

2009年度 統計関連学会連合大会市民講演会

学力調査と統計

～全国学力・学習状況調査の現状と統計的側面からの検討～

「大規模調査の今後の展望」

安野 史子

(国立教育政策研究所 教育課程研究センター 基礎研究部)

2009.9.6

「大規模調査の今後の展望」

1. はじめに
2. 国際学力調査の概要
3. 日本における全国規模テストの概要
－ 3つの全国規模調査を中心に－
4. 海外での事例: 全米学力調査(NAEP)
5. 日本における全国規模調査の結果比較
6. おわりにー今後の展望ー ???

はじめに

国際学力調査

	TIMSS	PISA
実施主体	国際教育到達度評価学会 (IEA) (1960年設立の国際学術研究団体、60カ国/地域の教育研究機関より構成)	経済協力開発機構 (OECD) (1960年設立の政府間機関、30カ国より構成)
参加国	4年生:25カ国/地域 8年生:46カ国/地域	57カ国/地域 (OECD加盟30カ国、非加盟27カ国/地域) ※PISA2006 41カ国/地域(加盟30カ国、非加盟11カ国/地域)PISA2003 32カ国/地域(加盟28カ国、非加盟4カ国/地域)PISA2000
調査対象	4年生および8年生 (TIMSS2007:4年生約16万人、8年生約22万人) (日本では小4, TIMSS2007:148校4,487人 中2, TIMSS2007:146校4,312人)	15歳児 (PISA2006:約40万人) (日本では高校1年生,PISA2006:約6000人)
調査項目	算数・数学、理科	読解力、数学的リテラシー、科学的リテラシー、問題解決能力(2003年～) ERA(Electronic Reading Assessment)(2009)
調査内容	学校のカリキュラムで学んだ知識や技能等がどの程度習得されているかを評価(選択肢が中心) 問題、生徒(児童)質問紙、教師質問紙	知識や技能等を実生活の様々な場面で直面する課題にどの程度活用できるかを評価(記述式が中心) 筆記型テスト、(コンピュータ型テスト、)生徒質問紙、学校質問紙
調査サイクル	4年 (2003年以降)	3年(中心分野:読解力⇒数学⇒科学⇒)



TIMSS 平均得点の国際比較

数学	小学校		中学校		理科	小学校		中学校	
第1回:1964 (FIMS)			2位 /12カ国	中1	第1回:1970 (FISS)	1位 /16カ国	小5	1位 /18カ国	中2
第2回:1981 (SIMS)			1位 /20カ国	中1	第2回:1983 (SISS)	1位 /19カ国	小5	2位 /26カ国	中2
第3回:1995 TIMSS1995	3位 /26カ国	小4	3位 /41カ国	中2	第3回:1995 TIMSS1995	2位 /26カ国	小4	3位 /41カ国	中2
第3回第2段階 TIMSS1999	-	-	5位 /38カ国	中2	第3回第2段階 TIMSS1999	-	-	4位 /38カ国	中2
TIMSS2003	3位 /25カ国	小4 565 点	5位 /46カ国	中2 570 点	TIMSS2003	3位 /25カ国	小4 543 点	6位 /46カ国	中2 552 点
TIMSS2007	4位 /36カ国	小4 568 点	5位 /48カ国	中2 570 点	TIMSS2007	4位 /36カ国	小4 548 点	3位 /48カ国	中2 554 点



OECD/PISA 平均得点の国際比較 読解力

	2000		2003		2006	
①	フィンランド	546	フィンランド	543	韓国	556
②	カナダ	534	韓国	534	フィンランド	547
③	ニュージーランド	529	カナダ	528	香港	536
④	オーストラリア	528	オーストラリア	525	カナダ	527
⑤	アイルランド	527	リヒテンシュタイン	525	ニュージーランド	521
⑥	韓国	525	ニュージーランド	522	アイルランド	517
⑦	イギリス	523	アイルランド	515	オーストラリア	513
⑧	日本	522	スウェーデン	514	リヒテンシュタイン	510
⑨	スウェーデン	516	オランダ	513	ポーランド	508
⑩	オーストリア	507	香港	510	スウェーデン	507
⑭			日本	498		
⑮					日本	498



OECD/PISA 平均得点の国際比較 数学的リテラシー

	2000	2003	2006
① 日本	557	香港	550
② 韓国	547	フィンランド	544
③ ニュージーランド	537	韓国	542
④ フィンランド	536	オランダ	538
⑤ オーストラリア	533	リヒテンシュタイン	536
⑥ カナダ	533	日本	534
⑦ スイス	529	カナダ	532
⑧ イギリス	529	ベルギー	529
⑨ ベルギー	520	マカオ	527
⑩ フランス	517	スイス	527



OECD/PISA 平均得点の国際比較 科学的リテラシー

	2000	2003	2006
① 韓国	552	フィンランド	548
② 日本	550	日本	548
③ フィンランド	538	香港	539
④ イギリス	532	韓国	538
⑤ カナダ	529	リヒテンシュタイン	525
⑥ ニュージーランド	528	オーストラリア	525
⑦ オーストラリア	528	マカオ	525
⑧ オーストリア	519	オランダ	524
⑨ アイルランド	513	チェコ	523
⑩ スウェーデン	512	ニュージーランド	521

日本における全国規模テスト

	小学校 (学力調査)	中学校 (学力調査)	高等学校	
			(学力調査)	(学力検査)
昭和20～ (1945)				進学適性検査
昭和30年～ (1955)	全国学力調査	全国学力調査	全国学力調査	
昭和40年～ 1965)				能研テスト
昭和50年～ (1975)	教育課程 実施状況調査	教育課程 実施状況調査		
昭和60年～				共通第1次 学力試験
平成元年～ (1989)	教育課程 実施状況調査	教育課程 実施状況調査		大学入試 センター試験
平成10年～ (1998)	教育課程 実施状況調査	教育課程 実施状況調査	教育課程 実施状況調査	
平成20年～	全国学力・学習状況調査			

日本における全国規模の学力調査

	全国学力調査	教育課程実施状況調査		全国学力・学習状況調査
		～1995年度	2001年度～	
実施主体	・文部省	・文部省	・国立教育政策研究所 (一部の業務を民間に外部委託)	・文部科学省 ・国立教育政策研究所 ・一部の業務を民間に外部委託
規模	抽出(中学:悉皆(4回)、自主参加校)	抽出 (約16,000人/学年)	抽出 (約16,000人/冊子)	悉皆
調査対象	小5, 6 中1, 2, 3 高3, 定4	小5, 6 中1, 2, 3	小5, 6 中1, 2, 3 高3	小6 中3
調査項目	国、社、数、理、 外、図、保体、技 術・家庭、教科以外	国語、社会、算数・数学、理科、英語		国語、算数・数学
調査内容	・ペーパーテスト	・ペーパーテスト	・ペーパーテスト ・児童・生徒質問紙 ・教師質問紙	・ペーパーテスト ・児童・生徒質問紙 ・学校質問紙
調査サイクル	毎年(調査教科、 学年は様々)	学習指導要領の全面実施の2年目 (2001小中・2002高のみ、実施の最終年度)		毎年

日本における全国規模の学力調査

- 全国学力調査
 - 小1956(昭和31)～1966(昭和41): 抽出(約4～5%→約20%)
 - 中1956(昭和31)～1966(昭和41): 抽出(約4～5%)⇒**悉皆**⇒抽出(約20%)
 ※**悉皆調査1961(昭和36)～1964(昭和39)年度の4年間**
 - 高1956(昭和31)～1962(昭和37): 抽出(約10%)
- 教育課程実施状況調査: 抽出
 - 小**昭和52年告示**→昭和55年全面实施 1981-1982(昭和56,57),
平成元年告示→平成 4年全面实施 1993-1994(平成5, 6), 2001(平成13)
平成10年告示→平成14年全面实施 2003(平成15)
 - 中**昭和52年告示**→昭和56年全面实施 1982-1983(昭和57, 58)
平成元年告示→平成 5年全面实施 1994-1995(平成6,7), 2001(平成13)
平成10年告示→平成14年全面实施 2003(平成15)
 - 高**平成元年告示**→平成 6年学年進行実施 2002-2003(平成14,15)
平成10年告示→平成15学年進行実施 2005(平成17)
- 全国学力・学習状況調査: **悉皆**
 - 小・中2007(平成19)～ ※2010以降は未定

日本における全国規模の学力調査 －全国学力調査－

	小学校										中学校										高等学校									
	学年		教科								学年			教科							教科(科目数)									
	5	6	国	社	算	理	音	図	家	他	1	2	3	国	社	数	理	英	他	国	社	数	理	英	体					
1956(S31)		✓	✓		✓								✓	✓		✓				✓		✓								
1957(S32)		✓		✓		✓							✓		✓		✓				✓		4							
1958(S33)		✓					✓	✓	✓	✓		✓	✓					3	2					✓	✓					
1959(S34)		✓	✓		✓								✓	✓		✓				✓		✓								
1960(S35)		✓		✓		✓						✓	✓		✓		✓				2		1							
1961(S36)		✓	✓		✓							✓	✓	✓	✓	✓	✓	✓						✓						
1962(S37)	✓	✓	✓		✓							✓	✓	✓	✓	✓	✓	✓					✓							
1963(S38)	✓	✓		✓		✓						✓	✓	✓	✓	✓	✓	✓												
1964(S39)	✓	✓	✓		✓							✓	✓	✓	✓	✓	✓	✓												
1965(S40)	✓	✓		✓		✓							✓	✓	✓	✓	✓	✓												
1966(S41)	✓		✓		✓		✓				✓		✓	✓		✓			3											

日本における全国規模の学力調査

～調査目的・趣旨～

- **全国学力調査**
 - 学力の実態をとらえる
 - 学習指導の改善に役立てる
 - 教育条件の整備に利用する
- **教育課程実施状況調査**
 - 学習指導要領に基づく教育課程の実施状況について、学習指導要領における各教科(科目)の目標や内容に照らした学習の実現状況の把握を通して調査研究
 - 指導上の問題点は何かなどを明らかにして、今後の教育課程や指導方法等の改善に資する
- **全国学力・学習状況調査**
 - 国が、全国的な義務教育の機会均等とその水準の維持向上の観点から、各地域における児童生徒の学力や学習状況をきめ細かく把握・分析することにより、教育及び教育施策の成果と課題を検証し、その改善を図る
 - 各教育委員会、学校等が、全国的な状況との関係において自らの教育及び教育施策の成果と課題を把握し、その改善を図るとともに、そのような取組を通じて、教育に関する継続的な検証改善サイクルを確立する
 - 各学校が、各児童生徒の学力や学習状況を把握し、児童生徒への教育指導や学習状況の改善等に役立てる

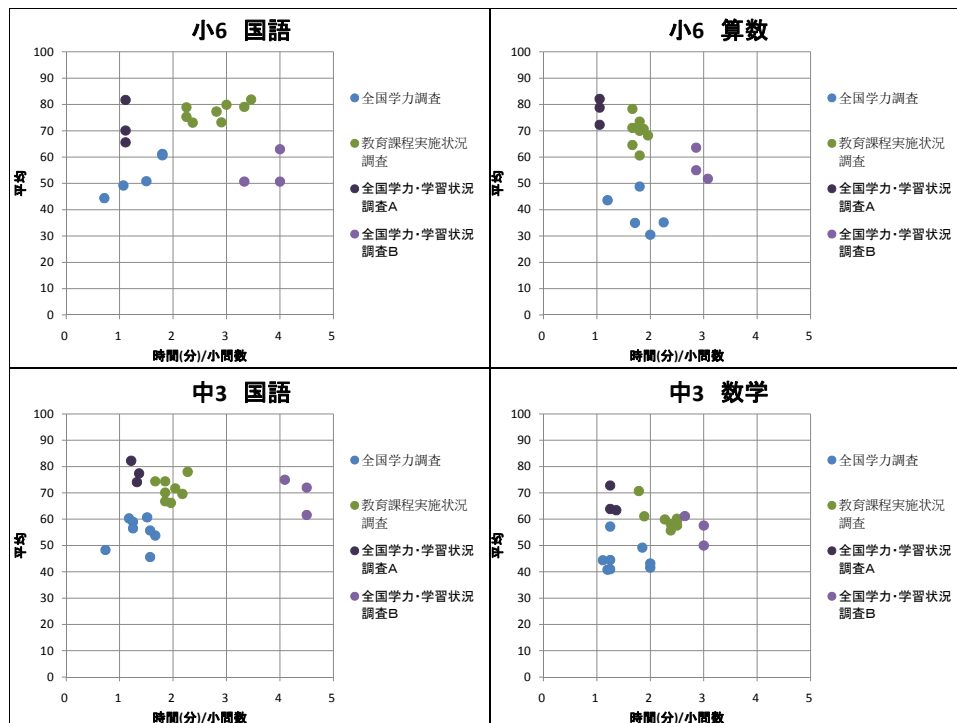
日本における全国規模の学力調査

～主な公表結果・分析(統計量の視点から)～

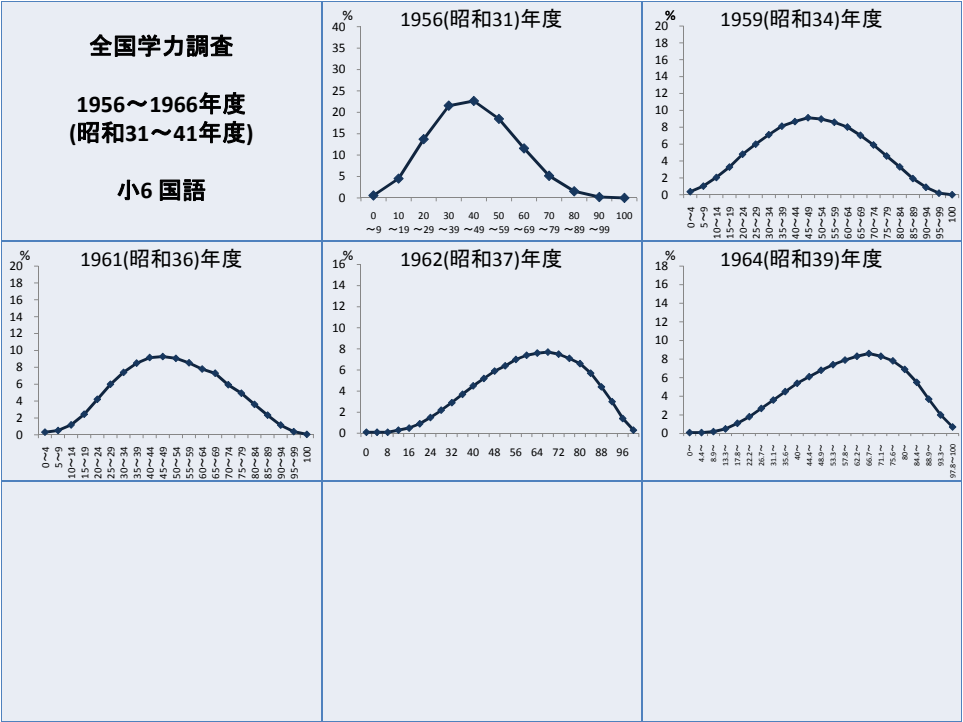
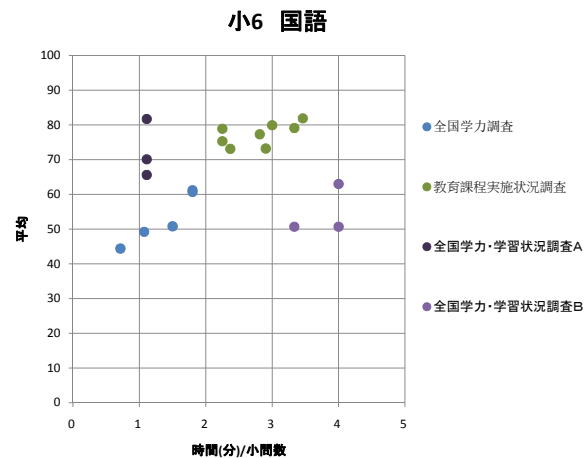
- **全国学力調査**
 - 全国平均点、問題別正答率
 - 平均点、標準偏差、変異係数
 - 個人得点分布、学校平均点分布
 - 地域類型別の個人得点分布、学校平均点分布
 - 鉱業市街、工業市街、商業市街、住宅市街、その他の市街、鉱山、漁村、山村、小都市、都市近郊農村、農山村、純農村、普通農村、その他、へき地
 - 標準化換算表
- **教育課程実施状況調査**
 - 各問題については配点を行わず、問ごとの「通過率」を求めて結果の分析
 - ① 個々の問題の通過率および解答類型ごとの反応率分析 ② 教科、学年別にみた問題ごとの分析 ③ (学習指導要領に示す) 内容、領域別にみた問題ごとの分析 ④ (指導要録に示す) 評価の観点別にみた問題ごとの分析 ⑤ (過去の調査と) 同一問題の通過率比較分析
 - ⑥ 設定通過率との比較分析 ⑦ 記述式の問題における問題ごとの分析
- **全国学力・学習状況調査**
 - 平均正答数
 - 全国、層別(公立、国立、私立)、都道府県別
 - 正答数分布
 - 全国、層別(公立、国立、私立)、都道府県別

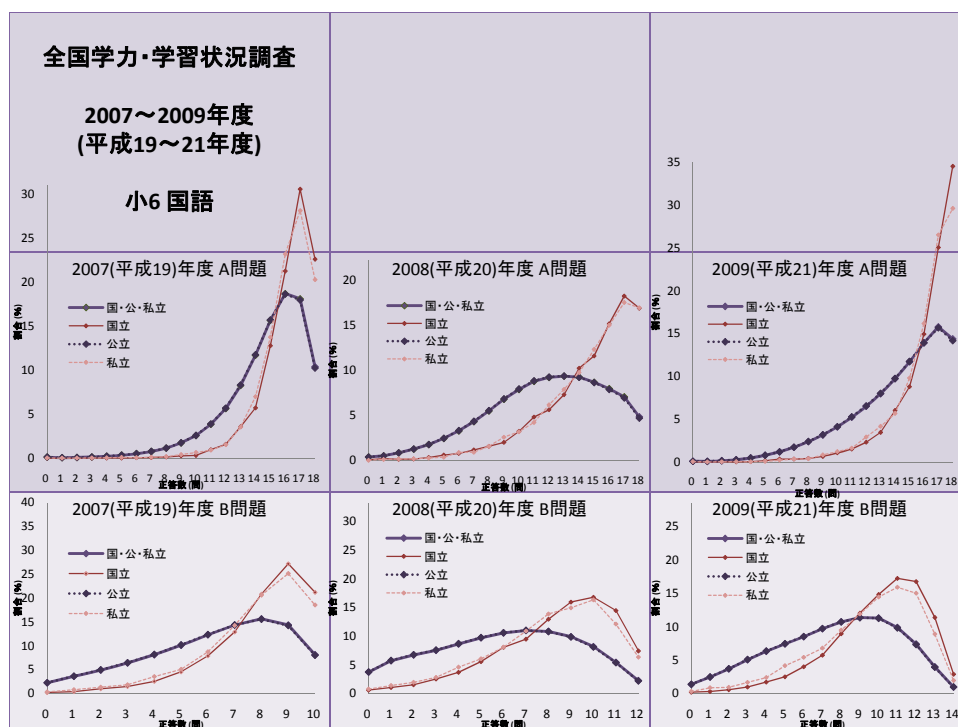
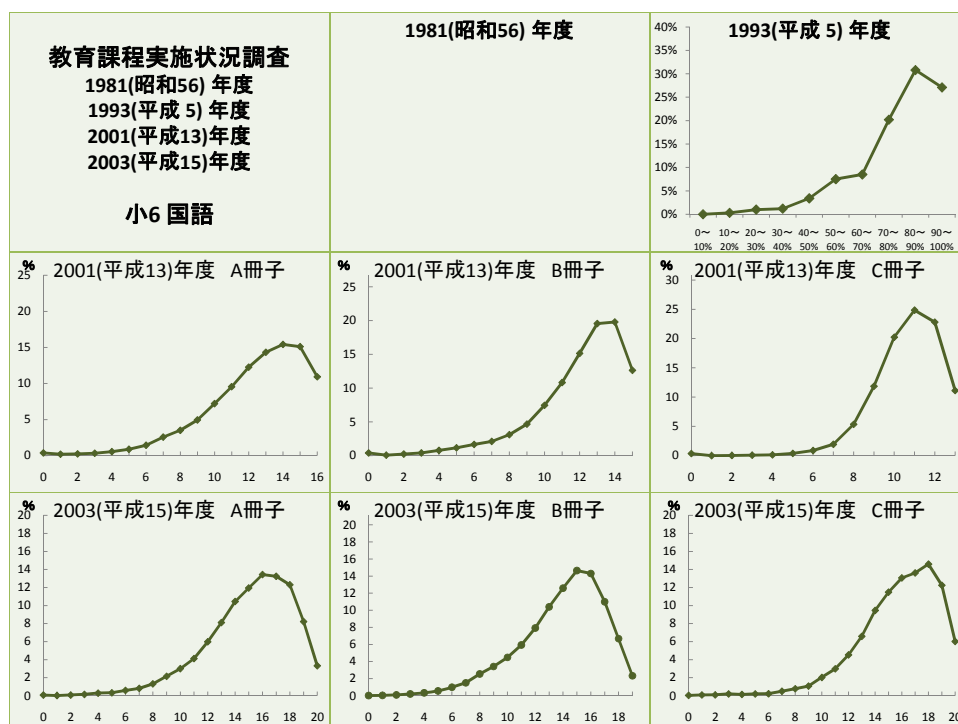
日本における全国規模の学力調査 ～統計結果の特徴を比較～

- 3つの調査(小6・中3 国語、算数・数学のみ)
 - 全国学力調査
 - 教育課程実施状況調査
 - 全国学力・学習状況調査
- 1問あたりの平均時間(分)と平均との関係
- 得点分布の特徴

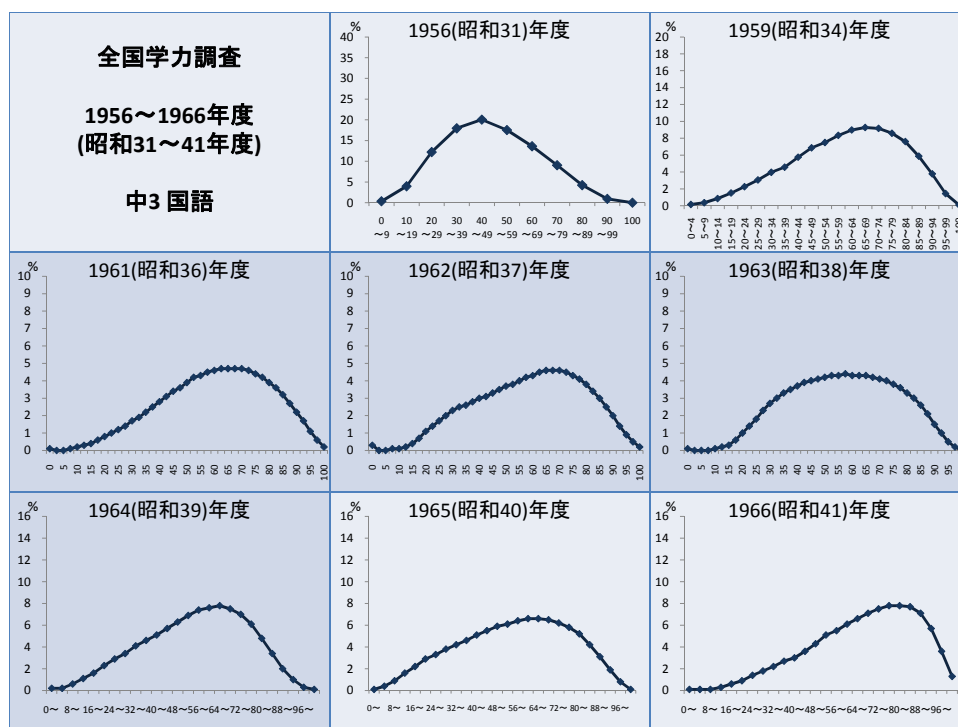
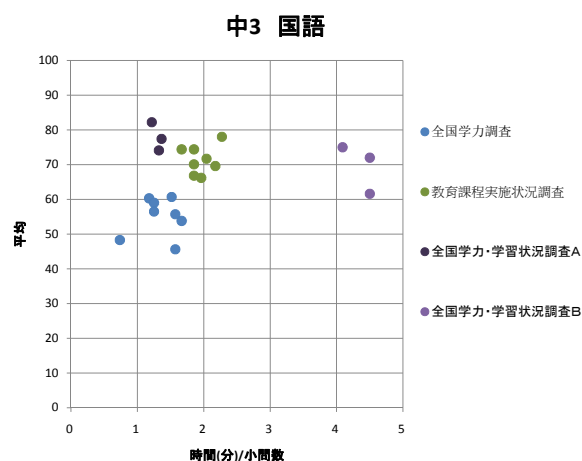


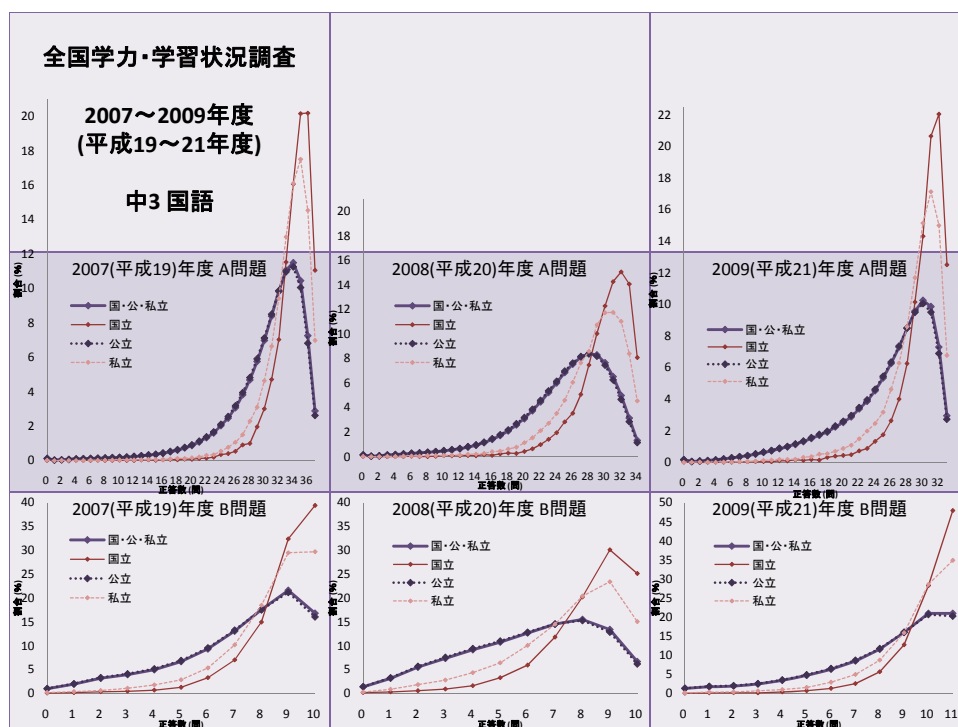
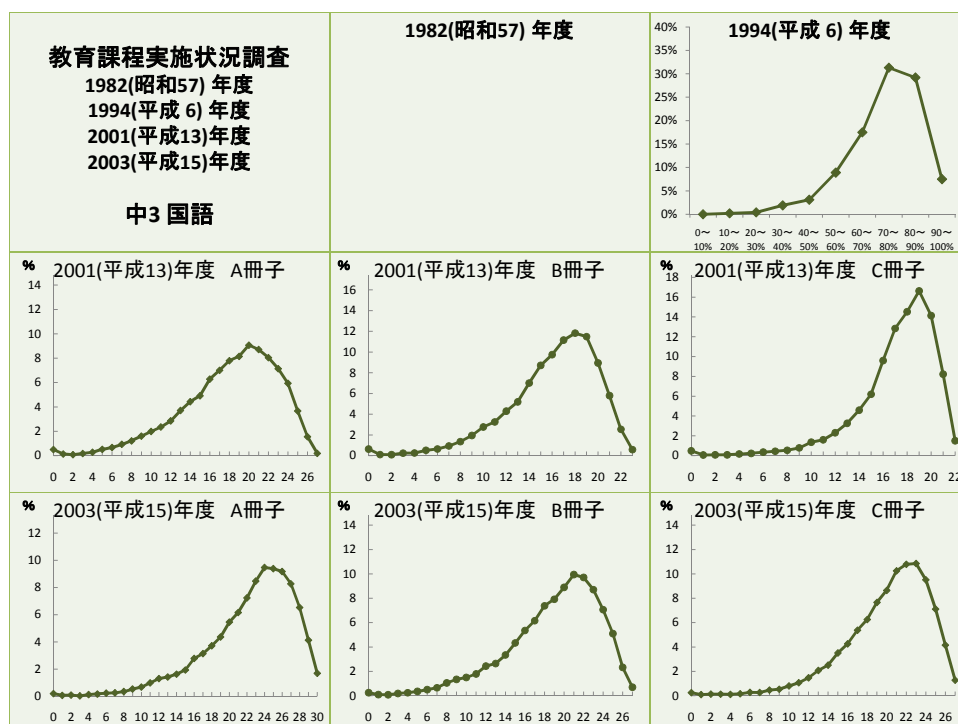
小学校 第6学年 国語



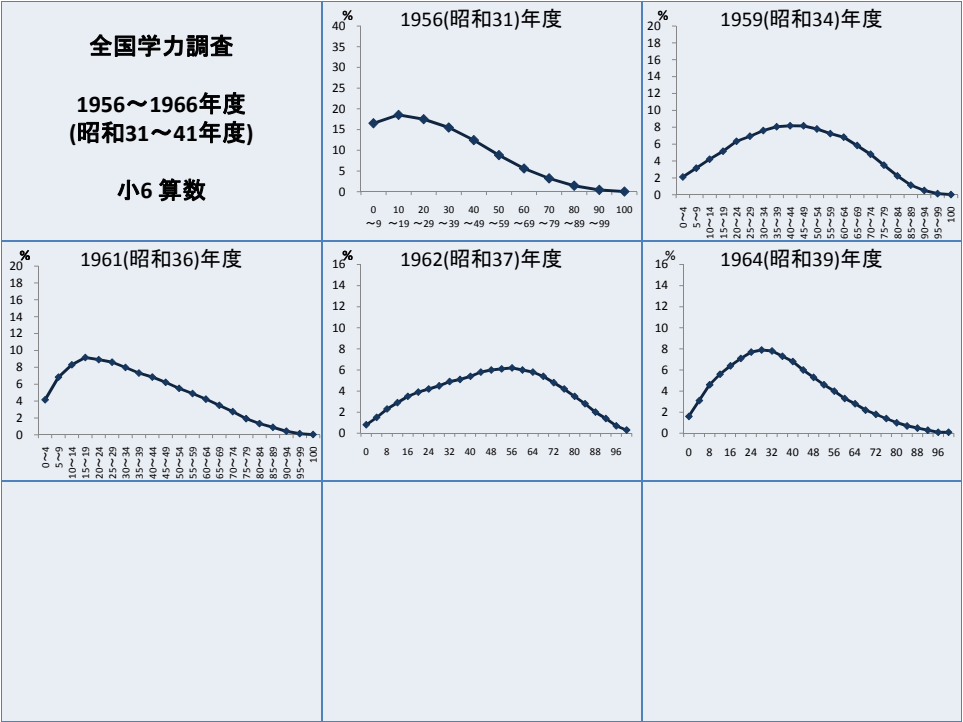
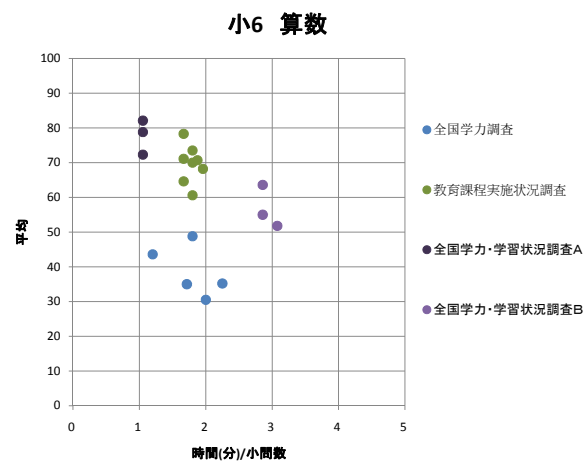


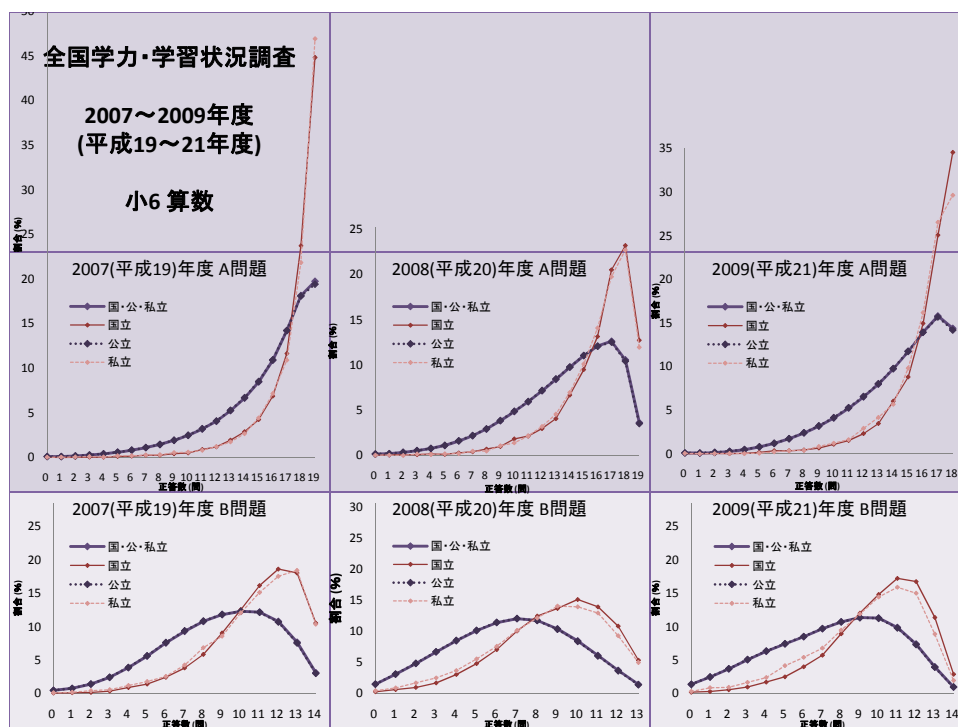
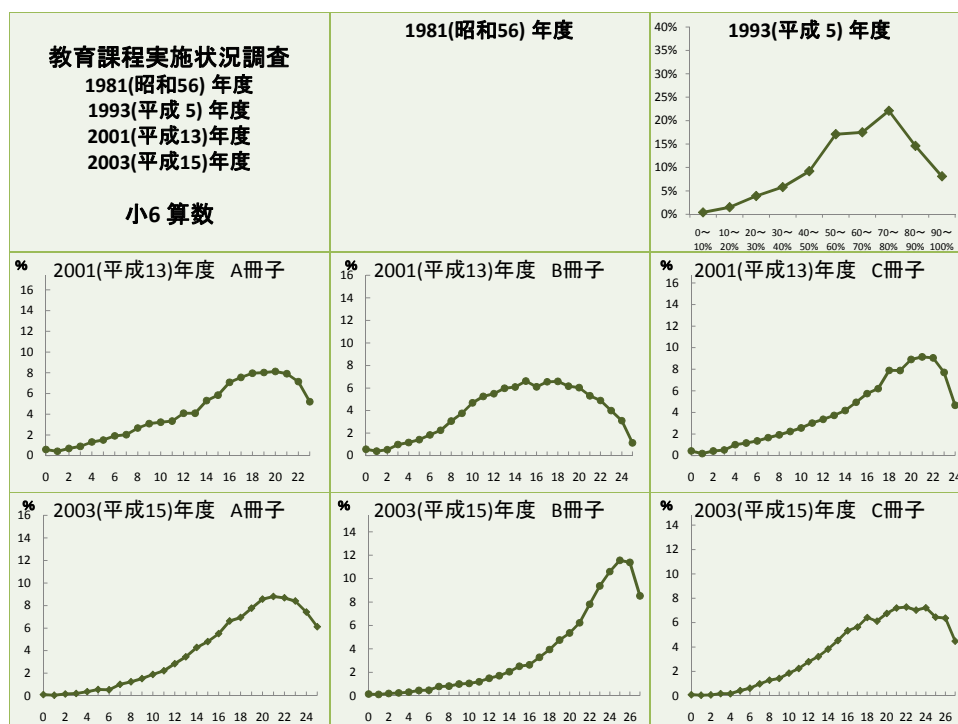
中学校 第3学年 国語



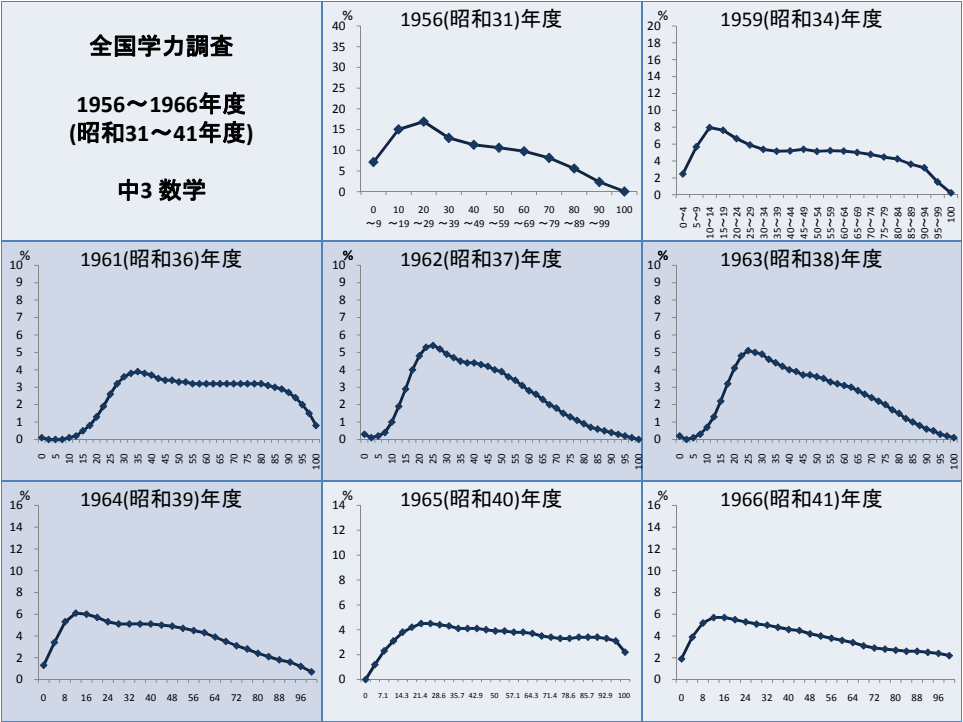
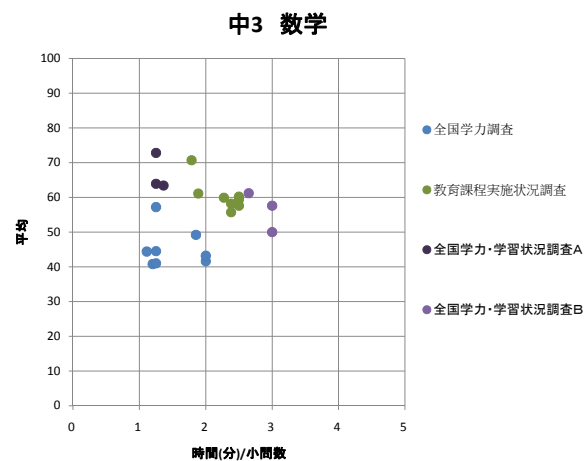


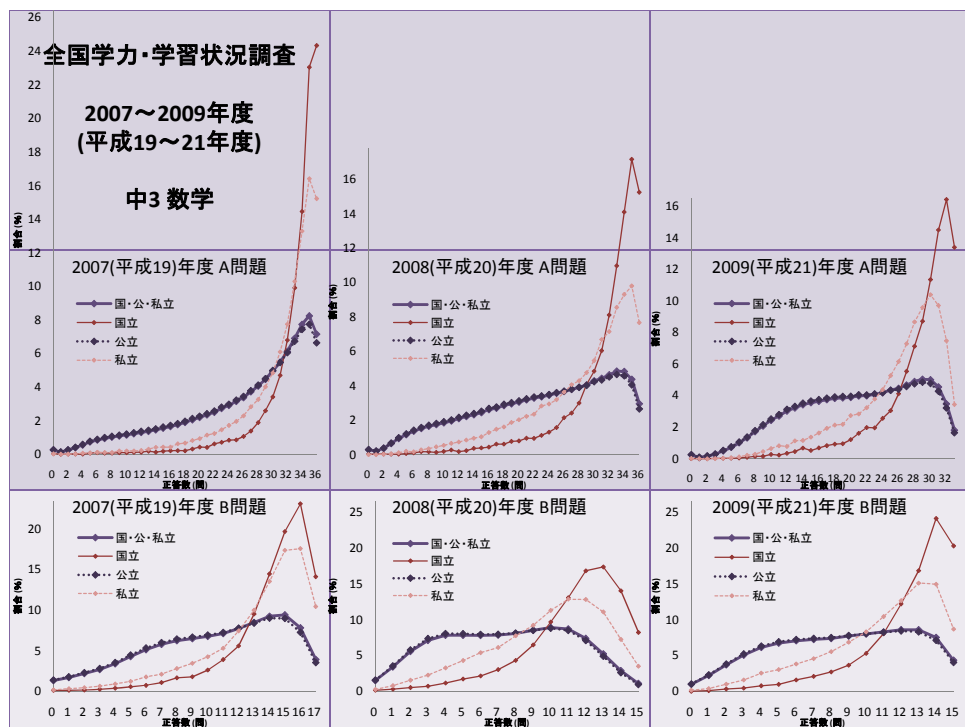
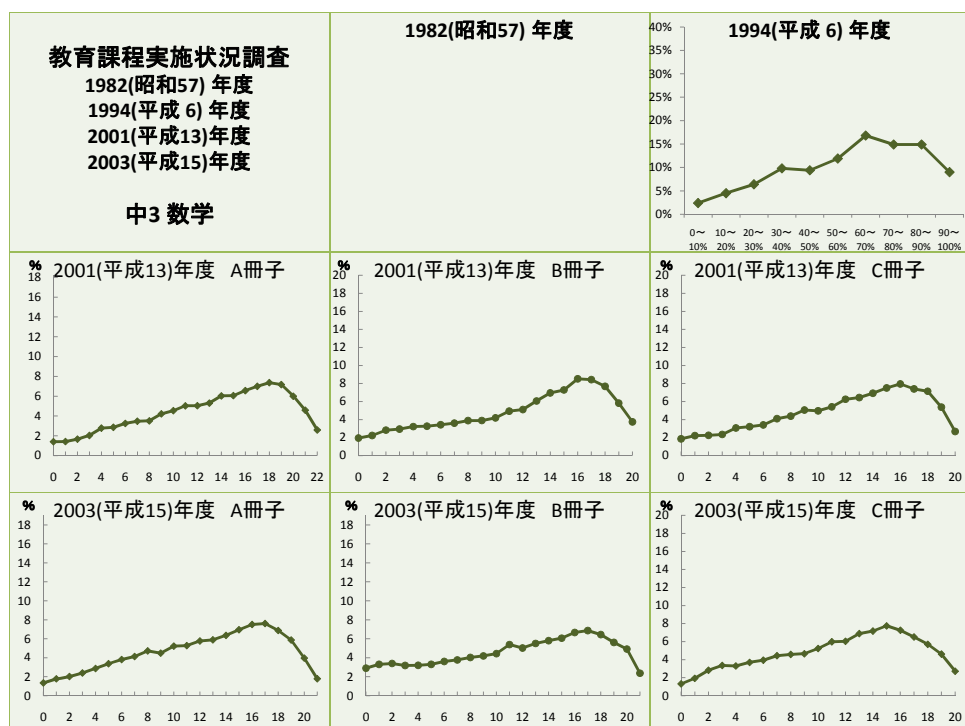
小学校 第6学年 算数





中学校 第3学年 数学





海外での事例: 全米学力調査(NAEP)

The National Assessment of Educational Progress

- 抽出調査
 - 学校数15,000 校以上、生徒数85 万人以上が調査に参加
- 1969 年から開始され、毎年何らかの調査が行われている
- 連邦議会で決定され、教育省が責任を持つが、調査方針は全米調査統括委員会(NAGB:National Assessment Governing Board) がたてる
- サンプルング、作問、調査実施、採点、分析といった実務は5 年ごとの入札制度によって、応札機関・業者を決定
 - 現在はETS(Educational Testing Service) という非営利団体が中心となりその他6 団体からなるチームで実施
- 結果公表は、ハイライト版の小冊子(NAEP Report Card) と Web での公開(生徒や学校に個別に結果を返すことはない)
- 地域別統計は、米国北東部、南東部、中央部、西部の大枠のみ
- 参考URL <http://nces.ed.gov/nationsreportcard/>

海外での事例: 全米学力調査(NAEP)

	Main NAEP	long-term trend NAEP	State NAEP
調査対象	第4,8,12 学年	9,13,17 歳	第4,8 学年
調査教科	数学、読解、作文、科学、合衆国史、公民、地理、芸術、(世界史、外国語) など	数学、読解、作文、科学、(公民)	数学、読解、作文、科学
調査内容	時代や社会の変化に応じて教育課題の優先順位も変わるため、それに焦点を当てた内容を測定	・時代や社会変化にかかわらず不変な基礎学力の測定 ・経年変化を測定	州単位の統計量の算出および全米の結果との比較のため
調査サイクル	毎年(学年、教科は異なる)		2 年ごとに実施

海外での事例：全米学力調査(NAEP)

	Main NAEP												Trend NAEP				State NAEP			
	数学	読解	作文	科学	国史	公民	地理	芸術	世史	公民	経済	外語	数学	読解	作文	科学	数学	読解	作文	科学
1999													✓	✓	✓	✓				
2000	✓	4		✓													4,8			4,8
2001					✓		✓													
2002		✓	✓														4,8	4,8		
2003	4,8	4,8															4,8	4,8		
2004												12	✓	✓						
2005	✓	✓		✓													4,8	4,8		4,8
2006					✓	✓					12									
2007	4,8	4,8	8,12														4,8	4,8	8	
2008								8					✓	✓						
2009	✓	✓		✓													4,8	4,8	4,8	
2010							✓		12											
2011	4,8	4,8	✓														4,8	4,8	4,8	
2012										✓		✓	✓	✓						

おわりに-今後の展望

- I. 調査設計
- II. 项目开发 - 学習指導要領との関係
- III. 項目評価
- IV. 冊子構成
- V. 結果公表
- VI. 評価

おわりに-今後の展望

I. 調査設計

- ◆ 調査の目的 → 中長期的な視点で
- ◆ 項目の枠組み =学習指導要領?
- ◆ 調査対象(規模,学年(年齢),教科(科目) など)、調査サイクル

II. 項目開発 — 学習指導要領との関係→ 予備調査等の利用

- ◆ 項目属性
 - ✓ 内容(領域、範囲、配分)、形式(客観式、記述式など)、測定的特徴(難易度、識別力など)
- ◆ 採点基準と手順
- ◆ 項目数

III. 項目評価

- ◆ 項目の評価基準 — 良質の項目を選択

IV. 冊子構成

- ◆ 項目の配列(冊子数)、レイアウト、再製、機密性
- ◆ 調査後の問題の扱い(公開、非公開)

おわりに-今後の展望

V. 結果公表 → 理想は、実施前に公表データ、表を確定

- ◆ 公表の単位
 - ✓ 個人・学級・学校・自治体・国、人口統計上の下位集団別、学校設置者別、性別、その他
- ◆ 公表のすべきか否か?
 - ✓ 社会的、政治的、教育的影響、誤解、データからの情報の価値、データの誤用の可能性
- ◆ 公表データの種類
 - ✓ 正答率、標準化得点、パーセンタイル(四分位、五分位) など
- ◆ 比較データの提供 (年度間比較、学年間比較など)
 - ✓ - 変化があったのか、なかったのか?(有意性)
- ◆ 公表の様式
 - ✓ 媒体、方法、受けてのニーズに合った様式

VI. 調査の評価

- ◆ 成功の定義は?
 - ✓ 目的どおりに教育課程や教育プログラムが遂行されているか? 効果は何か?
⇒ 調査設計やデータの解釈が重要な点で変わってくる
- ◆ データの操作
 - ✓ 調査内容と指導内容の合致度、テスト経験、不正行為
⇒ 悪用の可能性 (小:指導改善のためのフィードバック 大:報酬や賞罰に直結)
- ◆ テストの信頼性・妥当性