

測定・評価 926

特殊教育における単一事例実験法と統計的—臨床的有意に関する諸問題

成田 滋

国立特殊教育総合研究所

I 単一事例実験計画法

行動科学の発達と応用的行動療法に使われる技術の進歩によって、いろいろな研究上の計画法が考案されてきた (Baer, 1975). 行動分析における方法論も盛んに研究がなされ、例えば心理療法などで Leintenberg (1973), 精神病理では Barlow と Hersen (1973) らが時系列を使った研究を行い、ノンパラメトリックの分野で先駆的な役割を果たしてきた。今や $N=1$ (単一事例実験計画法) が行動分析の分野に広くゆき渡ってきたにもかかわらず教育学や心理学等の研究領域ではこの単一事例実験計画法はあまり一般的ではない傾向がある。これには研究における法則性や普遍性の究明という観点からは単一事例を扱う実験はこれらの目的を遂行することが極めて難しいということ、事例研究は記述的なものが多く、推測統計学等の対象となりにくいことなどにあると思われる。単一事例実験法によって得られたデータには母集団の特性を推測する手掛りはないからである。標準的な母数による統計法—パラメトリック法は、もちろん単一事例実験法では使うことが困難である。しかし、単一事例実験も実験計画に準ずるものとして近年研究が進み、その方法論が次第に確立されつつある。具体的には実験計画の誤差の制御、変数の操作、持越効果の偶然化、等の方略を導入して、独立変数と従属変数の関係を単一事例実験に応用しようとするものである。

Campbell と Stanley (1966) はいろいろな実験計画法を考察しながら単一事例実験計画法を "one shot case study", または "one group pretest--posttest design" と呼んでいる。これらの実験計画法は通称、時系列といわれるが、単一事例の時系列実験計画法はデータの収集手続については真の実験計画の場合とは違っているので準実験計画と呼ばれている。このような時系列を用いた実験の方法は特殊教育の研究にもいくつかの影響を与えてきた。それはベースラインと教育的処遇をセットにした反復測定であり、もう一つは実験上の妥当性を脅かすいろいろな要素に関する実験管理の原則にそった実験計画である (Kratowill, 1978)。

時系列を用いた行動変容に関する望ましい単一事例実験計画、あるいは少数被験者実験計画についてはいくつかの議論が存在する。この種の議論とは単に不適切な統計処理が行われるといったことではなく、どのような推測統計法が臨床的判断あるいは社会的妥当性などの対比で位置づけられるか、もう一つは追認の方法としての推測的統計法の手続はどのようなものであるべきか、ということである。本稿は前者の論点を考察するものである。

II $N=1$ 実験と妥当性

外的妥当性は被験者のグループに対する刺激提示条件を伴う実験における発見の一般化、普遍化にかかわることである。実

験上の性格から単一事例実験における外的妥当性の問題は深刻である。被験者を無作為に抽出することが無意味であるばかりでなく、刺激—状況設定が数限りない可能性から生れること、さらに被験者内の外的妥当性を考慮にいれることが不可能であることである。しかし、別の見方をすれば外的妥当性の要件は刺激の効果や影響がどのような重要性を有するかにあるともいえる (Kazdin, 1977)。例えば、治療教育における最も大きい関心は行動上の変化が被験者の生活においてはたして有意であるかということであるが、この場合の有意とはその変化が顕著であって統計的な検定を必要としないことである (Kazdin, 1977)。最近の研究では治療上の有意を教育的処遇との関連で問題にするときの基準とはなにかが盛んに論議されている。単なる変化や変容を指摘するだけでなく、ある処遇がその被験者にとって質的な変容をもたらしたか否か、能力的な側面からみて社会が求める基準的レベルに達するかどうかの問題にされる。

教育的処遇の効果をもどのように評価するかによっていろいろな臨床的有意の基準が提案されている。一つは実験的基準ということで処遇が行動にどのような変容をもたらしたかにかかわる。もう一つは臨床的基準であって行動上の変容が社会的にあるいはその被験者の日常生活上、意味あるものにかかわる。この基準の設定は容易ではない。というのは臨床的基準を満たしたかどうかの判断は行動の変容がいまだ偏っているか、あるいは許容される範囲かの解釈にかかっているからである。

実験的基準は多くの場合、被験者中心の行動の評価ということである。すなわち、ベースラインと処遇のフェーズに明らかな違いがあるかどうかである。評価基準は、実験者があらかじめ設定する。行動の変容がベースラインとは反対の方向に向かう、あるいは、明白な違いがグラフ上にあらわれている、などの場合である。また、統計的有意が二つのフェーズに見出される場合も含まれる。

一方、臨床的基準には主観的評価と基準的評価がある。主観的評価は家族、教師、友達など被験者の身の回りの者によってなされ、行動上になんらかの変化がみられるかどうか報告されることである。被験者を取り巻く者の観察はもちろん大事なことには違いないが、それだけに依存することも危険である。例えば、静に椅子に坐る行動はなにかを学習する際の前段階行為と考えられ、教師にとって望ましい態度にはちがいないが、静かに坐れるようになることが成績を向上させることには必ずしも結びつかないことがわかっている。むしろ、何かを学ぶことによって、静かに坐ることが多くなるということが少しづつ判明している事実があるので主観的評価の意味は薄れる (Kazdin, 1977)。

基準の評価は主観的评价よりも厳しく、その被験者の属するグループの規範、社会的に期待される基準にそった行動ができるかということである。しかし、この基準は通常、実験ではあまり採用されない。子供の障害の程度に照らしてむしろ不適切なことが多い。障害が重い場合はなおさらである。

III 実験

実験 I

被験者Cは職業訓練センターに通いながら21の工程から成る軟水器の組立作業に従事するCA 14:6, MA4:8の男性である。強化方法は軟水器の1セットの完成ごとに1セントを自分で貰う(自己報酬)とスーパーヴァイザからの社会的賞讃の二種類である。訓練セッションでは強化法は午前と午後は無作為に振り分けられ、その結果は次のようになった。

週	1	2	3	4	5	6
自己報酬	18	16	17	20	22	23
社会的賞讃	15	13	16	15	16	20

(値は週的最頻値)

ウィルコクソン検定の結果、両側検定で有意 ($.01 < P < .05$) となり二つの強化方法によって作業能率に違いがあることがわかった(詫間, 成田, 中村, 1986)。

実験 II

同じ訓練センターでCA, 施設生活, 訓練プログラム参加期間など、かなり等質な被験者13名を対象に軟水器の組立作業を与えその完成数が性差によって影響があるかを調べた。

男性	1	2	3	4	5	6	
組立数	13	15	16	12	20	15	
女性	1	2	3	4	5	6	7
組立数	18	16	17	20	22	23	18

(値は月の最頻値)

結果はマン・ホイットニ検定で組立作業に男女の差があることが判明した ($.01 < P < .05$) (詫間, 成田, 中村, 1986)。

以上の例は単一事例実験に準じた実験における観測データが統計的に処理・検定をおこない、統計的有意だけを問うものである。しかし、統計的有意が明かになったからといって独立変数の違いが作業能率に実質的にあらわれたか、強化方法が真の意味で作業行動に影響をもたらしたかは早計に判断することは困難である。これらの関係は以下の四分割表に示すことができる(成田, 1985)。

		統計的	
		有意	非有意
臨床的	有意	A	B
	非有意	C	D

桁A, Dは統計的、臨床的に合致し、異論はない。しかし、Bについては臨床的に有意であるが、統計的検定では有意とは

ならない。逆にCは統計的には有意であっても臨床的に有意ではない場合である。前述の例では結果は共にCの領域における問題であることがわかる。

実験は被験者にとって日常生活とかけ離れた環境でなされることがしばしばあるので、観測データが日常場面での出現とは異なることが考えられる。実験条件の均一化や制御などによって実験誤差が少ないことは実験室などの作作的条件下における実験の強みであるが、被験者の行動変容はいろいろな対人関係において意味を有するとする考え方にたてば教室とか家庭など自然な環境での実験的試みが本来望ましいといえるかもしれない。統計的有意では実験誤差を除去または制御という内的妥当性(Campbell & Stanley, 1963)は満たされているが、社会的適合性、社会的妥当性という視点からは弱い。一方、被験者の日常生活における実験ではさまざまな誤差は混在してくるが、長期間に渡る観察、複数の観察者、異った場面での観察など、実験誤差を無作為化することによって内的妥当性のある程度満たすことができる。四分割表におけるCについての解釈はKazdin(1982)が指摘するように臨床的有意の統計的有意に対する優位という立場が妥当と思われる。かといって統計的有意を軽視してもよいということではない。統計処理は主観的観察だけでは不十分で偏った解釈を防ぐという点で臨床的有意(非有意)を補完する役割を有するといえることができよう。

文 献

- Baer, D.M. (1975). In the beginning there was the response. In E. Ramp & G. Semb (Eds.). Behavior analysis: Areas of research and application. Englewood Cliffs, N.J.: Prentice-Hall.
- Barlow, D.H. & Hersen, M. (1973). Single case experimental designs: Use in applied clinical research. Archives of General Psychiatry, 29, 319-325.
- Campbell, D. T. & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.
- Kazdin, A.E. (1977). Assessing the Clinical or Applied Importance of Behavior Change Through Social Validation. Behavior Modification, 4, 427-451.
- Kazdin, A.E. (1982). Single-case research designs: Methods for clinical and applied settings. London: Oxford Press.
- Leitenberg, H. The use of single-case methodology in psychotherapy research. Journal of Abnormal Psychology, 82, 87-101.
- 成田 滋(1985). 障害児教育のためのやさしい統計情報処理入門: ノンパラメトリック法を中心に. 横須賀, 国立特殊教育総合研究所.
- 詫間晋平, 成田滋, 中村均 (1986). 単一事例実験法におけるノンパラメトリック統計手法の特殊教育への応用に関する研究. 国立特殊教育総合研究所 研究紀要, 13巻, 1-7, 横須賀.