

11.A Collaborative Convergence on Studying Reasoning Processes : A Case Study in Statistics

Marsha Lovett: A Collaborative Convergence on Studying Reasoning Processes : A Case Study in Statistics
Cognition and Instruction, pp.347-384 (2001).

この章は記憶実験で始まります。そこのあなた，読者が参加者です！次の短いストーリーを読んで覚えておいて下さい。後で再生の質問をします（知ってる人もいるでしょうが [Gick,Holyoak 1983]）

将軍

小国が独裁者の圧制にあった。独裁者は頑強な要塞に籠城して国を支配していた。要塞は国の中央にあり、周囲は農場と村に囲まれていた。要塞からは複数の道が放射状に伸びていた。ある時、将軍が独裁者を倒すため、軍隊を集結して要塞を陥落させる作戦を考え、自軍が一斉攻撃をかければ、要塞を陥落させることができる。しかし、要塞への道には地雷がしかけられ、大勢で通ると重みで地雷が爆発してしまう。そこで将軍は自軍を小勢に分割し、各隊を異なる道から同時に突撃させ、要塞陥落に成功した。

11.1 INTRODUCTION

- 本節の焦点
 - (若い)成人の統計推論スキルの改善
 - 学習場面外への転移
 - 統計推論の学習は相当困難 [Garfield,delMas 1991; Pollatsek,Konold,Well,Lima 1984]
- 知識の適切な表現の学習 (特殊・汎用偏らず) - 教授・認知理論の統合
 - 単一側面のみ持つ3つのアプローチ
 1. 単純な統計推論のパフォーマンスを説明する認知理論
 2. 現場で、生徒に単純な統計の問題を解かせる
 3. 新しい教授技術発展のための教室での研究 (理論導入・貢献なし)

本章の概要

- 複数のアプローチの収束
 - [Gick,Holyoak 1983] の類推適用
- 複合領域アプローチ [Brown 1992, Palincsar,Brown 1986]
 - 教授の改善に記憶・学習研究の理論
 - 教授介入, 実験室・教室でテスト
- 統計推論の定義
- 過去30の歴史
 - 70s: 統計推論の誤りを説明する認知理論
 - 80s: 生徒の統計推論を調べる実験的研究
 - 90s: 授業実践中で生徒にとって統計推論の困難な点

11.2 WHAT IS STATISTICAL REASONING?

「統計推論」

- データの要約・予測・結論付けのために統計ツール・概念を使用すること
- 例題1: ある町には2つの病院がある。大きいほうの病院では一日あたり約45人、小さいほうの病院では一日あたり約15人新生児が誕生している。新生児の50%は男の子であるが、毎日このパーセンテージは変わる。
ある年のある期間、どちらの病院でも60%前後男の子が生まれた日の記録が何日分がある。大きい病院と小さい病院、どちらのほうがこのような日が多いだろうか。あるいはこのような日の数はどちらも同じであろうか。 [Kahneman,Tversky 1972, pp.443]
解: サンプル数が少ないことから、小さい病院のほうに男の子が60%となった日が多いことが期待される。観察される分布もより分散するはず。
 - 統計推論の定義に適合
 - (a) 統計概念 (サンプル数, ばらつき), (b) 与えられたデータからの予測

- 例題 2：硝酸銀を雲に散布することで雨量が増加するかどうかを調査する，天候修正実験がフロリダで行われた．散布を行う / 行わない雲はランダムに割り当てられ，それぞれの雲からの総降雨量をデータとして収集した．変数「群」は散布した / しない雲，変数「雨」はそれぞれの雲からの降雨量である．この散布は降雨量を増やすか？この質問に答えるため，与えられたデータから適切な統計分析を行い，結果を解釈せよ．
解（の一部）：探索的・確証的なデータ解析が必要である．
- 統計推論の定義に適合
 (a) 統計ツールや概念，(b) 分析に基づく結論付け（ Fig 11.1）

11.3 HISTORICAL REVIEW

（理論の 1970 年代，実験の 1980 年代，教室ベースの 1990 年代）
比較のテンプレート（4 つの質問）

- 研究者は誰か？
- 認知科学者・教育研究者・教師・領域専門家，それらの組合せ
- 目標・方向性・手法に関係
- 目標は何か？
- 理論の発展・拡張，データ収集，教育的介入の試用，教育成果の評価 ... 等
- 文脈は何か？
- （文脈自体も多義だから）多様なレベルの回答の可能性
- どう適用されるか？
- 暗黙的（教授法の改良等）教育への応用

1970 年代の理論的焦点 ~ 何かしら心的表象の香りがした

- 不確実性下での判断：ヒューリスティックスとバイアス [Kahneman, Slovic, Tversky 1982]
- 確率・統計推論からの人間の行動（主に誤り），行動を導く心的プロセスを仮定
- 例題 1 の例：半数の被験者はどちらの病院も同じと回答
- 目標：エラーを規定する理論の構築
- 例：代表者ヒューリスティックス
 : A が B をどの程度代表するかにより，イベント A-B 間を確率的に判断
 : 実験・7 割が技術者，3 割が弁護士のものであるプロフィールの一枚を判断
 Dick は 30 歳男性，結婚していて子供はない．能力とモチベーションのある者には成功を約束する．彼は同僚からも慕われている．
 （Dick の職業を判断する情報がない，ならば技術者である可能性が高いのに...）
 被験者の 5 割が Dick は技術者，5 割が弁護士と判断
- 1970 年代アプローチに対する「4 つの質問」回答要約（ Table 11.1）
- 理論ベース，現実場面が注目されていない

Table 11.1 成人の統計推論の研究の歴史区分の対比

質問	1970s	1980s	1990s
研究者？	認知心理学者 社会科学者	認知心理学者	心理学者 教育者，指導者
目標？	ヒューリスティックスや バイアス使用を仮定する 理論の発展・検証	理論検証 しかし生徒の能力も解明	統計推論の困難さを解明
文脈？	心理実験室 実験課題の問題	教室外での統計学習 偽物っぽい実問題	教室で統計の学習
応用？	理論の構築・発展	理論の検証 同時に教授に応用	指導者に情報提供

1980 年代の実験的焦点 ~ 科学・数学教育研究方面の興味が認知に

- 認知理論・結果を応用した教授改善 [Klahr 1976]
- 統計推論でも生徒がどう学ぶか等が... [Fong, Krantz, Nisbett 1986] 等
- 従来授業以上の学習効果を予測，しかし生徒は教室外で活用できず [Fong et al. 1986]
 例：一般に，MLB で新人賞を取った選手は翌シーズンの成績が下がるというが，過去 10 年

を見ると明確である。AL では新人賞の 8 人は翌シーズンの成績が低下し、向上したのは 2 人だけである。NL でも 10 人中 9 人が翌シーズンの成績のほうが悪かった。なぜ新人賞を取った選手は翌シーズンの成績が下がると思うか答えよ。[Fong et al. 1986, pp.279]

統計推論の解答は事前 16% ~ 事後 37% に留まる、統計推論自体困難？

- 歴史的な重要性
 - × 動機は生徒の統計推論の理解
 - 動機は「人は統計概念を適用した抽象的規則を使う」仮説の検証
 - 授業教授の改善を直接指向していない
- まとめ (再度 Table 11.1)
 - 70 年代同様の理論ベースだが、より現実的な状況
 - 大半が大学生受講者、大学講師や教授デザインに

1990 年代の教室授業焦点 ~ 公式暗記から実践重視のテキストへ

- 実問題の解決の練習
 - データ解析用の統計ソフト
 - 統計的な現象のシミュレータを利用
- 研究上の疑問のシフト
 - 生徒は何を学び、どんな問題を解けるようになったか？
 - 生徒の得意 / 苦手の評価、指導技術の評価ではない
 - 例： [Garfield, delMas 1991]
 - : コイントスの系列で、最も起こりやすい / 起こりにくいものは？
 - : 発疹になる確率が 15% という注意書きのある薬をどう解釈する？
 - 事前・事後で成績向上、しかし特定の誤概念が残る
 - 生徒にとって困難な特定の統計概念 [Melvin, Huff 1992]
 - 統計ソフトの導入 [Cohen, Chechile 1997] 等

- まとめ (また Table 11.1)
 - 教室授業における生徒の統計推論に焦点
 - (なぜ困難なのか等) 理論面の試みが少ないが、教師に貢献

1990 年代中盤以後の統合的研究の島々

- これまで：単一アプローチ、プロセスより学習・推論結果に注目
- 複合アプローチ、プロセス指向の研究
 - 例：より一般的な心理学理論による確率推論の誤りのレビュー [Garfield 1995; 1994]
 - 解を導くプロセス・概念への注目
- DIAGNOSER [Hunt, Minstrell 1994]
 - 生徒は解・理由双方を選択、知識の面 (facet) ・理解が診断可能に
 - : 生徒の内的状態を推定 (70 年代風)
 - : 生徒に合わせたフィードバックによる介入も
 - 理論構築かつ実験、生徒の理解に注目、力学 確率に適用領域変更

11.4 A CONVERGENT ASSAULT ON STATISTICAL REASONING

(著者自身の統計推論学習に関する、CMU での最近の研究)

- 研究概要
 - 生徒の学習プロセスの理解、推論能力を改善する環境構築の双方を指向
 - 認知的理論を教授目標に適用、理論・実験・教室のアプローチを統合
 - 多岐にわたる専攻の学部一年生対象の講義
- 講義概要
 - 学習項目 (1,4 が研究の焦点)
 - (1) 探索的データ分析技術、(2) 標本のばらつきを理解、(3) 研究デザインの効果の評価、(4) 推理の統計手続の使用・解釈
 - 科学的でポリシーのある問題
 - 例：例題 2、どちらの薬がより効果的？、女性教授は収入が少ない？
 - 毎週 2 時間の講義セッション + 1 時間の計算機セッション
 - ペアで課題演習、類題を宿題として提示
 - 協同、実践課題などの 90 年代新手法導入

- 統計推論学習のモデル構築
 - (1) 同一問題を解く統計推論の認知モデル
 - (2) 検証済みの新指導法
 - (3) 計算機学習環境

タスク分析

- 知識とスキル
 - 認知理論による知識表現，知識の学習・使用の予測
 - 例：ACT-R [Anderson,Lebiere 1998]
 - :意味ネットワーク上の宣言的事実，生成規則中の手続き的知識
 - :使用された知識辺のトレースが問題解決ステップを記述
- 知識の分解（プロトコル分析）
 - 先の学習項目 (1) 探索的分析，(4) 推理的統計
 - 問題解決プロセスのステップ系列（例題 2）(Fig 11.2)
 - 生成規則と宣言的事実を作成，生徒のシミュレーション
- 知識の一般性 / 特殊性
 - 生成規則の例 (Table 11.2 一般的な規則の一部)
 - 疑問：生徒が学ぶルール的一般性のレベルは？
 - :特殊ルールを数多く学習，他の文脈に適用？
 - :特殊条件に直接適用できない一般的なルールを学習可能？

Table 11.2 統計推論の認知モデル中の一般的な生成規則

IF ゴールがデータ集合 d で質問 q を扱うことができる and 関係する変数が x と y and x のタイプは x-type and x は説明変数 and y のタイプは y-type and y は反応変数 and y-type 対 x-type データのグラフツールがグラフ g THEN x 対 y のグラフ g を作成
IF ゴールがデータ集合 d で質問 q を扱うことができる and 関係する変数が選択されていない and データ d 中に変数 x と y があり q と関係がある THEN x と y を分析するデータとして選択
IF ゴールがデータ集合 d で質問 q を扱うことができる and 関係する変数が x と y and 変数 x のタイプ (量的 / カテゴリ的) が判明していない THEN x のタイプを調べるサブゴールを設定

大規模な評価

- 評価手段のデザイン
 - 学習環境改善のための情報
 - 生徒の得意 / 苦手，学習場面外での振舞い
 - 複数選択問題を実施 (Fig 11.3)
 - :実験群 (受講) - 統制群 (テストだけ)
 - :事前 (講義前) - 事後 (講義後，他の統計講義にて)
 - 結果概要 (Fig 11.4)
 - :パネル A：結果全体，実験群は向上
 - :パネル B：実験群だけが向上した項目のみの結果 講義の成果があったスキル
 - :パネル C：両群とも向上しなかった項目のみの結果
 - :パネル D：天井効果が出た項目
 - 結果から講義の改善へ
 - :B=講義で習得：記述や推理の解釈，統計用語定義 (相関等)
 - :C=講義で未修得：適切なデータ表示の選択 (Fig 11.3)，結果考察
 - :D=特に学習せずとも容易：箱型図，散布図

生徒個々の詳しい学習

- 実験室実験（プロトコル分析）
 - 講義中とほぼ同じ内容，計算機使用

- プロトコルと計算機の操作をあわせて分析

問題例：男子ゴルフではプロは（51歳未満なら）通常のツアー，ないし（51歳以上なら）シニアツアーで競技する．貴君の友人が両ツアーでの賞金総額の違いを知りたがっている．友人は両ツアーの上位30人の賞金を記録した．変数「金」は各選手が昨年獲得した賞金，変数「ツアー」は選手が出場したツアー，通常のツアーは1，シニアツアーは2．変数「ランク」は選手の順位，優勝者は1である．

- プロトコルのコーディング

:ステップ（各ステップは Fig 11.2）(Fig 11.5 例)

1) おお，よし．で，必要なのは，彼は賞金総額，2つのツアーで選手が獲得した金額の合計，に違いがあるかどうか
 2) で，これが賞金だと思う，う，金は各選手の獲った賞金なわけだ．ツアーはどっちのツアーで選手が競技したか．えーと，ランクは実際いらぬよな，これを解くには，だよな？こうなつて，ええと，分からない．
 4) うう... 箱型図やるか
 5) (被験者は統計パッケージで箱型図を作っている) お，かっちょええ(笑)．やったじゃん．
 6) いいぞ，う，で，平均見りゃよし．シニアツアーの人たちが金少なそうに見える．うん，けっこう総額の変化は少ない．どっちにも2，3ほど外れ値っぽい，分からないえど，何人かがとつてるような，一回でかなりみたいな．

Fig 11.5 (Fig 11.2の)主要ステップのコードを貼ったプロトコル．ちなみにステップ3と7は出なかった．

- プロトコルの分析

- 最初の3ステップ，最後のステップはあまり頻繁に行われぬ
- エキスパートはステップ4から開始
- 最初の3ステップを飛ばした生徒はステップ4で間違えがち
- ステップ4の間違い

: (1) 最初の3ステップを飛ばす，(2) 適切な分析の選択で間違え

: 統計パッケージにより推測-テスト方略が可能

事例 a: お，よし．うん，よ～分からない，何が必要なのうちょっとこんな，グラフ，いいのか？うん折れ線，多分... お，う，ヒストグラム，棒グラフ，多分，箱型図？う，げ... うう，ヒストグラム，うん，データ表，うん...

事例 b: お，えっと，たぶん，ん... これ(データの変数)全部取り上げたら，それから，たぶん箱型図，統計の授業で，いつも使ってた，貼るとき，箱型図作って，で何が起こったか見る．だから，う，これ箱型図に，うん， $y \times x$ ．うお，変数ランクは30のカテゴリ，続けていいのか？いつもこれ悪かった，それで中止して，こうならないだろうから．

- 量的分析

- 1問あたりの分析回数：11回
 - : そのうち適切な分析：3回
 - : 一度行った分析の繰り返し：3回
- 生徒は適切な検索方略を使用しない
 1. 適切なグラフ選択の手続は未修得？
 2. 理解なしに選択を繰り返す最適でない方略を学習？

学習の実験的研究

- 計画プロセスを促進する学習環境

- 適切なグラフ選択，相当するスキル
- 生徒の計画プロセスの scaffolding

- 実装のための統制実験

- 生徒の計画スキル学習のデータ収集
- scaffolding が学習を支援するかテスト

- 手続き

- 2条件：最終のみ直接フィードバック
- 4ステップ：事前テスト，教示，問題解決 (Fig 11.6)，事後テスト (2種類の被験者：統計授業参加なし / 以前に45分の授業参加あり)
 - 最終のみ：グラフ表示，正 / 否のフィードバック
 - 直接フィードバック：グラフ選択前にフィードバック，変数選択を強制的に正解へ

- 結果
 - 事前～事後 (Fig 11.7A): 授業ではこの種の選択が自明
 - 正答フィードバックを得るまでの試行回数：終了のみ群が現象
- 1. 生徒はフィードバックからスキルを獲得した
- 2. 中間的なフィードバックは学習を支援した
- 通常の統計パッケージは
 - (1) 生徒のグラフを解釈する能力に依存
 - (2) 生徒の選択のどの側面が正しくないのか示さない
 - (3) 生徒の間違いを正す情報を与えない

11.5 A CONVERGENT ASSAULT: PUTTING IT ALL TOGETHER

- 多様な側面からの統計推論学習の理解
 - 複合アプローチ (理論と実験, 実験と教室など)
 - 主要な結果
 1. 一般的スキルに注目したタスク分析：様々な場面に転移可能
 2. 大規模評価：特定のスキルのための向上
 3. プロセスに基づく生徒にとっての困難さ：適切な分析の選択
 4. 実験的研究：中間的なフィードバックによる支援
 - 学習環境への導入
 1. 汎用性の高いスキーマへの注目のため，目標を明示 (Fig 11.8 インタフェース)
 2. 生徒の計画プロセスの scaffolding
- 「将軍の方略」を動かす
- そつえば最初のストーリー，議論してませんでしたね
 - 共同のタイミング：作成成功には各進撃隊のコミュニケーションが必要
 - 協同研究も同じ