

第37回中部地区英語教育学会・三重大会
2007年6月24日

英語教育研究法セミナー2
実験研究をすすめるときに

酒井英樹(信州大学)
sakaih@shinshu-u.ac.jp

ウェブサイトはこちら (資料 DL が可能です)
<http://www.urano-ken.com/research/seminar/>

研究を進めるときに

- リサーチクエスチョン(研究課題)の設定
 - 経験による疑問、先行研究のギャップ
- 研究の計画
 - 参加者、道具(テスト、アンケート)、分析方法
- 結果
 - 記述統計、推測統計
- 考察
 - リサーチクエスチョンに対する回答

変数 (variables)

- 独立変数 (independent variables)
 - 実験研究における処遇 (treatment)
 - ○○の効果
- 従属変数 (dependent variables)
 - 観測する変数
 - 何を測定しようとするのか。△△に及ぼす影響
- 研究課題
 - 独立変数と従属変数の関係
 - 独立変数が従属変数に及ぼす影響

なぜ研究するのか

- 理論・仮説を検討すること
 - 知の創造 (creation of knowledge)
 - 知を疑い、自ら検証して、知の構築、または補強に貢献しようとする
- 理論・仮説から導き出される具体的な手法等を検討すること
 - 例えば、音読の効果調べたとする。
 - 黙読よりも音読が、読解力向上に効果的であった。
 - なぜか。黙読と音読のプロセスに違いはあるか。その違いのうち、どこが読解力向上につながったのか。どんな読解力が向上したのか。

I家の食卓は、なぜ面白いのか。

- はがきやビデオで送られるアイデア(個別事象)
 - 個別事象の報告
 - 現象の描写
- 実証
 - 再現性
 - 信頼できる事象ということ
- 説明
 - 理論的説明
 - なぜそうなるか、という説明
- 応用
 - 理論に基づく現象の応用
 - 理論に基づく現象の予測
 - この場合も同じようにいかせるが、この場合にはうまくいかない。

操作的定義 (operational definition)

- 構成概念 (constructs) と変数 (variables)をつなぐ定義
 - 構成概念
 - 抽象的
 - "the underlying constructs that they [variables] represent" (Brown, 1988, p. 8)
 - 変数
 - 具体的
 - "a variable is something that may vary, or differ" (Brown, 1988, p. 7)
- 操作化 (operationalization)
 - Explanation of "how each variable is being defined with respect to the construct in question" (Brown, 1988, p. 8)
 - 操作的定義 (operational definition)

操作的定義 (operational definition)

- 構成概念 (Constructs)
 - まずは、何を調べようとするのかを明確にする
 - 例 「読解力」→「まとまった英文を読み、書かれた事実関係を把握する能力」
- ↓ 操作化 (operationalization)
- 変数 (Variables)
 - 構成概念を、観察可能な、量的に表現されるように、操作的に定義したものに関するデータ
 - 例 「英文を読ませ、書かれた事実関係について質問を与え、選択肢の中から適切なものを選ばせる英検式のリーディングテストの合計得点によって示される力」

実験研究の種類

- 中学校3年1組の英語を担当することになった。あるリスニングの指導法を3ヶ月間毎回実践した。7月に、英検3級リスニングの過去問題 (20問) を実施した。平均点が16.0点であった。この指導法は、効果があったか。



もともとリスニング力が高かったかもしれない

・指導の前に、事前テストを行おう。



実験研究の種類

- 4月に平均点が10.0点だったのに、7月には平均点が16.0点であった。この指導法は、効果があったか。



そのリスニング指導を受けていなくても、リスニング力が向上したかもしれない。(maturation, Brown, 1988, p. 32)
事前テストを行ったので、事後テストは二回目。練習効果があったかもしれない。(practice effect, Brown, 1988, p. 35)

- ・リスニング指導以外の点では同じ指導をする統制群を設けよう。
- ・別のリスニング指導を行う対照群を設けよう。



実験研究の種類

- 実験群 10.0 点 → 16.0 点 6点アップ
 - 統制群 7.0 点 → 10.0 点 3点アップ
- 実験群の方が6点アップしている。実験の処遇が効果があったといっけよいか。



実験群には良い生徒が集まっていたのかもしれない

- ・実験群と統制群を同等にしよう (実験研究の場合)。random assignment.
- ・実験群と統制群をあらゆる点から同等かどうか確認しよう (準実験の場合)。matching (Seliger & Shohamy, 1988, p. 142)



実験研究の種類

	要素	前実験研究1	前実験研究2	準実験研究	実験研究
1	事前テストと事後テスト	○	×	○	○
2	統制群と実験群	×	○	○	○
3	Random assignment	×	×	×	○

前実験研究2は、例を出さなかったが、他のクラスと7月のテストだけで比較をしようする場合がある。

研究計画における留意点

- 十分な人数の確保
 - 少ない人数 → 検定力の低下
 - 特に、統計的処理を行う場合
 - 「正規分布」の確保のため
 - ばらつき確保のため
 - 例、被験者内計画 20人
- 信頼性の高いテスト・アンケートの使用
 - 低い信頼性 → 検定力の低下
 - 低い信頼性 → 何を測定しているのかがあいまいになる
 - クロンバック α (手計算、SPSS、ウェブ上のソフト)

信頼性と妥当性

● テストの信頼性、妥当性

- 信頼性・・・「安定した結果を提供しているか」
 - 採点者間信頼性(一致度、相関係数)
 - 手順の詳細なマニュアル化
 - 練習
 - 内的一貫性(クロンバック α 、項目・合計相関 など)
- 妥当性・・・「測定したい力を測っているか」
 - 表面的妥当性(テスト項目の吟味)
 - 構成的妥当性(構成概念との対応)
 - 並存的妥当性(他のテストとの相関)

物語作りタスクの手順

- ・チャートを読む。
- ・「これから、『ユキは昨日何をしたか』という物語作り・タスクを行ないます。チャートは、昨日の朝8時から夜10時までの時間が示されています。」
- ・「絵を裏返しにして、チャートの上に並べなさい。」
- ・「時間ごとに並べた絵は、昨日のユキの行動を表しています。1枚ずつ絵をめくりながら、昨日ユキが何をしたか、順番に英語で話してください。」
- ・「それぞれの絵の間の行動や流れを想像して、物語を付け加えてください。」
- ・「例えば、3時に勉強しました。4時に、おながすいたので、料理をして、食べました。または、3時に勉強しました。4時に、お昼を食べていなかったことを思い出しました。レストランに行き、スバグッティーを食べました。というように、あなた自身言葉を加えてください。」
- ・「ユキは昨日7時に起きました。『Yuki got up at seven yesterday.』という英語から始めてください。それでは、『ユキは昨日何をしたか』 What did Yuki do yesterday? という物語作り・タスクを始めます。」

アイデアユニット・() の番号は生徒番号

1. Mike has a part-time job	<ul style="list-style-type: none"> ● マイクは新聞配達をしている 1×20 (741) ● 毎朝は早い時間に新聞を配っている 1×20 30一語ずついるけどよしとする (731) →3が一つ前の文と合体している。音声情報を利用できない? ● マイクは新聞配達のパートアルバイトを毎日の早朝にやっている。102 030 (701, 729, 717) →3が一つ前の文と合体している。 ● マイクの父は新聞記者です。1×2× (739) → part-time = father? deliveryの意味が理解できない。 ● マイクは毎朝は早くから自転車で新聞配達をしている少年だ。1×203 040 (725) ● マイクは新聞の仕事をしている。1×2× (708) ● マイクは自転車に乗って朝早くから毎日新聞を配達している男の子です。1×203040 (527) →1-4が一緒になった。
2. He is a newspaper delivery boy.	<ul style="list-style-type: none"> ● マイクはニュースを読んだ 2× (740) → deliveryの意味がわからない。 ● マイクはよく新聞を読む人です 2× (511) → deliveryの意味 ● マイクはボランティアの新聞を読んだ。1×2× (528) ● 私の仕事は新聞記者です。1×2× (525) ● マイクは毎日新聞を見ることが日曜だった。1×2×3× (532) ● マイクは新聞会社に勤めていて × (536)

SPSS 信頼性

Reliability Statistics

Cronbach's Alpha	N of Items
.892	40

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
PD1T	15.19	59.048	.588	.887
PD2T	15.33	59.018	.585	.887
PD3T	15.46	61.142	.388	.891
PD4T	15.12	60.400	.423	.889
PD5T	15.43	60.189	.455	.889
PD6T	15.40	58.922	.520	.896
PD7T	15.55	61.890	.256	.892
PD8T	15.51	60.392	.481	.888
PD9T	15.60	61.507	.355	.891
PD10T	15.55	60.725	.440	.890
PD11T	15.30	58.744	.619	.886
PD12T	15.19	59.128	.575	.887
PD13T	15.44	60.608	.397	.890
PD14T	15.80	61.847	.331	.891
PD15T	15.37	61.189	.301	.892
PD16T	15.50	61.486	.292	.891
PD17T	15.47	59.923	.503	.888
PD18T	15.19	61.824	.217	.893
PD19T	15.16	61.882	.213	.893
PD20T	15.16	61.886	.212	.893
SD1	15.35	58.573	.513	.896
SD2	15.02	60.015	.535	.888
SD3	15.08	59.885	.512	.888
SD4	15.23	59.195	.560	.887
SD5	15.22	59.203	.561	.887
SD6	15.57	61.637	.309	.891
SD7	15.38	59.550	.523	.888
SD8	15.38	59.878	.479	.890
SD9	15.65	62.338	.250	.892
SD10	15.19	61.188	.303	.891
SD11	15.44	62.811	.096	.895
SD12	15.12	61.119	.324	.891
SD13	15.17	60.268	.428	.888

テスト作成のときに・・・

- 項目数の確保
- 適切な困難度
 - 項目困難度(正答率 0.10 以下、0.90以上)
 - 項目弁別度(上位群の平均 - 下位群の平均)

予備実験の目的

- 実験手続きの確認
 - 時期の検討、テスト手続きの検討、処遇の検討。
- テストの改良。
 - 簡単すぎたり、難しすぎでは意味がない。
 - 項目分析
- テストの同等性の検討。
 - ←実質難しい…。
- 実験の可能性。
 - 統制群なしの計画で行ってみる。
 - 理論的枠組みの再検討が必要になることも。

検定の考え方 (Brown, 1988)

標本 (sample)		母集団 (population)	
記述統計		推測統計	
Group 1 ≠ Group 2 ? $M_1 \neq M_2$?		Population 1 ≠ Population 2 ? $\mu_1 \neq \mu_2$? (95% confidence interval)	
Group 1	n_1 M_1 / SD_1	$SE_1 (SD_1 / \sqrt{n_1})$	$\mu_1 = M_1 \pm 2SE_1$
Group 2	n_2 M_2 / SD_2	$SE_2 (SD_2 / \sqrt{n_2})$	$\mu_2 = M_2 \pm 2SE_2$

n = 人数、M = 平均、SD = 標準偏差、SE = 標準誤差、CI = 信頼性区間

データのスクリーニング

- 正確なデータ入力
- 欠損値
- 外れ値 (一変量、多変量)
- 正規性
- 直線性 (linearity)
- 等分散性など
- 用いようとする統計手法で要求する条件

データの正規性の前に

- ところで、標準偏差と偏差値の違いって???
- A: M = 50, SD = 10 の分布で、70 点とりました。
- B: M = 50, SD = 20 の分布で、70 点とりました。
- 同じ???
- 標準化してみよう
 - まずは、平均との差を計算 (70 - 50 = 20 点)
 - 平均の差を標準偏差で割ってみよう (20 / 10 = 2, 20 / 20 = 1)
 - つまり、A は平均から +2SD なのに対して、B は平均から +1SD
 - z 得点は、A = 2.0, B = 1.0。(M = 0, SD = 1 に標準化)
 - t 得点(いわゆる偏差値)は、A = 70, B = 60。(M = 50, SD = 10 に標準化)

データの正規性

- さて、応用です
 - z 得点。A = 2.0 は、全体の分布で、どの程度に位置してる? (2.27%)
 - z 得点。B = 1.0 は、全体の分布で、どの程度に位置してる? (15.87%)
- z 得点
 - 5% 水準。z = ±1.96
 - 1% 水準。z = ±2.33
 - 0.1% 水準。z = ± 3.29

データの正規性の確認

- ヒストグラムを描いてみよう
- 平均と標準偏差を眺めてみよう
- 平均、中央値、最頻値を比べてみよう
- 傾きや尖り具合を確認してみよう
 - 歪度と尖度 (Tabachnick & Fidell, 2001, pp. 73-74)
 - |z-skewness (Skewness/SE)| と |z-kurtosis (Kurtosis/SE)|
 - SE for Skewness = $\sqrt{6/N}$, SE for Kurtosis = $\sqrt{24/N}$
 - < 1.96 at 5%
- 外れ値 (outlier)
 - z 得点に換算 (z = (得点 - 平均)/標準偏差)
 - 3.29 at 0.1%

結果の書き方

Table 4
Descriptive Statistics for Identification of Errors and Correction in the Test Without Hints, and Rule Statement and Correction in the Test with Hints

	Identification without hints	Correction without hints	Rule statement	Correction with hints
M	4.61	3.64	7.39	8.68
% of the 12 items	38.4%	30.3%	61.6%	72.3%
95% CI				
Lower limit	3.72	2.53	6.48	7.79
Upper limit	5.50	4.75	8.30	9.57
SD	2.30	2.86	2.35	2.29
Skewness	0.40	0.55	-0.88	-1.06
SE of skewness	0.44	0.44	0.44	0.44
Kurtosis	0.04	-0.54	0.93	1.11
SE of kurtosis	0.86	0.86	0.86	0.86

Excel の例・入力は回答そのものを

番号	1	2	3	4	5	6	7	8	9	10
1	4	3	2	2	2	3	4	2	1	4
2	1	3	2	3	2	3	4	3	1	1
3	4	1	2	2	2	3	4	3	1	4
4	4	2	2	1	2	3	4	3	3	1
5	4	2	2	3	2	3	4	3	1	4

選択肢ごとの回答者数はcountif関数を使って[=countif(B2:B5, 1)]

	1	2	3	4	5	6	7	8	9	10
1	17	29	4	37	23	3	6	8	112	62
2	9	41	146	35	94	4	7	9	15	6
3	5	74	4	50	30	139	21	141	25	4
4	129	16	6	37	9	14	126	2	7	88

Excel の例・採点は if 関数を使って [if(B2=B\$1,1,0)]

正答	4	3	2	1	2	3	4	3	1	4	合計
1	1	1	1	0	1	1	1	0	1	1	8
2	0	1	1	0	1	1	1	1	1	0	7
3	1	0	1	0	1	1	1	1	1	1	8
4	1	0	1	1	1	1	1	1	0	0	7
5	1	0	1	0	1	1	1	1	1	1	8

Excel の分析ツール

- 「分析ツール」がインストールされているか確認してみよう
 - 「ツール」→「分析ツール」
 - もしなければ、「ツール」→「アドイン」により、「分析ツール」をインストールする。
- 「分析ツール」・・・かなりの分析ができる
 - 例 「基礎統計量」と「ヒストグラム」

平均	6.7875
標準誤差	0.14579739
中央値 (メジアン)	7
最頻値 (モード)	8
標準偏差	1.844207317
分散	3.401100629
尖度	-0.163510296
歪度	-0.420091332
範囲	9
最小	1
最大	10
合計	1086
標本数	160
信頼区間(95.0%)	0.287949121

