

## 第六章 研究デザインと内的妥当性への脅威

目的：主要な前実験、疑似実験、真の実験研究デザインを提示しクリティークする。そして、よく出会う内的妥当性を脅かすような配置の例を示す。

内的妥当性とは？

- 内的妥当性とは、従属変数の変化の原因が独立変数の操作に帰せられる度合をいう。従属変数の変化は独立変数以外の要因（実験統制外の要因）によってもたらされる可能性もあり、「そうしたことがない」と断言しうるほど内的妥当性が高いという。

研究デザインとは？

- 独立変数の効果（一時的分散）を研究するための実験的デザイン。

---

### 前実験デザインと疑似実験デザイン

---

#### 記述的研究

- 対象の特徴、人や生物の下位クラス、または自然あるいは人工の出来事や事象を記述する非実験的研究（ex. 行動観察など）

その下位クラスが何を共有しており、何がそれらを他の下位クラスと区別しているかを明確に示さなくてはならない。

#### 記述的研究を読む場合

選択されたサンプルや事象は、特定の下位クラスの真の代表である保証されることを望む。

記述的情報がどのように得られたのか（面接、観察、調査質問紙）をしり、その客観性と妥当性が確信されることを望み。

記述が、その下位クラスの成員に共有される証拠を探しその特定の下位クラスが他と明確に違い、独特で示唆的である証拠を探す。

-----  
一群・事前テストと事後テスト  
-----

例題

- ある研究者は、不安の主訴により治療に参加している患者のグループへの長期心理療法を行う。サンプルは40人の長期治療者（1年以上と定義される）を完了した患者で構成される。完了したというのは、患者と治療者の双方が互いに同意した集結を意味する。査定は、終結時にテイラー顕在性不安検査(Taylor, 1953)によって実施される。グループの平均スコアは、正常域にあると報告され、研究者はポジティブな効果が示されたと主張する。

この例題の中で扱われている査定をどのようにすれば有効に導き出せるだろうか？

- 事前テストと事後テストの比較

・**1群・事前テストと事後テスト**：患者の開始時の症状の重篤度をあらかじめ測定し（事前テスト）、治療が完了し後の測定（事後テスト）と比較することで患者の何らかの変化の経験を測定することができる。 事後テストのみであるならば比較することができない

図式化すると・・・

事前                  事後  
O    X    O

\* Xは処理、Oは観察、または測定を示す

ここでは、事前と事後での変化を測定することはできるが、それが治療以外の要因に起因する可能性もある。そこで、治療を受けていない統制群を加える必要がある。

・**2群・事前テストと事後テスト**：ここに心理療法を経験していない患者により構成される無治療者の統制群を加える。  
この無治療者の群の場合も、治療前の査定を行い、治療後の査定と比較して等価であるということを示さなければならない

図式化すると・・・

| 事前 |   | 事後 |
|----|---|----|
| O  | X | O  |
| O  | - | O  |

\* ダッシュ（-）は無処理を示している

治療による要因を測定することができるが、不等価統制群であるためサンプルをランダムにしなければ、妥当性のあるデータとは言えない。

-----  
実験デザイン  
-----

例題2 例題1を発展させて、心理療法Aと心理療法Bではどちらが効果があるかを測定する。

**ランダム配置・2群処理群・事前テスト・事後テスト**

: 不等価統制群のデザインを避けるために、参加者をランダムにその半数を治療法Aに、残りの半数を治療法Bに割り当て、事前テストと事後テストを行うことにより、サンプルの統制をとることができて、しかも治療法Aが治療法Bよりも優れていることを示すことができる。

図式化すると・・・

| 事前 |       | 事後 |
|----|-------|----|
| 0  | $X_1$ | 0  |
| 0  | $X_2$ | 0  |

\*  $X_1$ は治療法A、 $X_2$ は治療法Bを示す。

ここでも、無治療者の統制群を比較しないと、その要因を心理療法に起因することができない。

**ランダム配置・3処理群・事前テスト・事後テスト**

| 事前 |       | 事後 |
|----|-------|----|
| 0  | $X_1$ | 0  |
| 0  | $X_2$ | 0  |
| 0  | -     | 0  |

ここで、治療法Aと治療法Bが比較することができる。

**ソロモン4群デザイン**

: 事前テストそれ自体が独立変数になる可能性もあり、ソロモン4群デザインは、このような効果を統制する効果がある。

- ソロモン4群デザインでは、2つの実験群と2つの統制群を設定する。いずれの場合にも、それぞれの群の一方が事前テストを受け、もう一方は受けない。参加者はランダムにランダムに割り当てられる。これによって、研究者は事前のテスト効果を見ることができる。

図式化すると・・・  
 ランダム割り当て：

| 事前 |       | 事後 |
|----|-------|----|
| 0  | $X_1$ | 0  |
| -  | $X_1$ | 0  |
| 0  | $X_2$ | 0  |
| -  | $X_2$ | 0  |

\*  $X_1$ は実験群、 $X_2$ は統制群を示す。

このように事前テストを受ける者、受けない者をランダムに割り当て比較することで、事前テストの効果を統制することができる。

#### ランダム配置・2群・長期間繰り返し測定

- 実効的なデザインには、1回以上の時間間隔をおいた事後テストが行われる。これは、人に変化をもたらす、単なる一時的な効果に終わらないようなデザインされた処理では、複数のフォローアップが有効である。例えば、禁煙プログラムなど。

| 事前 |       | 事後 | フォローアップ |
|----|-------|----|---------|
| 0  | $X_1$ | 0  | 0       |
| 0  | $X_2$ | 0  | 0       |

このデザインにより、その効果が一時的にすぐ消えてしまうかどうかを測定できる

#### 交差デザインとカウンターバランシング

- 心理療法のような変数の効果の研究には適さないが、より直線的でより一時的な効果には有効である。

| 事前 | 処理1   | 処理2   |   |
|----|-------|-------|---|
| 0  | $X_1$ | $X_2$ | 0 |
| 0  | $X_2$ | $X_1$ | 0 |

異なる処理にさらされていることによる干渉が内的妥当性の脅威になりえる。たとえば、順序効果、あるいは、処理×順序の交互作用効果などが起こりえる。

### 多重独立変数デザイン

- 3つ以上の独立変数を研究する際にこのようなデザインも用いられれば研究できる。2条件XとYの下で2つの処理をする2×2のデザインにおいて、その選択肢は、以下のように示される。いずれの場合も処理Aと処理Bの間の差や、条件Xと条件Yの間の差は、部分的の部分的な群差によって比較される。

|     | デザイン1          |                | デザイン2          |                | デザイン3          |                | デザイン4          |                |
|-----|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|     | X              | Y              | X              | Y              | X              | Y              | X              | Y              |
| 処理A | G <sub>1</sub> | G <sub>2</sub> | G <sub>1</sub> | G <sub>1</sub> | G <sub>1</sub> | G <sub>1</sub> | G <sub>1</sub> | G <sub>2</sub> |
| 処理B | G <sub>3</sub> | G <sub>4</sub> | G <sub>1</sub> | G <sub>1</sub> | G <sub>2</sub> | G <sub>2</sub> | G <sub>1</sub> | G <sub>2</sub> |

\* 処理と条件という用語は、ここでは大まかな交換可能なものとして用いる。  
 (Ex. Xはストレスあり、Yはストレスなし、または、Xは男性、Yは女性など)

起こりうる「誤差」

- ・ デザイン1：処理の比較は、G<sub>1</sub>+G<sub>2</sub>対G<sub>3</sub>+G<sub>4</sub>のみ。条件に関しては、G<sub>1</sub>+G<sub>3</sub>対G<sub>2</sub>+G<sub>4</sub>となる。  
 推定される「誤差」は統計的目的の為にさまざまな群内の変動を算出することにより推定される。
- ・ デザイン2：「誤差」は起きないが、人が完全に一貫していることはなく、処理や条件が同一であっても2回目は異なる遂行をするかもしれないことから生じる。
- ・ デザイン3：処理 (G<sub>1</sub>+G<sub>1</sub>対G<sub>2</sub>+G<sub>2</sub>) の被験者内、条件 (G<sub>1</sub>+G<sub>2</sub>対G<sub>1</sub>+G<sub>2</sub>) の被験者間で生じうる。
- ・ デザイン4：デザイン3の逆である。

### 1事例デザイン

- 1事例実験のデザインは、研究者が実験処理あるいはそれに相当するものを制御し、結果を説明し得るすべて剰与要因を統制しようとする実験である。これによって研究者は、この1事例、あるいはこの1群においての因果的推論ができる。

手続き

- 安定したベースラインを確立するまで行う
- 計画していた介入
- この後ベースラインAに戻る

ex. 理解できない言葉を話し続ける患者への治療的介入

- (A) 理解できないどんな言葉を話すのかを記録しベースラインを確立する
- (B) 計画していた治療的介入
- (A) 治療を中断し、ベースラインと比較

以上の査定の連続的な記録から、実験者は変化が介入の導入と除去に関連しているかどうかを決定することができ、グラフにすることで傾向線、傾き、データの変動性、つまり、この例でいうならば個人の治療的な有効性を検討することができる。

### 1事例デザインの变化形

ex.

BAB：介入 - 無介入（ベースライン） 介入  
ABCB：ベースライン - 随伴強化 - 半随伴強化 - 随伴強化  
AA1B1A1B：非服薬ベースライン - (A) - 偽薬(A<sub>1</sub>) - 服薬(B) - 偽薬(A<sub>1</sub>) - 服薬(B<sub>1</sub>)

などの变化形も考えられる。

-----  
1-5-2 内的妥当性への脅威の検出  
-----

#### 履歴

- 履歴とは、参加者に生起し影響する外的事象で、これは、研究の内的妥当性を損ねるよう作用しうる。

#### 履歴 × 割り当て

- 生起する事象は、参加者の割り当てと交互作用し、偏った影響をもたらす。

#### 成熟

- 成熟は、研究期間中の参加者内部の変化をいう。

#### テスト体験（練習効果）

- テストを受けること、すなわち通常の規準測度としてのテスト使用は、それが自体が影響する。人々は心理的（かつ生理的）にテストに反応する。初回の経験から学び、次の機会によりよく遂行する者をいる。不安から、より悪い遂行になるものもいる。生理学的なテストに順応する者も言えれば、順応しない者もいる。

#### 多くの処理による干渉

- 参加者に2回以上の実験処理をうけることを求める実験もある。最初の処理は2番目の処理に影響し、その効果について何らかの結論を引き出す事を難しくする。

#### 測定用具（評定者）

規準の測定が評定者によってなされるとき、評定者が訓練され、動機付けられ、誠実で

認知心理学課題演習  
10月16日  
発表者 金田一賢顕

あり、一貫して信頼性があり妥当性のある評定を行えることが重要である。

抽出 × 平均への回帰

- 参加者選択と平均への回帰の交互作用が奇妙な効果を生み出しうる。