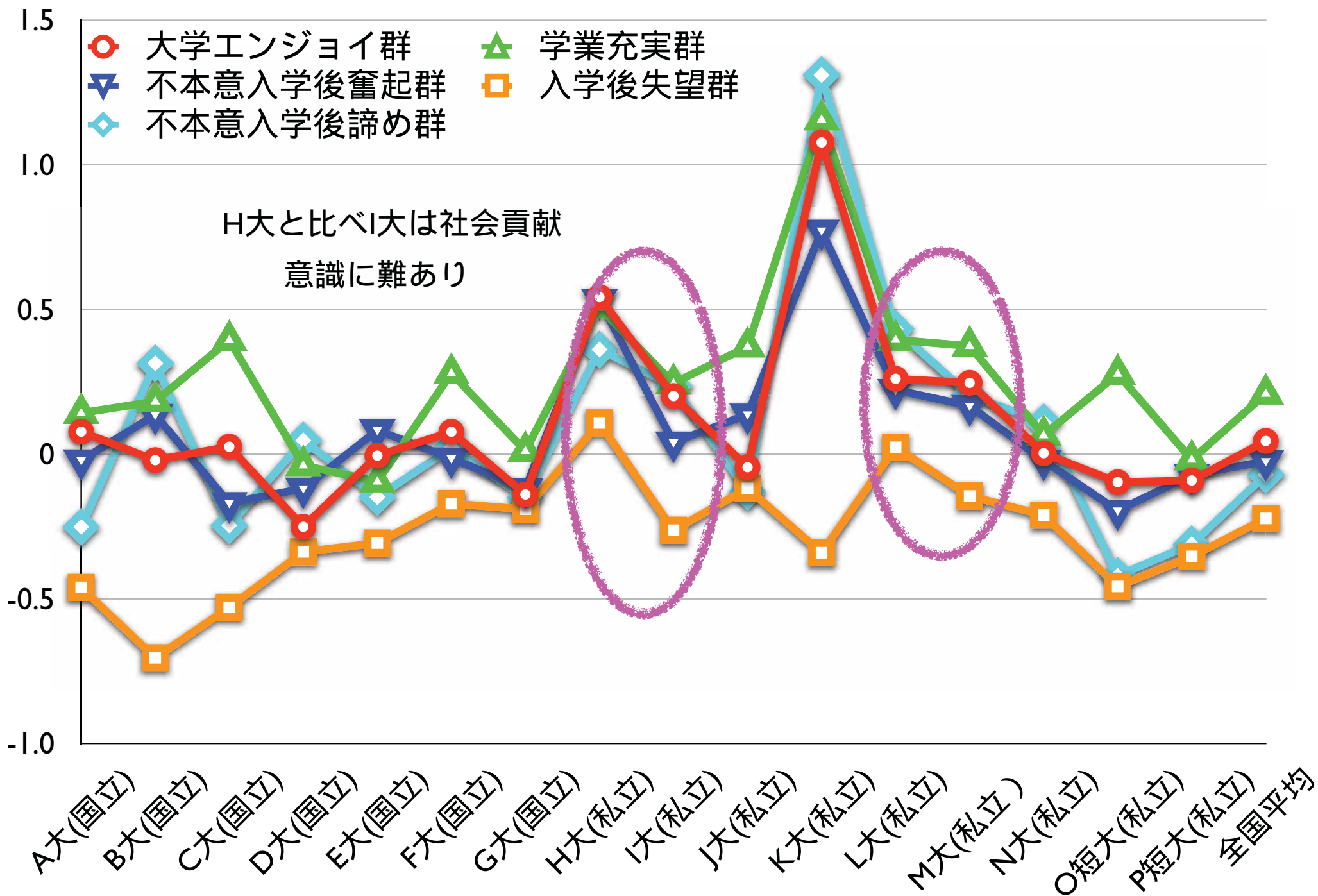
因子得点

異文化コミュニケーション力



因子得点

社会貢献

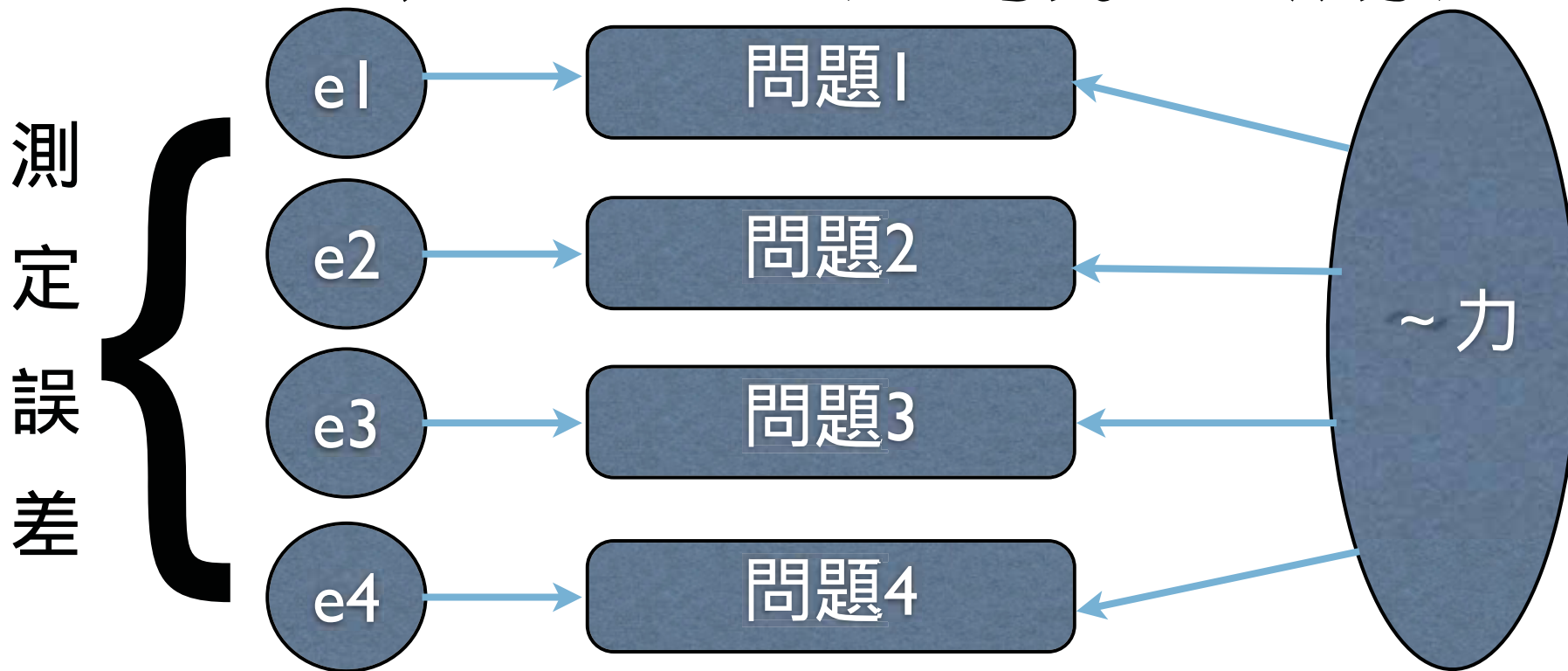


# 話題構成

- ラーニングアウトカムズ(LO)の測定に関する各種の取り組みの概要と分類
- ラーニングアウトカムズ(LO)の測定方法が有する利点と限界
- ラーニングアウトカムズ(LO)の測定に関する取り組みと、学術会議で策定しようとする参照基準との関係



# 「～力」の定義と測定



高いテスト得点となる → ～力が高い



「～力」の命名(定義)は帰納的のはず

ずれ↑↓

測定モデル

高いテスト得点である ← ～力が高ければ

政策・教育論

測定方法を開発せよ ← ～力があるはず

測定の要求はいつも一方的(演繹的)

# 分野別参照基準と

## ラーニングアウトカムズの測定

- 策定予定の分野別参照規準→ラーニングアウトカムズ(LO)の測定と連動ではない：賛成
- 例)「学士力」の【仮想】測定  
学士課程教育で身に付けて欲しい学生像  
→ 結局、GPAの高い「学業充実群」が最も身に付けている能力と一致 (添付資料参照)
- GPAで測っているものを改めて細かく測定するに過ぎない→二度手間になる
- 各LOの測定方法の特徴を踏まえ、各分野・各大学ごとに取捨選択し、各大学の自主的・自律的な教育課程の編成が望ましい

# まとめにかえて

## 【測定業界が抱える実状】

- 今後の質保証に「テストの専門家」は必須
- 日本はその養成に失敗
- 世界では「テストの専門家」を巡る争奪戦
- 現在は、必要最低限の配置 & ぎりぎりの状況
- 「テストの専門家」養成機関の設立のお願い
- 統計的に間違った測定結果の粗製濫造は回避すべき
- 確固とした技術水準に適った測定・評価ができる環境の構築が必要

HOME PAGE MY TIMES TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

# The New York Times

## Education

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION

EDUCATION WASHINGTON

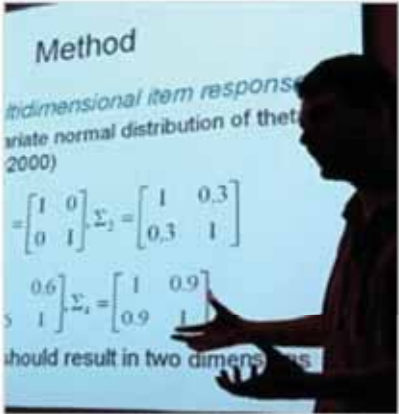
Intel has brought together over 1,500 students w  
research will lead to tomorrow's breakthroughs.

### As Test-Taking Grows, Test-Makers Grow Rarer

By DAVID M. HERSZENHORN  
Published: May 5, 2006

Sz-Shyan Wu is not a Cuban baseball star or a dissident musician. But in urging the United States government to grant him a work visa, the New York State Education Department is arguing that Mr. Wu, too, has talents so rare that bureaucracy must be cut and a red carpet rolled out.

Enlarge This Image



Mr. Wu is a psychometrician or, in plain English, an expert on testing. And testing experts are in high demand.

With federal law requiring wider testing of schoolchildren, the nation faces a critical shortage of people like Mr. Wu with the mathematical, scientific, psychological and educational skills to create tests and analyze the results. The problem has sent states, testing companies and big school districts into a heated hiring competition, with test companies offering salaries as high as \$200,000 a year or more plus perks.

A result is a peculiar outcome of the No Child Left Behind act. Psychometrics, one of the most obscure, esoteric and

SIGN IN TO E-MAIL THIS

PRINT

SAVE

ARTICLE TOOLS SPONSORED BY

(500) DAYS OF SUMMER

Carmel Zucker for The New York Times

Jonathan Weeks, a graduate student at the University of Colorado, presenting a measurement-research project this week.

# T大学大学院教育学研究科 修士論文の専門領域の変遷

	教育測定 統計	教育哲学・史	教育社会学	学校教育学	教育行政学	社会教育	図書館	教育心理学	臨床心理学	特殊教育
1954d 1959	2(1.9%)	<b>26(25.2%)</b>	7(6.8%)	18(17.5%)	13(12.6%)	13(12.6%)	1(1.0%)	17(16.5%)	3(2.9%)	3(2.9%)
1960d 1964	0(0.0%)	14(14.9%)	12(12.8%)	19(20.2%)	12(12.8%)	6(6.4%)	2(2.1%)	<b>25(26.6%)</b>	3(3.2%)	1(1.1%)
1965d 1969	2(1.6%)	<b>27(20.9%)</b>	15(11.6%)	21(16.3%)	21(16.3%)	12(9.3%)	0(0.0%)	24(18.6%)	4(3.1%)	3(2.3%)
1970d 1974	2(1.4%)	23(16.0%)	20(13.9%)	20(13.9%)	29(20.1%)	12(8.3%)	2(1.4%)	<b>31(21.5%)</b>	4(2.8%)	1(0.7%)
1975d 1979	1(0.8%)	<b>22(16.8%)</b>	17(13.0%)	21(16.0%)	<b>22(16.8%)</b>	9(6.9%)	7(5.3%)	15(11.5%)	11(8.4%)	6(4.6%)
1980d 1984	2(1.5%)	23(17.2%)	17(12.7%)	<b>36(26.9%)</b>	12(9.0%)	9(6.7%)	6(4.5%)	19(14.2%)	9(6.7%)	1(0.7%)
1985d 1989	5(3.2%)	27(17.2%)	19(12.1%)	<b>36(22.9%)</b>	10(6.4%)	12(7.6%)	5(3.2%)	32(20.4%)	8(5.1%)	3(1.9%)
1990d 1994	9(6.0%)	22(14.6%)	22(14.6%)	<b>33(21.9%)</b>	10(6.6%)	14(9.3%)	7(4.6%)	22(14.6%)	11(7.3%)	1(0.7%)
1995d 1999	9(4.1%)	29(13.1%)	39(17.6%)	<b>44(19.9%)</b>	21(9.5%)	17(7.7%)	10(4.5%)	33(14.9%)	18(8.1%)	1(0.5%)
2000d 2005	5(1.7%)	41(14.2%)	52(18.0%)	<b>60(20.8%)</b>	20(6.9%)	28(9.7%)	11(3.8%)	30(10.4%)	39(13.5%)	3(1.0%)
修了者合計	<b>37(2.4%)</b>	254(16.4%)	220(14.2%)	308(19.8%)	170(10.9%)	132(8.5%)	51(3.3%)	248(16.0%)	110(7.1%)	23(1.5%)



昭29年	測定統計	哲学・史	教育社会	学校教育	教育行政	社会教育	教育心理	臨床心理	特殊教育	合計
教授	7	69	8	25	19	7	55	1	2	193
助教授	4	105	14	26	15	9	91	2	6	272
講師	6	42	11	15	8	10	54	11	15	172
総計	17	216	33	66	42	26	200	14	23	637
昭40年	測定統計	哲学・史	教育社会	学校教育	教育行政	社会教育	教育心理	臨床心理	特殊教育	合計
教授	5	114	20	20	13	7	96	6	11	292
助教授	6	100	24	31	28	14	118	7	15	343
講師	1	19	5	0	7	1	26	4	14	77
助手	0	1	0	1	0	0	3	0	0	5
総計	12	234	49	52	48	22	243	17	40	717
昭50年	測定統計	哲学・史	教育社会	学校教育	教育行政	社会教育	教育心理	臨床心理	特殊教育	合計
教授	3	118	37	27	31	17	130	7	56	426
助教授	4	79	20	18	25	16	114	6	79	361
講師	0	20	9	2	8	1	37	0	27	104
助手	0	7	1	1	0	0	18	0	11	38
総計	7	224	67	48	64	34	299	13	173	929
平12年	測定統計	哲学・史	教育社会	学校教育	教育行政	社会教育	教育心理	臨床心理	特殊教育	合計
教授	4	93	51	59	34	27	131	57	127	583
助教授	1	59	35	48	19	24	106	48	83	423
講師	1	15	10	10	4	5	11	10	15	80
助手	0	4	9	9	1	4	13	6	7	54
総計	6	171	105	126	58	60	261	121	232	1140



ご静聴有り難うございました

[kimura-t@nagsaki-u.ac.jp](mailto:kimura-t@nagsaki-u.ac.jp)