

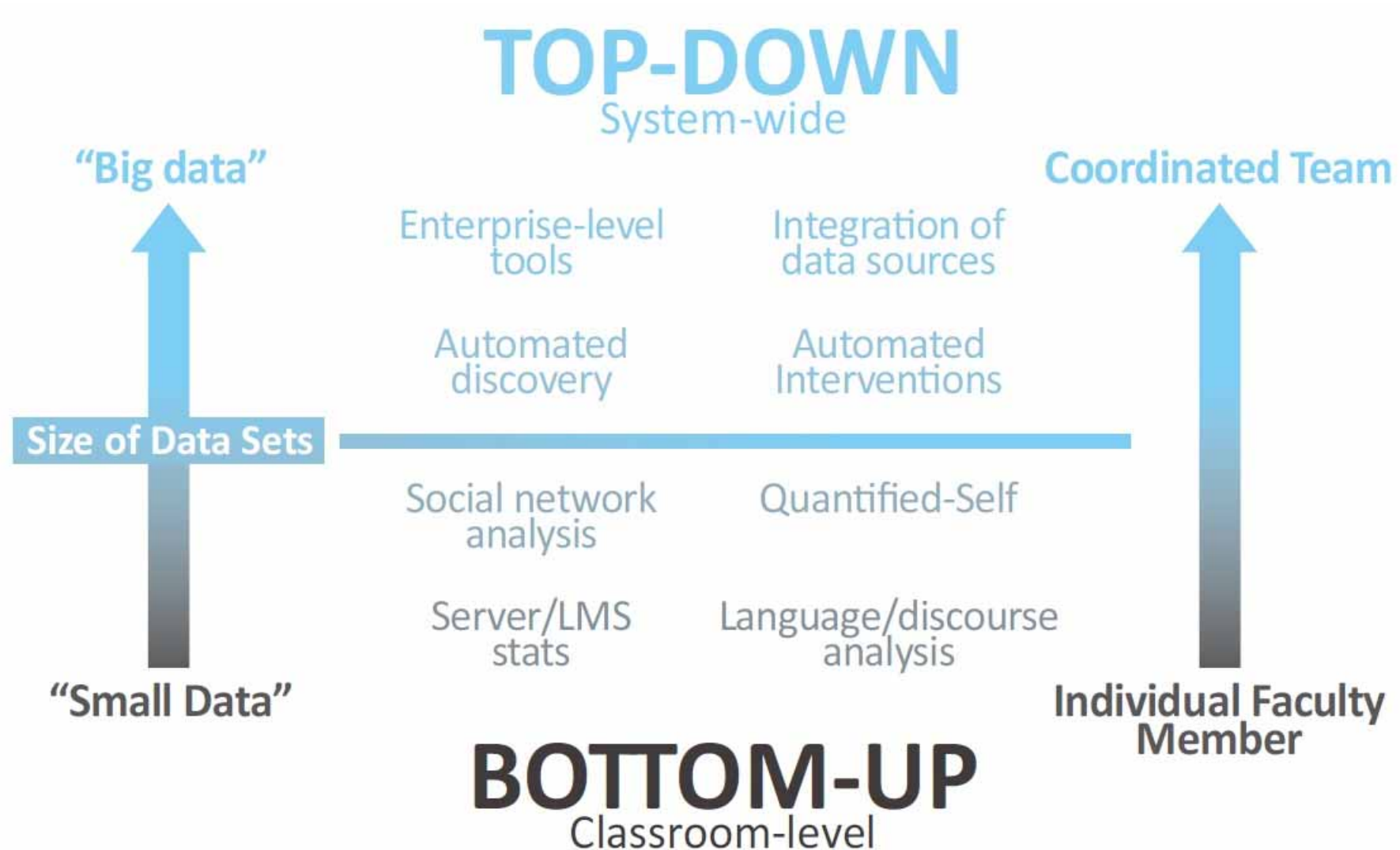
Learning Analyticsと教学IRなどの
教育ビッグデータの現状と限界
および企業の貢献の可能性
～フレックスの事例をもとに～

福井県立大学
学術教養センター
山川 修

本日は言いたいこと

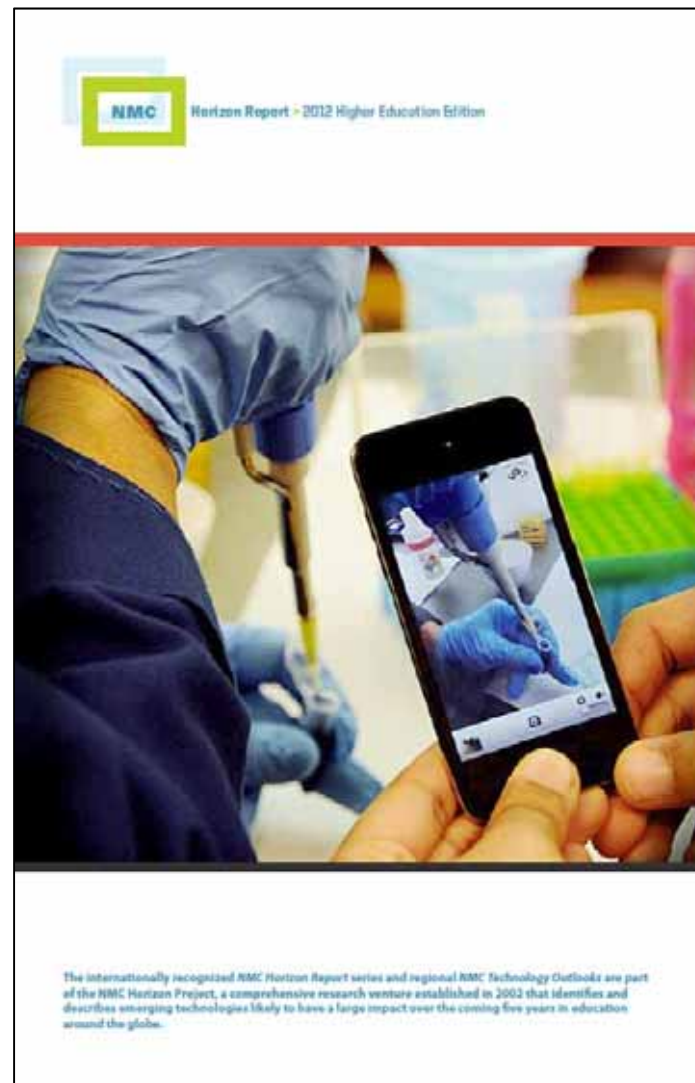
- 教育データはビッグデータとしては成立しにくい
 - 従来のビッグデータ分析とは違う方法論が必要かも
- しかし、データを基礎に教育を改善していく取り組みは意味がある
 - 学生がいかに学んでいるかに関して、いままでにない視点から可視化できる
- 多くのデータを統合し有効利用するためにはモデルが必要ではないか
 - コンテキストをモデルで乗り越える
- 企業がこの分野に貢献できることは多数ある
 - 可視化ツールとウェアラブルセンサー

LA-levels



Horizon Report 2012

- Educause Learning Initiative (ELI)
- 毎年、高等教育における1年から5年後の技術を予測
- **Learning Analytics**が昨年版から登場（4～5年後）
- 今年は2～3年後の技術として掲載



Learning Analytics (LA)

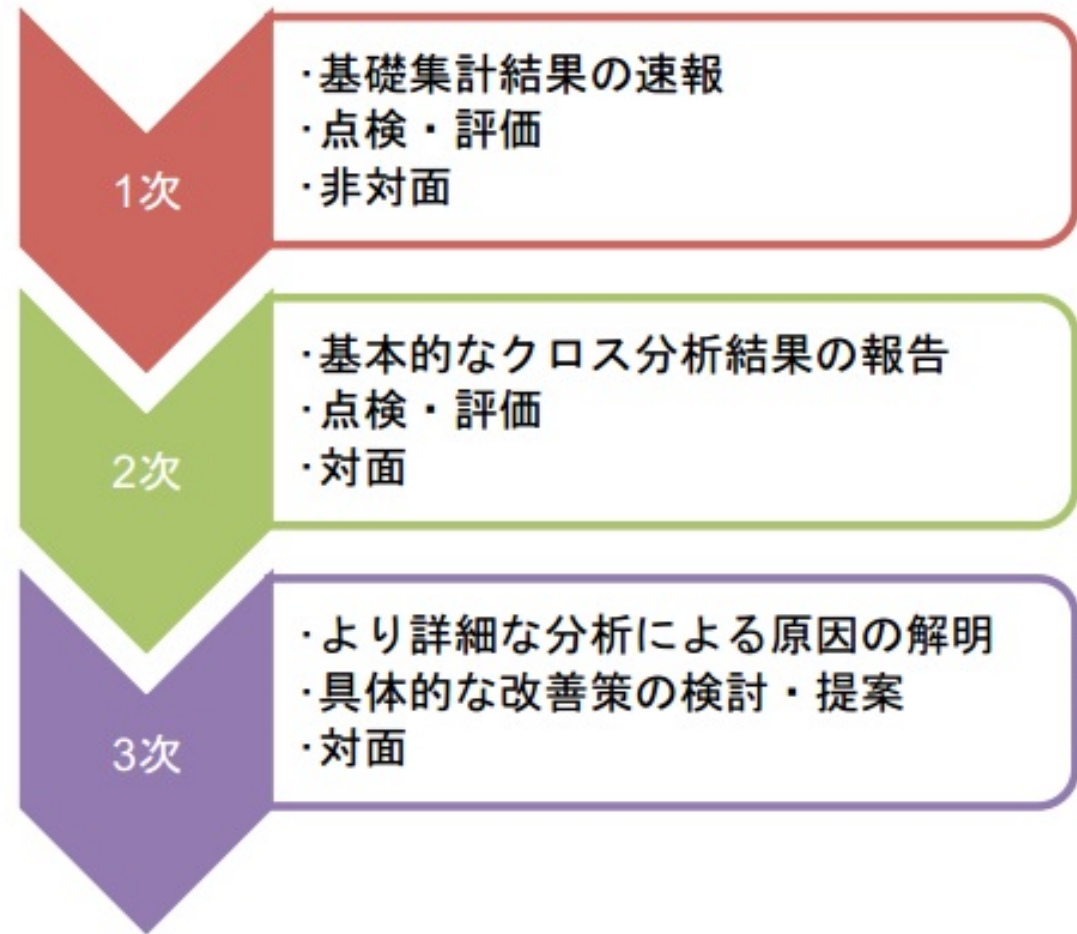
- 学生によって生み出された様々なデータの意味づけを行う
 - 学生の達成度を評価
 - 将来的な能力を予測
 - 隠された問題を見つける
- ゴール
 - 様々なレベルやニーズの学生に合わせた教育を（リアルタイムに近い形で）可能にする

教学IRに関して

- **IRの定義**：機関の計画策定，政策形成，そして意思決定を支援する情報を提供するために，高等教育機関で行われる調査研究（Saupe，1990）
- **教学IR**とは：特色を持った教育活動を行っていくため，定期的にデータ収集，分析を行い，教育の改善につなげていく活動
- 日本の場合、教学IRとLAは取り組んでいるコミュニティが違うので**あまり交流がない**
- 目的は違うが、分析結果を相互利用することによりもっと精度の高い分析ができる可能性があるので、**情報交換をもっと盛んにすべき**

教学IRのステップ

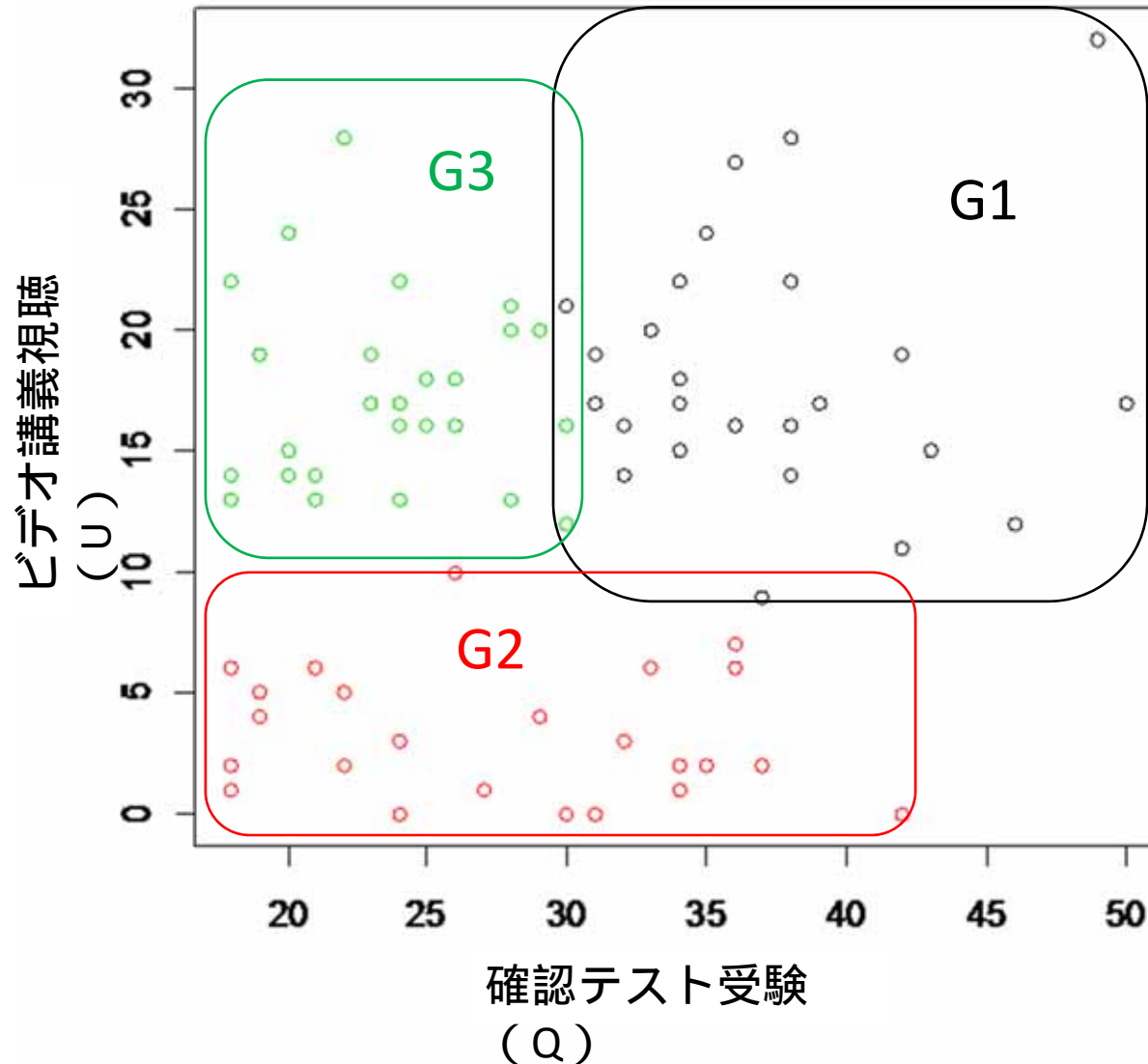
- 1次のステップは大学共同で実施することが可能で、意味がある
- 他の大学と比較することができる
- 2次以降のステップは共同実施が難しい



LMSの学習行動の分析

(対面の学習行動との対応)

学習行動によるグループ化

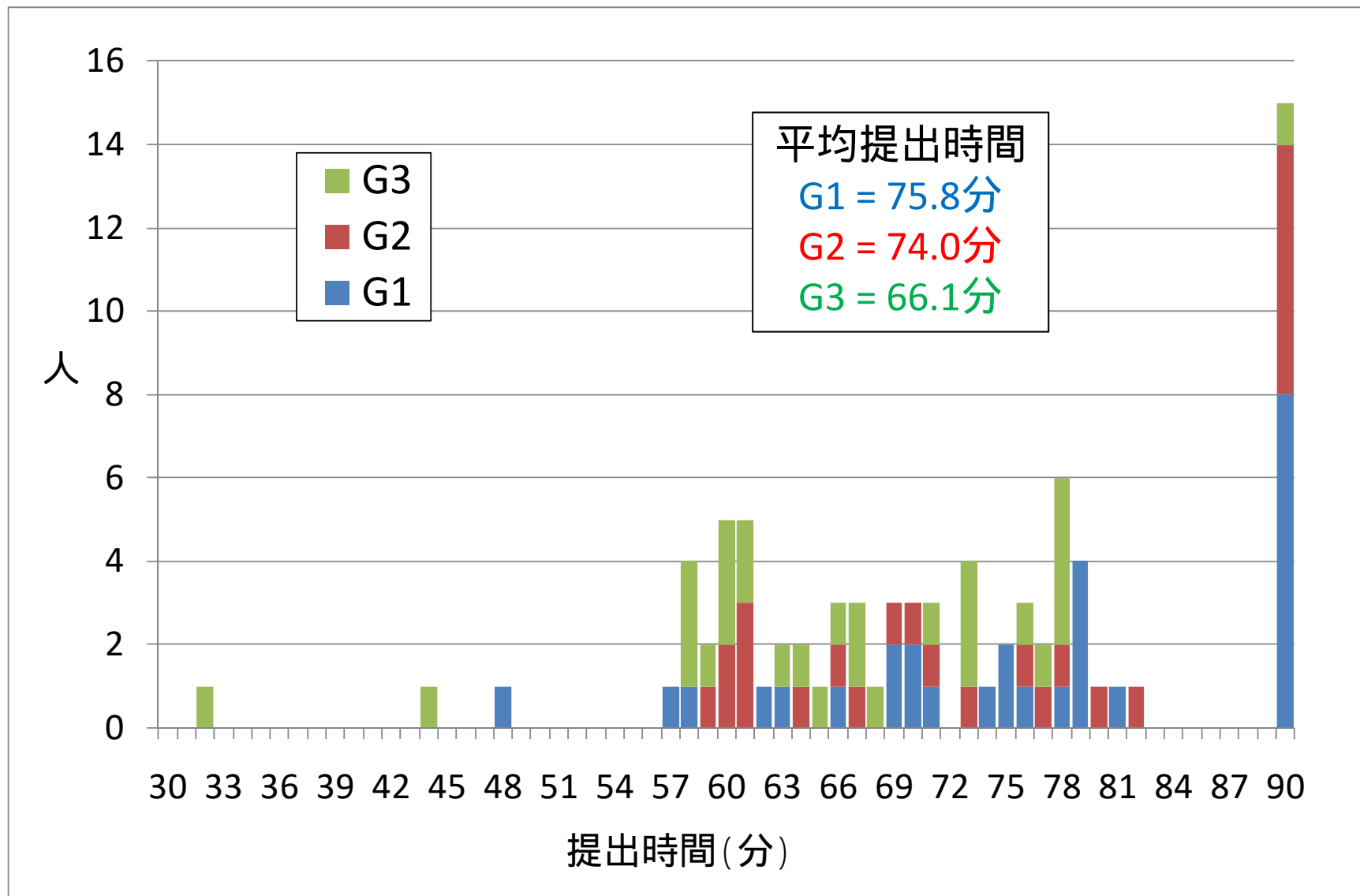


G1:
ビデオ講義視聴も確認テスト受験も多い

G2:
ビデオ講義を視聴せずに確認テストを受験

G3:
ビデオ講義視聴は多いが、確認テスト受験は少ない

期末試験の提出時間との対応










2012年度後期

大福帳

コミュニケーションカード

講師 山川 修	科目名 情報科学	アドレス
学部 工学部	学籍番号	ふりがな
		名前

月/日	言いたいこと、聞きたいこと、あなたからの伝言板	あなたへの伝言板
1 10/3	教科書もこの先生が書いていたのだからとメモ びっぴりした。 どうもありがとう。	みました 
2 10/17	おもしろいことまでして、7/11の乱用はと今まで 自分が知らなかったことをたくさん教えてもらった。 今後はこれからのことに気をつけてながらネットを 利用していきたい。	そのちょうし! 
3 10/17	電子メールが公開鍵暗号で二人の間でしかやり取りが できないようにして、暗号暗号は二人の間で見られる程 系統のようになっている。電子署名は送信者だということが 証明できるような名前ではないもの。 その理解はちがひです。	GOOD! 
4 10/31	授業講義やメールなどを利用してした場合 署名が重要で相談することが大事。特にネットでの問題が 生じた場合は自分一人で解決するのはなかなか難しい 相談など積極的に利用するのが最善の方法である。	OK 
5 10/31	なにかしらとされたらどうに個人情報の保護を しっかりと、不審なメールは容易に削除して確認を とる。送付ファイルも下取り。 （授業はとてはとて）（おまかせ） LANから出る、（おまかせ）	もうすこし! がんばって! 
6 11/13	人は周囲の環境に影響をうける人ばかりで、た まに映画や本などに影響をうけて、福祉に対する 考えがなれる。 だから、これからは映画や本、小説の分野でも、 価値観をよりよくしていきたいと思った。	OK 
7 11/13	主張：おまかせの新設が不認可のため、今回の資料を 取り上げ直した。 根拠：おまかせの数が少く、就職難になっている。 具体的には理由がある。→おまかせが不足 数の問題を解決はつづけてほしい。→おまかせが不足	よかったです! 

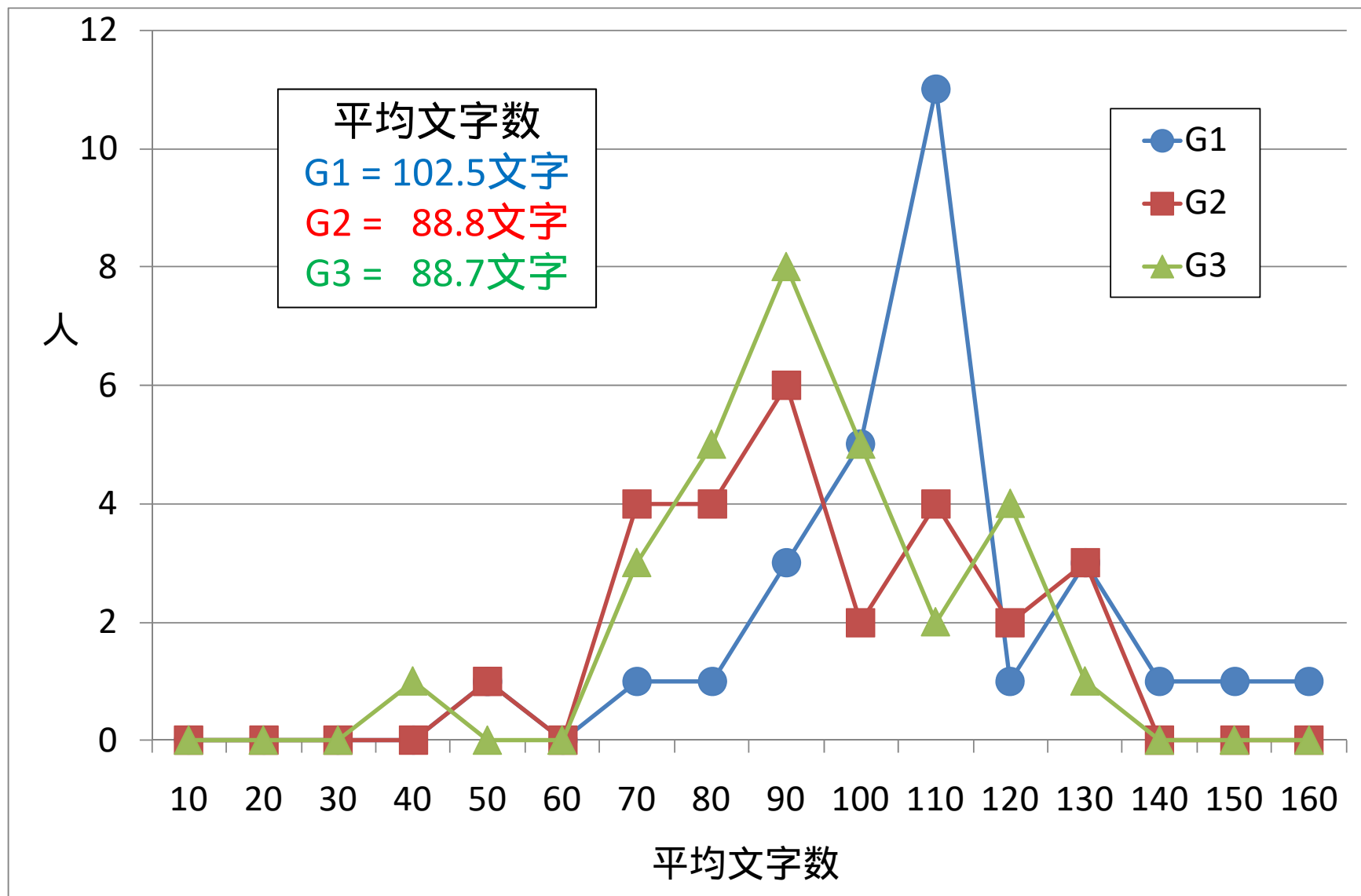
3

12

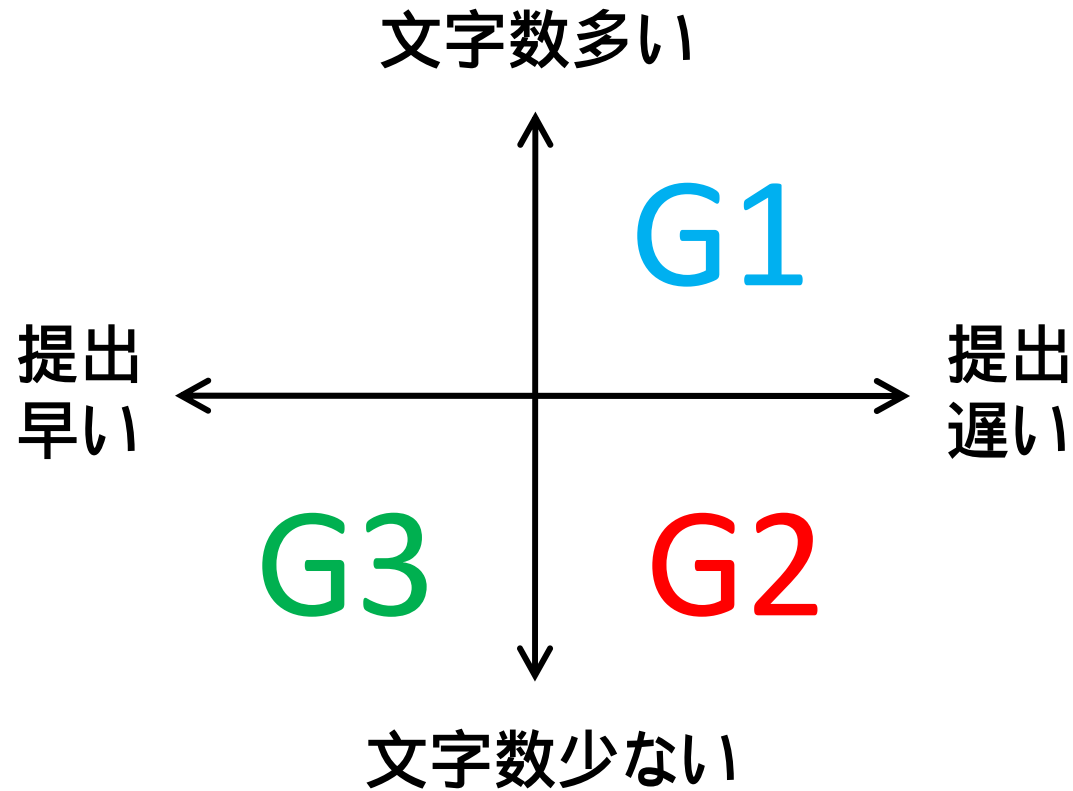
1/2

11

大福帳の平均文字数



3グループの対面における特徴



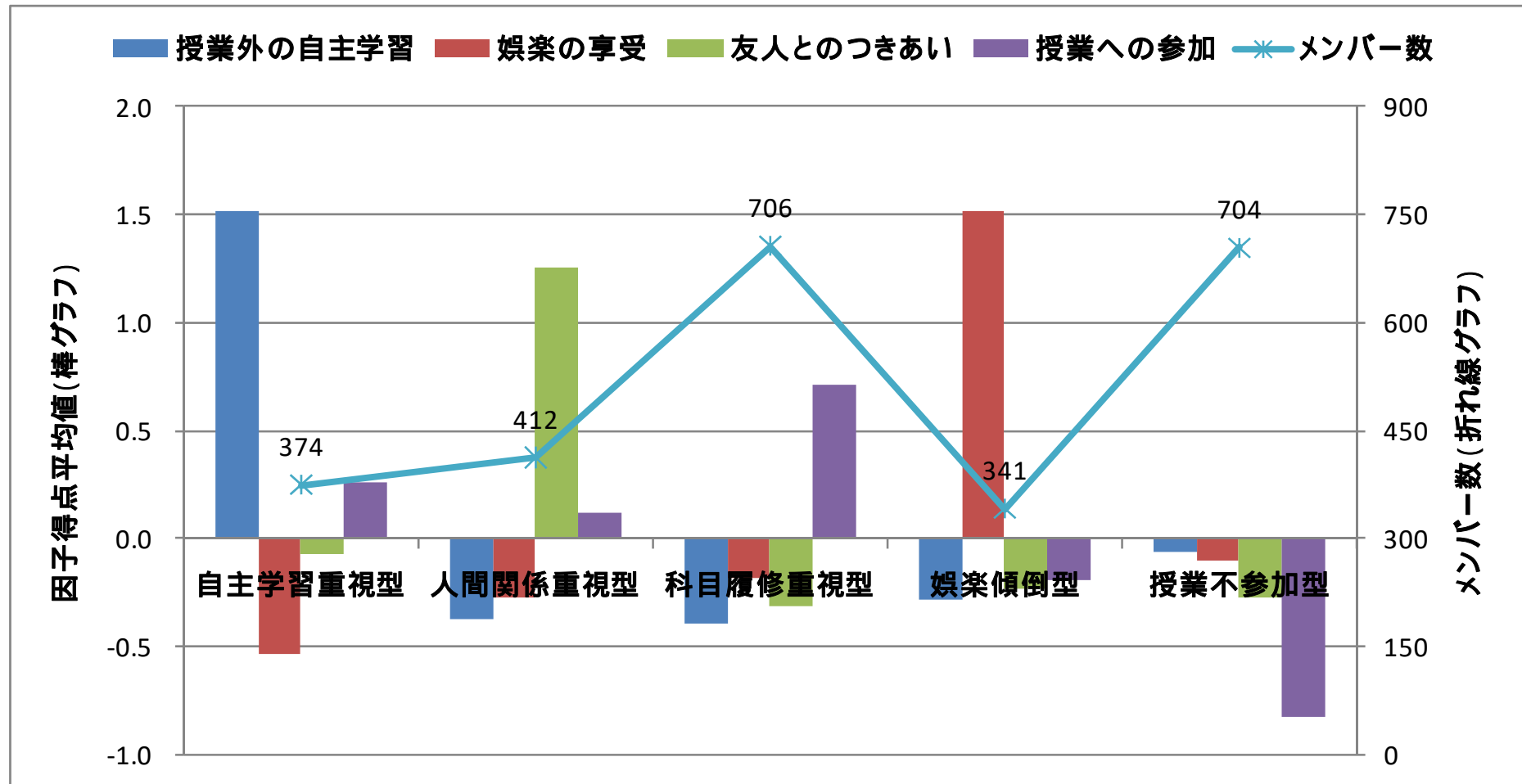
アンケートデータの分析 (教学IRとの連携)

福井県内の大学連携プロジェクト フレックス

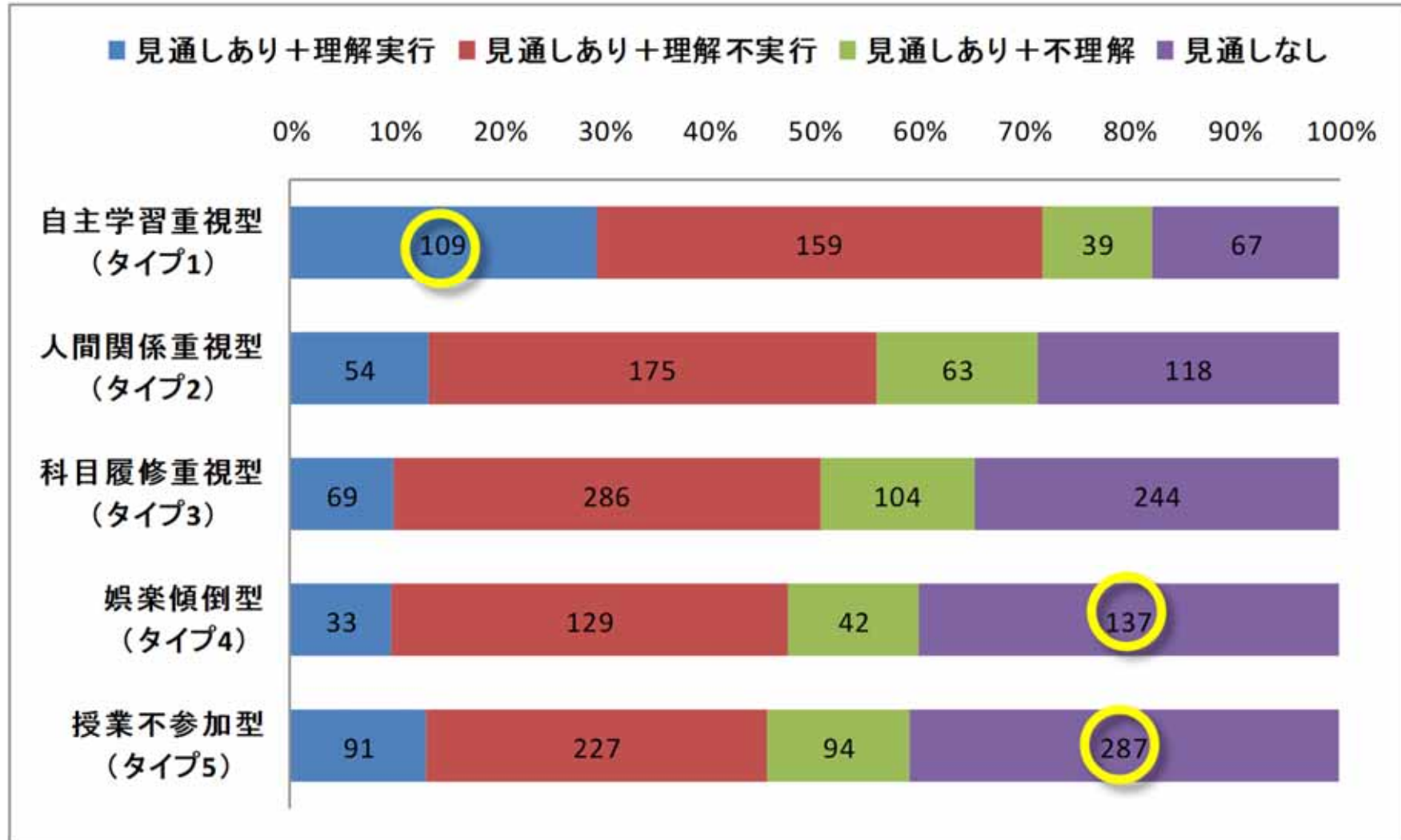
- 「学習コミュニティ」をキーワードに平成20年度から連携を開始
- **Fukui LEarning Community Consortium(F-LECCS)**
- 福井県は恐竜が有名,
T-Rexからのもじり
- Tレックス→フレックス



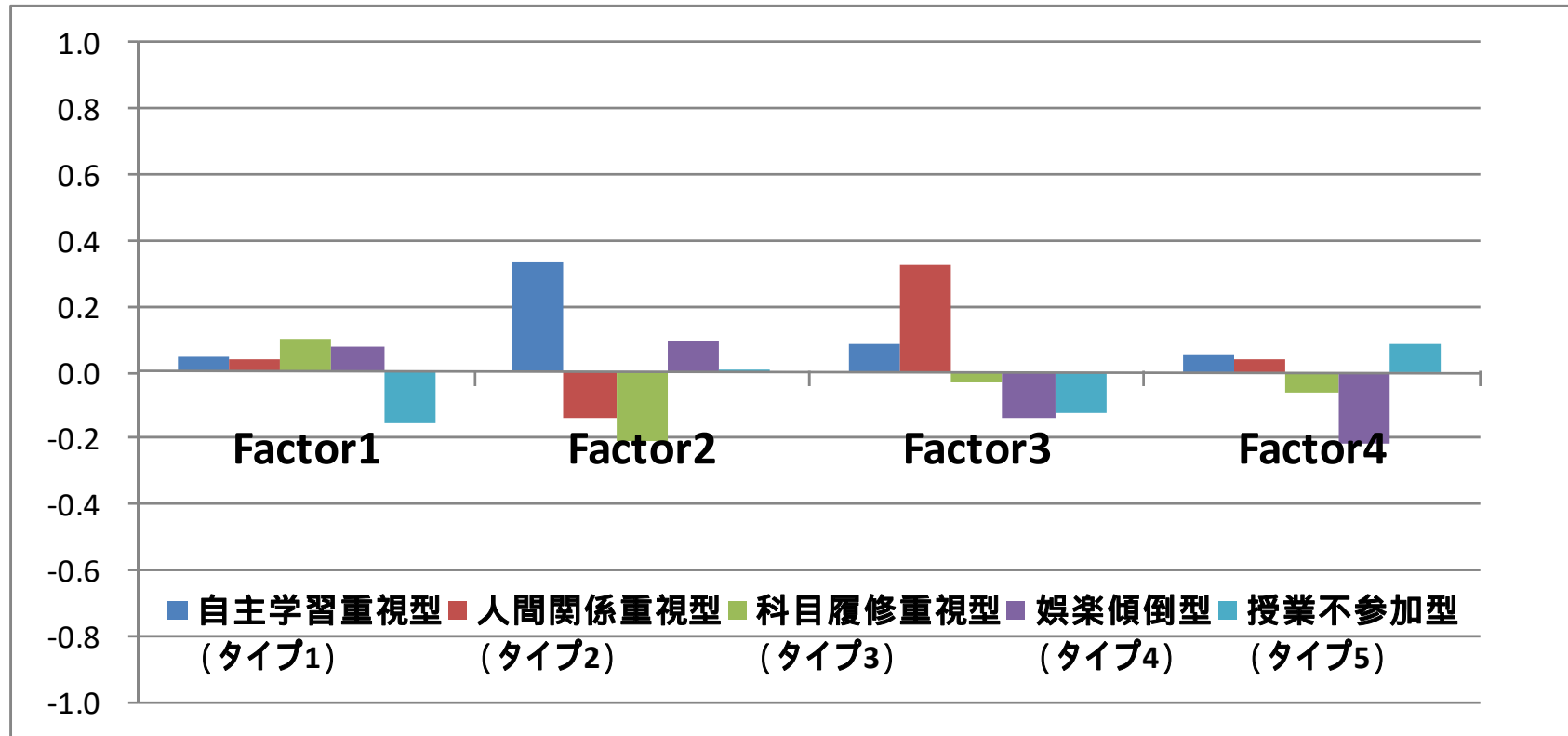
全体の分析：学生タイプ n=2526



全体の分析：将来展望



全体の分析：獲得能力



Factor1: コミュニケーション能力（授業内），知識の活用能力（授業）

Factor2: 起業家精神（授業内外），知識の活用能力、数学的思考、外国語（授業外）

Factor3: コミュニケーション能力，競い合う姿勢（授業外）

Factor4: 競い合う姿勢（授業内）

年度間比較・経年変化の調査

記名式で実施している一部の機関（機関A, B）において、
2年間または4年間共通して回答した学生をもとに調査

調査年度		2012	2013	2014	2015
回答件数 (全体)		2,739件	2,796件	3,131件	3,992件
共通 回答 件数	機関A (4年間)	101件			
	機関B (2年間)	51件		68件	
		70件			

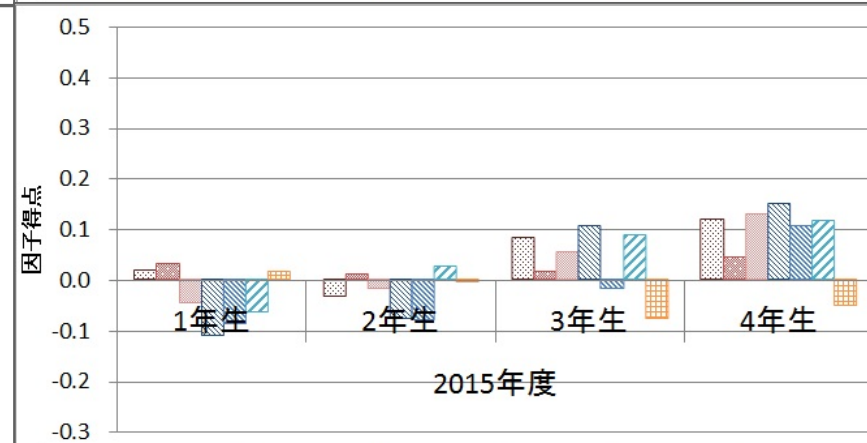
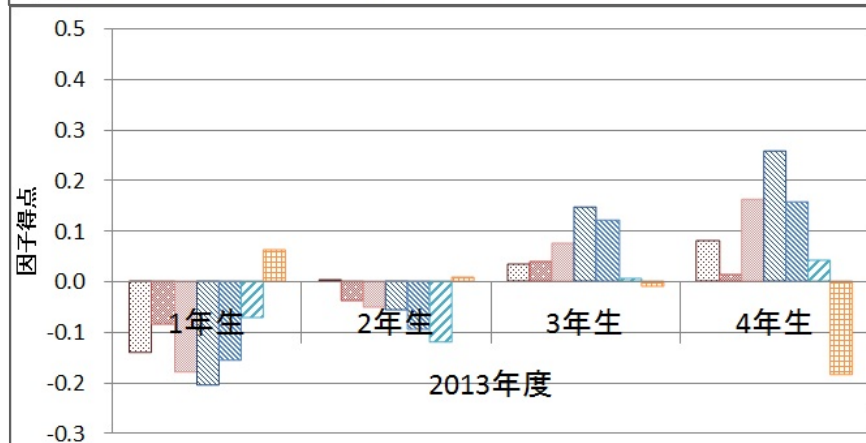
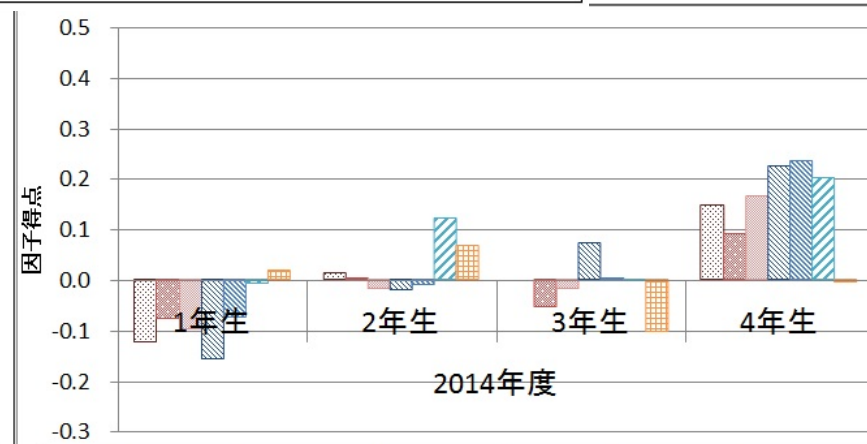
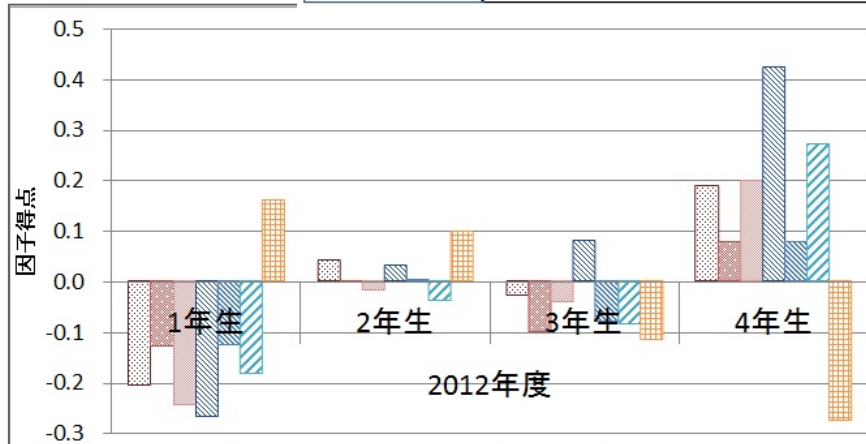
年度間比較（知識・技能の獲得）

- 知識・技能の獲得（授業内外各23項目，計46項目）の回答データを因子分析（最尤法）
- 因子分析には無記名回答も含む4年分の回答12,658件（うち，回答不備を除く）を利用
- 解釈可能な7因子を抽出

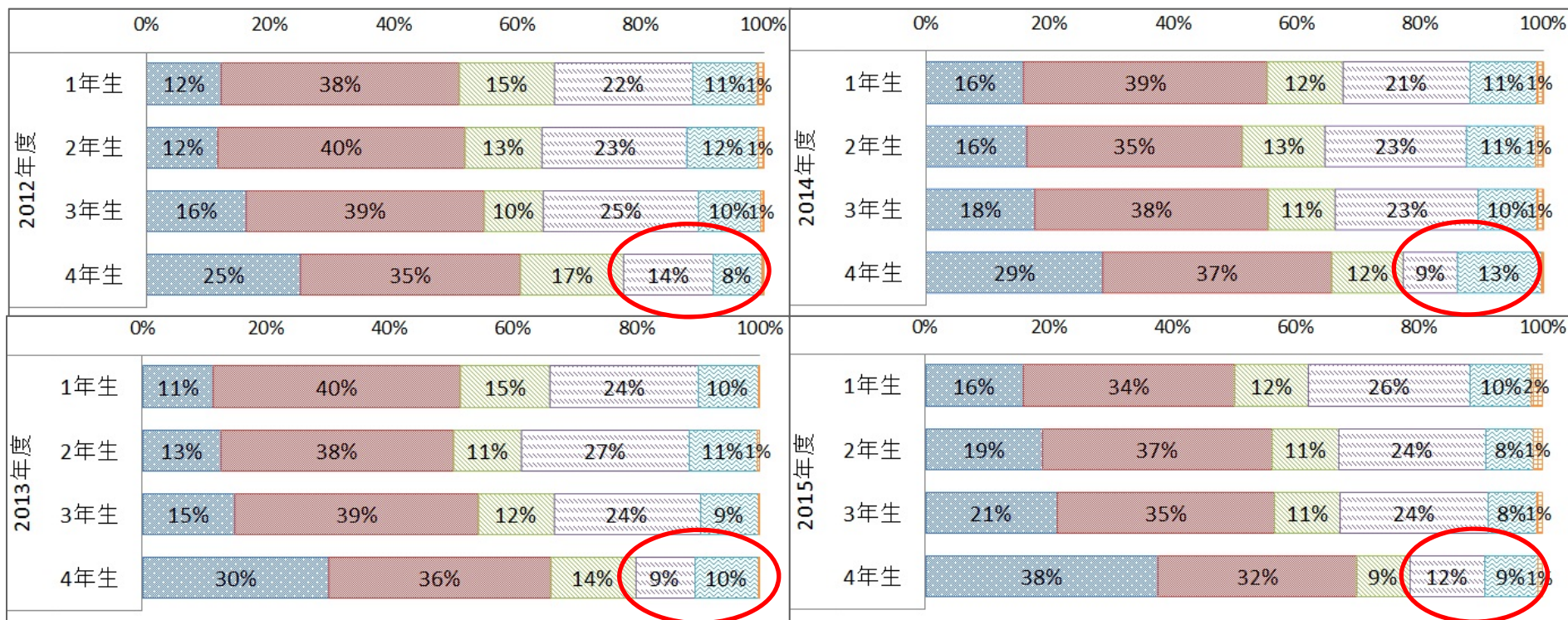
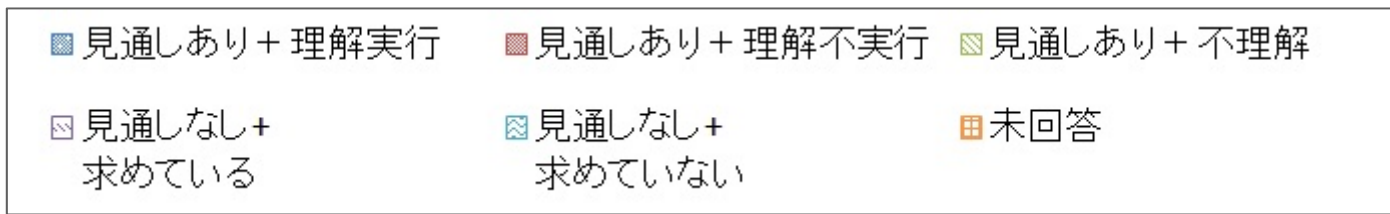
授業で身に付いた知識・技能	因子5	コミュニケーション能力
	因子3	競い合う姿勢，課題遂行能力，自己理解・自信
	因子4	知識の活用能力
授業以外の活動で身に付いた知識・技能	因子1	コミュニケーション能力，課題遂行能力，競い合う姿勢
	因子2	一般知識の活用能力
	因子7	専門知識の活用能力
授業内外	因子6	起業家精神

知識・技能の獲得 年度間の比較

授業内	<ul style="list-style-type: none"> 因子5: コミュニケーション能力 因子3: 競い合う姿勢、課題遂行能力、自己理解・自信 因子4: 知識の活用能力
授業外	<ul style="list-style-type: none"> 因子1: コミュニケーション能力、課題遂行能力、競い合う姿勢 因子2: 一般知識の活用能力、外国語でのコミュニケーション能力 因子7: 専門知識の活用能力
授業内外	<ul style="list-style-type: none"> 因子6: 起業家精神

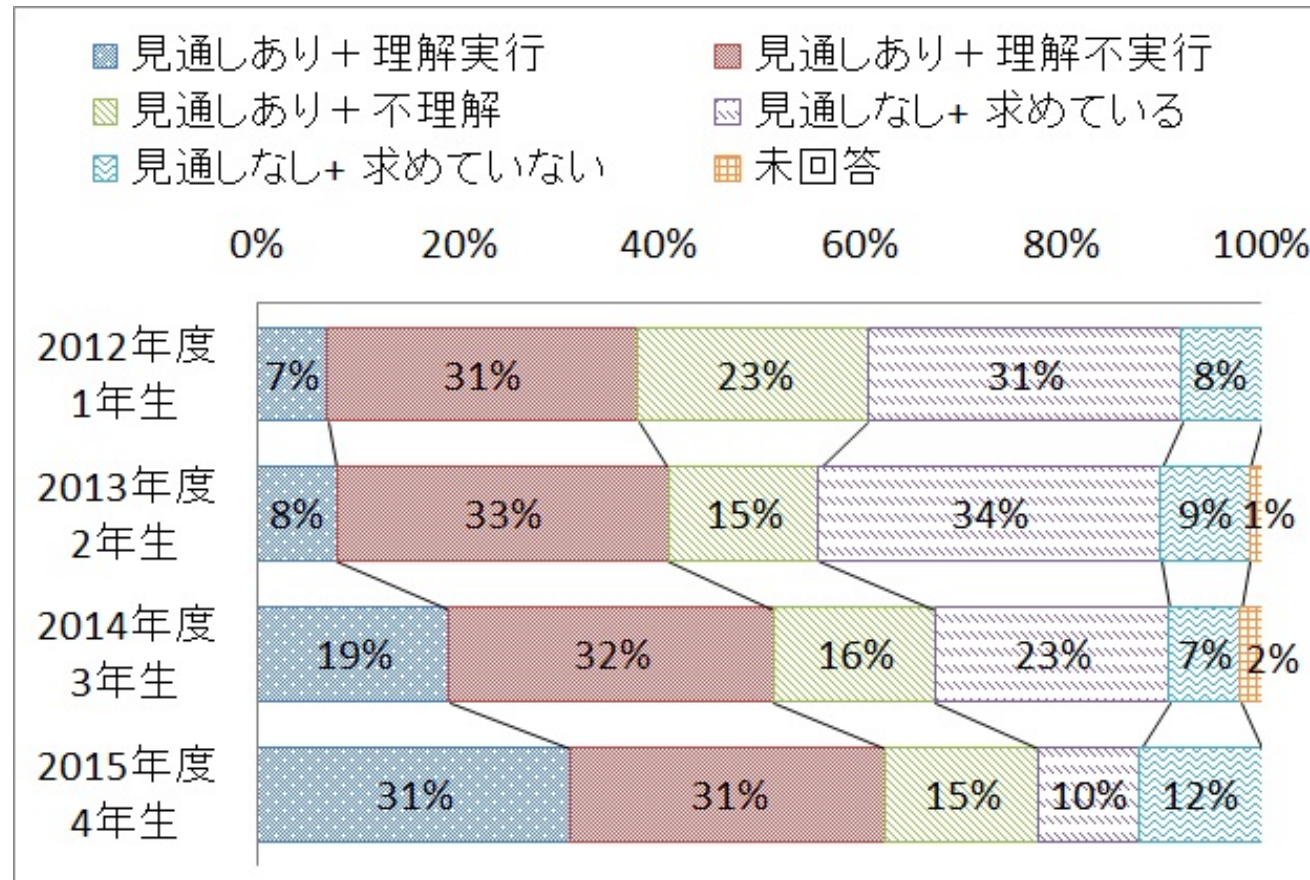


2つのライフ（将来展望と現状）年度間の比較



- 4年生になると「見通しあり+理解実行」が増加
- 一方、4年生の段階でも「見通しなし」と回答している学生はどの年度も20%近く存在

2つのライフ（将来展望と現状）の 経年変化（機関Aの結果）



- 「見通しなし」の割合が1年生よりも2年生の方が高い
- 3年生以上では「見通しなし」が減少するも、4年生の段階で依然20%程存在

教育データはビッグデータ
としては成立しにくい

通常のビッグデータ

- Googleのネコ
 - 200×200ピクセルの画像を1000万枚
 - 1.2×10^{12} B



- Twitterのツイート数
 - 350,000ツイート/分
 - 約 1.3×10^{13} B/year
 - 70B/ツイートの場合



教育のデータ

- 九州大学LA-C
 - 180,000logs/day
 - 1log→100Bytes
 - 6.57×10^9 B/year
 - ただし60クラスの蓄積
- 通常のビッグデータから比べると、データサイズが小さい



教育データは統合が難しい

- コンテキスト（文脈）が違えば，違う授業のデータを統合して分析することの意味はなくなる
 - 文脈：授業設計，授業環境，受講学生群の特性
- 文脈を包含したモデルをつくることにより，文脈が違うデータを，統一的にみることが可能になるかもしれない。

どうしてデータのサイズが 問題になるのか？

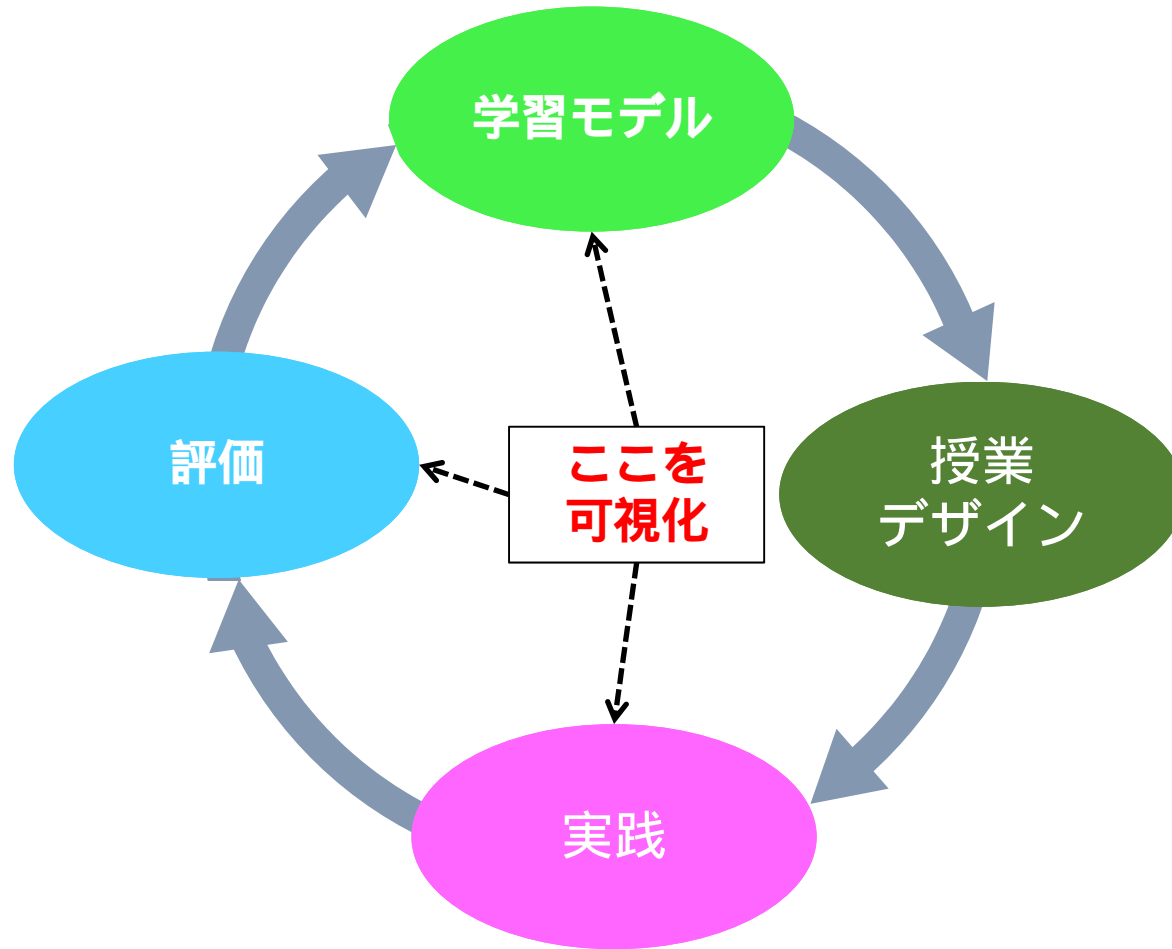
- データのサイズによって、仮説とデータの関係が変わってくる。
- データのサイズが小さい場合は、RQ (Research Question) から出発して、それを考えるためにデータを収集する
- データのサイズが大きい場合は、データにして語らしむ (つまりデータマイニング) という、データから仮説を見つける方法がとれる場合がある
- 教育データの場合、ある程度、データマイニングは可能。しかし、文脈依存性があるので、他の文脈でどうなるかはやってみないとわからない。

データを基礎に教育を改善し
ていく取り組みは意味がある

そうはいっても、 教育にとってデータは重要

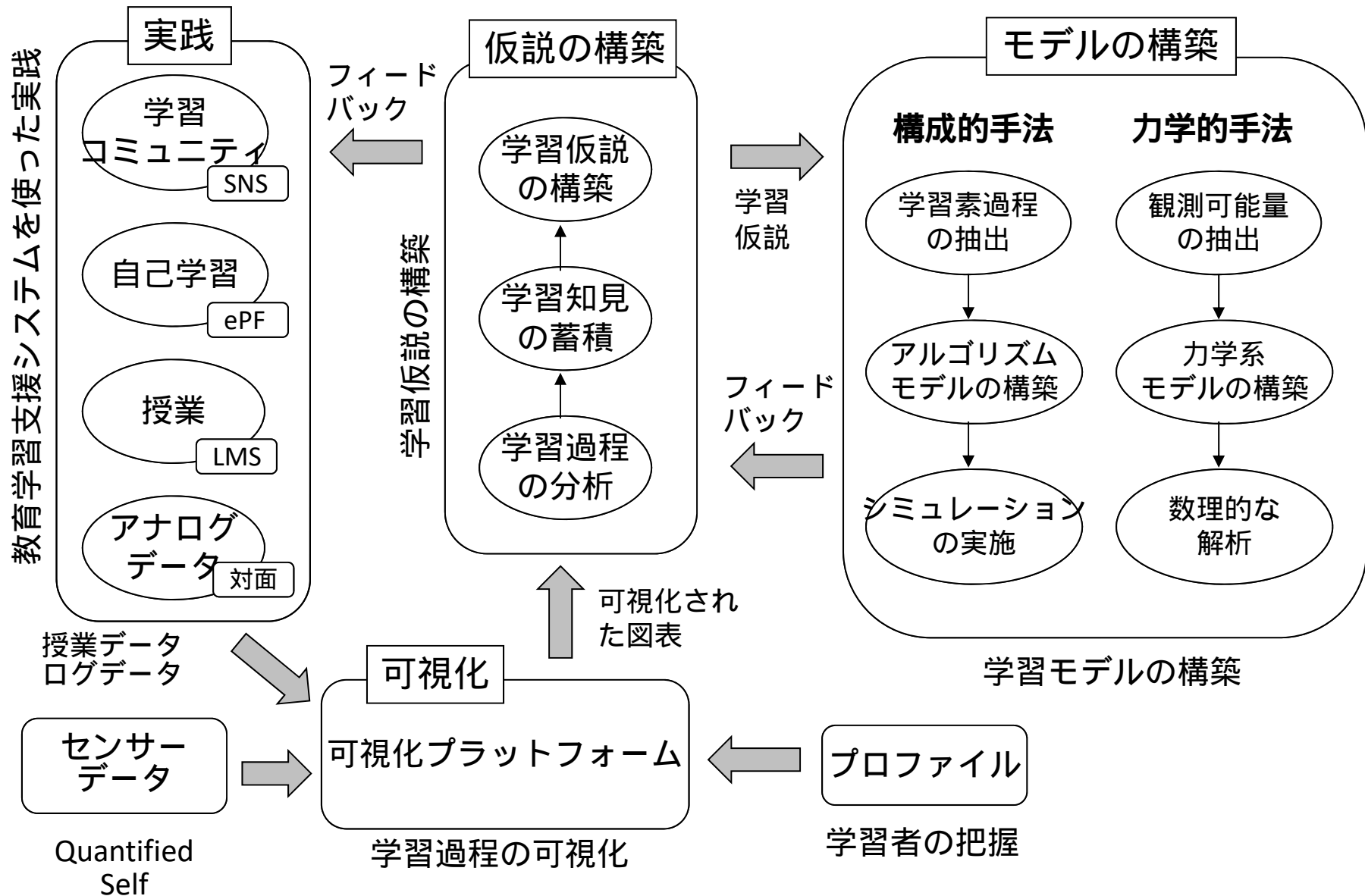
- **可視化**が重要
 - 今まで見えなかったものが見えてくる
 - 議論の共通基盤としての可視化
- **人間が判断**するための材料を提供
- Learning Analytics
 - 目的：授業の改善
 - 対象者：教員，学生
- 教学IR
 - 目的：教育システムの改善
 - 対象者：学長，部局長，教授会，事務局

一般的な教育サイクル



多くのデータを統合し有効利用するためにはモデルが必要

実証的な教育・学習研究方法論



企業がこの分野に貢献できる
ことは多数ある

大学側からの勝手な希望 1

- 汎用（簡単）可視化ツール
 - 現在は、統計解析等ができる教員が研究としてLAを実施している。
 - 教育を改善するためなら、だれにでも使えるようにしないと広がらない。
 - 統計解析等が得意な教員がつくった、解析用のモジュールを組み込めば、自分の授業でも同等の解析ができるような可視化ツールは作れないものだろうか
 - さらに、全国、全世界の教員が作った解析モジュールの集積場所があり、そこからモジュールをとってきて、自分の授業を可視化するといったことができるとうれしい

大学側からの勝手な希望 2

- 学習研究用ウェアラブル・センサーの開発
 - ウェアラブル・センサーを使うと、身体からの情報を利用し、学習状況をリアルタイムにモニターできる可能性がある
 - しかし、市販の3D加速度センサーや心拍計が入った、デバイスからの出力は、ある程度処理された後のデータしか手に入れることができない。（たとえば歩数）
 - 学習状況のモニタリングに使う場合、（少なくとも研究段階では）生データが必要となる。
 - また、デバイスとスマホまたはパソコンが1対1で対応していて、複数のデータを1箇所のパソコン等で取り扱えないので不便。

本日は言いたいこと

- 教育データはビッグデータとしては成立しにくい
 - 従来のビッグデータ分析とは違う方法論が必要かも
- しかし、データを基礎に教育を改善していく取り組みは意味がある
 - 学生がいかに学んでいるかに関して、いままでにない視点から可視化できる
- 多くのデータを統合し有効利用するためにはモデルが必要ではないか
 - コンテキストをモデルで乗り越える
- 企業がこの分野に貢献できることは多数ある
 - 可視化ツールとウェアラブルセンサー