

心理学研究における実験的方法の意義と 限界(4)単一事例実験法をいかに活用するか

長谷川 芳典

本稿は、単一事例実験法に関するこれまでの研究や解説書等を概括し、いくつかの誤解や適用・解釈の誤りを指摘し、新たな活用の道を探ることを目的としている。大まかな構成は以下の通りである。

- ・なぜ単一事例実験なのか？
- ・単一事例実験法の問題点
- ・単一事例実験法の将来

1. 単一事例実験法の定義

単一事例実験法 (single-case experimental design) とは、1人の被験者または1匹の被験動物を対象に、何らかの実験的介入を行い、その前後の行動変化に基づいて、介入の有効性を確認する研究法のことを言う。但し、被験者が複数であってもごく少数で群間比較ができない場合、もしくは、1つの集団全体に対して実験的介入を行う場合もこれに含めることがある。

Barlow & Hersen. (1984) によれば、心理学や生理学の分野で行われた最初の実験は、単一個体の実験であった。例として、ブローカによる言語中枢を見出した実験、フェヒナーによる精神物理学の手法、エビングハウスによる記憶の実験、パヴロフの条件反射の実験などが挙げられている。しかし、フィッシャーの推測統計学が登場した20世紀前半頃からは、対照群を設定して統計的検定を行う群間比較法が実験心理学研究法の主流となっていた(岩本・川俣, 1990, p.15)。もっとも、単一事例実験法は決して廃れたわけではない。B・F・スキナーおよび行動分析学の学派の研究では、単一事例実験が方法論的に確立し、数々の成果をあげていった。

なお、各種の心理学研究法解説書では、単一事例実験法(もしくは単一事例研究法)の扱いの比重はマチマチであり、全く取り上げられていない場合もある*1。

*1 南風原・市川・下山(2001)の『心理学研究法入門 調査・実験から実践まで』では、本文全体240頁のうち9頁が割り当てられており、比較的詳しく解説されている。いっぽう、高野・岡(2004)の『心理学研究法 一心を見つめる科学のまなざし』では、本文全体346頁の中で実質1頁が割り当てられているのみであり、「この実験の方法論は成熟しているが、心理学全体のなかでは、やや限定された分野で用いられているように思われるので、ここではくわしくは述べない。」としている。また、大山・岩脇・宮埜Z(2005)の『心理学研究法 データ収集・分析から論文作成まで』では、単一事例実験法についての言及は見当たらない。

2. なぜ単一事例実験法なのか？

この問いかけに対しては、2つのタイプの理由がある。

(1) 現実の制約のもとで仕方がない、という理由

第一のタイプの回答は、実験研究は、できるだけ多数のサンプルを集めて無作為割付を行い群間比較を行うことを望ましいとしつつ、現実の制約のもとでやむなく単一事例、もしくは少数事例の実験を行わざるを得ないとする立場である。

長谷川 (1994) が指摘したように、そもそも実験研究というのは、未知の諸要因の影響を無作為な割付によって取り除き、群間の有意差を見出すことにより、当該の独立変数の効果を確認する研究方法のことである*2。単一事例実験は、実験法のロジックの根幹である無作為割付の手續を踏んでおらず、しばしば準実験の延長線上に位置づけられていることに留意する必要がある(南風原, 2001参照)。

ではどのようなケースで、「やむを得ず」単一事例実験を行う必要があるのか？

Barlow & Hersen (1984) が精神科領域における治療法の有効性を検討するにあたって挙げた理由を要約引用すると以下ようになる。

- ・倫理的な問題：治療を行わない対照群を設けて比較するという実験は、対照群への治療放棄であり倫理的に問題がある*3。
- ・実施上の困難：精神科来院者の中で適切なサンプル数を確保することはきわめて困難。例えば、20人の強迫神経症患者のサンプルを集めるには2万人の患者を診なければならないと言われている。

また教育現場において学習指導法の効果を検証するというような場合、ひとつの学校内において複数の指導法を比較することは困難であるし、仮に複数の指導法が導入できても生徒をランダムに割り付けることが難しい点などがあり(南風原, 2001, p.124)、「やむを得ず、単一事例実験を行うことがある。

(2) 集団の平均値比較では解明できない問題がある、という理由

以上(1)では、「現実的制約のもとでは単一事例実験に頼らざるを得ない」という消極的な理由を挙げたが、これに対して、「群間比較よりも単一事例実験法のほうが優れている面があり、積極的に活用すべきである」と考える立場もある。

これに当てはまる理由として、Barlow & Hersen (1984) は、

*2 医療場面では、RCT (ランダム化比較試験、Randomized Controlled Trial) と呼ばれることが多い。

*3 一部のグループへの介入をずらす (=待機させる) という「待機法」により未介入群の設置を回避する方法もある。

- ・ 平均値算出の問題点：精神科領域では、そもそも均質なグループはあり得ないとする立場。
グループの平均値ではなく、個々の変化に注目したほうが適切であるとする。
- ・ 個体差の意味：個体のばらつきこそが臨床には大切なのであるとする立場。

などを挙げている。このうちデータを平均してしまうと個々の特徴が見えにくくなるという指摘は、以下の図1の仮想事例で分かりやすく説明することができる。

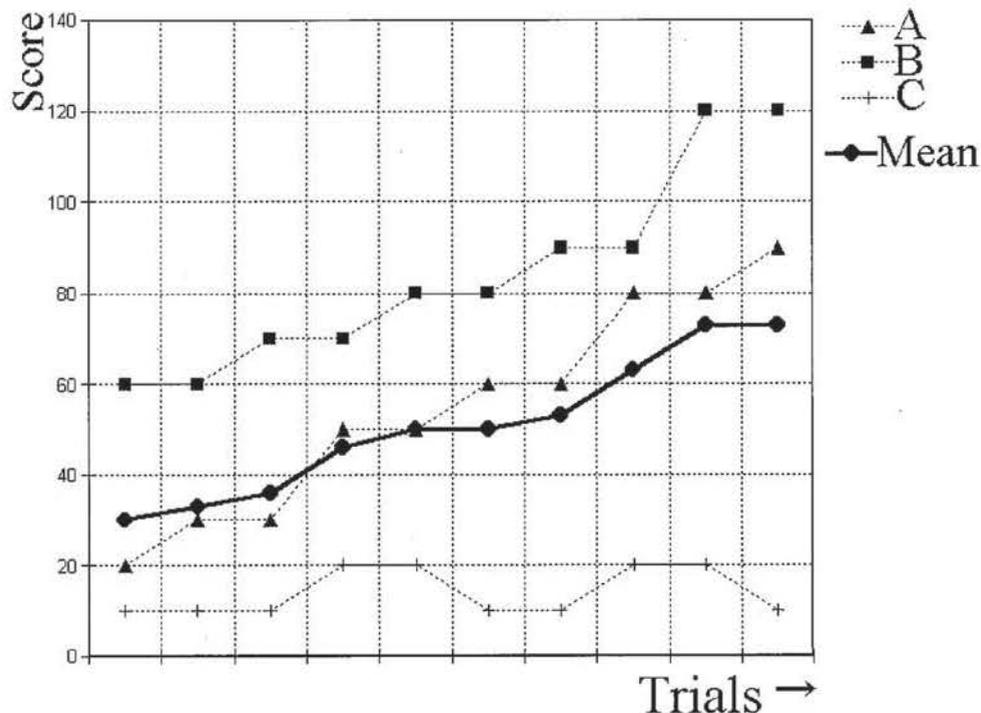


図1 試行 (Trials) の反復に伴う指標値 (Score) の変化についての仮想事例。それぞれの折れ線グラフは、仮定の被験者A,B,Cの個別の変化、および平均値 (Mean) を表す。

図1は、試行 (Trials) の反復に伴う指標値 (Score) の変化をグラフ化したものであり、データはすべて筆者がこしらえた仮想データである。この図では、A、B、Cという3人の被験者の変化が図示されており、太い実線はそれら3者の平均値の変化を表す。

ここで重要な点は、3者の平均値だけでは、

- 試行の反復によって指標値はしだいに増加した。但し、一部、例外者もある。

という程度の結論しか示せない。いっぽう、3者の個別の変化を見ると、被験者AおよびBは、単純に増加傾向を示しているばかりでなく、いずれも「停滞」と「上昇」という階段状の変化を示していることに気づく。また、Cは何ら一貫した増加傾向を示していないが、平均してしまえ

ば単なる例外として扱われてしまう。この場合、増加傾向が統計的に有意であるという実験結果を期待する研究者は、被験者D、E、F... というようにさらに被験者を集めて、「多数派」の平均値の増加をもって一般的法則が実証されたと主張する（反面、被験者Cはノイズとして切り捨てられてしまう）。

いっぽう、以上述べたように、個別の変化を観察した場合には、とりあえず

- 被験者AとBは、いずれも指標値を増加させた。その変化は階段状である。
- 被験者Cでは増減傾向は認められない。その原因は別途説明する必要がある。

という結論を出すことができる。もちろんそれだけでは一般化はできないが、次のステップとして、

- 他の被験者D、E、F、... でも増加するのか
- 増加のパターンはすべて階段状なのか、それとも複数種類のパターンがあるのか
- 被験者Cはどのような条件を揃えれば増加するのか。

というような研究課題が生まれてくる。平均値の増加のみに拘って強引にサンプルを増やそうとする立場に比べれば、個体の変化を重視した後者の立場のほうが遙かに生産的であると言えよう。

以上に述べた仮想の事例およびその基本的な考え方は、じつはすでにスキナーによって1930年代から提唱されている。じっさい、Skinner (1938, p.442) やFerster & Skinner (1957) では、オペラント反応の累積記録が最も重要なデータとして扱われているが、それらはすべて個体ごとの記録であり、そこからFIスキヤロップ**など、強化スケジュールに対応した特徴的な反応パターンが発見された。

O'Donohue & Ferguson (2001, pp.78-80) は著書『スキナーの心理学』の中で、物理学と行動の科学の違いについて次のように指摘している。

- ・物理学の対象は、一般に個別性は存在せず、個別性を無視しても差し障りは出ない。
 - ・いっぽう生活体の場合はそれぞれ独自の存在であり、2人が同一ということはありません。
- 個別性をふまえた上で一般原理が見出される。スキナーは単一事例研究法こそその最も有効な研究法であると主張した。

* 4 定時隔強化スケジュール (FI: fixed interval) において、オペラント行動が強化された後、しばらく行動が休止し、その後、行動を再開すると、時間経過と共に行動のペースが速くなり、次に強化される直前でペースは最も速くなる現象。反応の累積曲線がホタテ貝の形に似ていることから「scallop」と呼ばれる。

なお、以上に述べた単一事例研究の優位性には、じつは、個体の一貫性についての暗黙の仮定があることを指摘しておきたい。以下に述べるように、単一事例実験研究では、反復実験を重視しているが、そこには、実験的方法によって操作する独立変数以外の部分はすべて同じという暗黙の前提がある。この前提がどの程度妥当なものかについては、4.(5)で改めて論じることにしたい。

(3) 現実社会での「単一事例研究」

以上、研究方法としての単一事例実験を取り上げてきたが、現実社会では、実質的に「単一事例」として処置、運用、治療、分析等が行われている例は少なくない。

まずそもそも、個人に対して何らかの治療、セラピー、学習法、健康法などを導入し、その有効性を確認しようとする場合は、必然的に単一事例実験を行わざるをえないという点を指摘しておきたい。

例えば医師が患者に薬を処方する場合、当該の薬の一般的な効能は臨床試験^{*5}により確認されているが、患者個人への効き目の確認は単一事例研究法に頼らざるを得ない。通常、医師はまず、副作用の有無を確認した上で、薬Aを処方する。薬Aが治療に有効であった場合はその処方を継続、いっぽう、必ずしも効果が期待できなかった場合は、量を増やすか、別の薬Bに切り替える。もっとも、外來の場合、医師の指示通りに薬を飲むかどうかは患者本人の自己管理に任せられ、時として、医師に対して虚偽の申告がなされる場合もある。

ある子どもの勉強向上のために家庭教師Xさんを雇用するという場合も単一事例実験となる。雇用後、その子どもが勉強に励み成績が向上するようであれば雇用を継続するが、思うような向上が認められなかった場合は家庭教師Yさんに交代させる。なお、この事例については、4.(2)で改めて論じる。

このほか、ある組織における適材適所の配置、あるプロジェクト実施の有効性、新商品の開発、高速道路料金割引の社会実験などは、最終的には単一事例実験に準じた「試行→本格実施」の反復により導入されていくことが多い。

いっぽう、現実場面での「単一事例型」導入は、しばしば迷信や思い込みの原因ともなる。ある「おまじない」を始めたところ病気が治った人は、おまじないに効果があったと信じて、その後も終生、同じ行為を続けようとする。健康食品の場合も同様であり、医学的根拠の有無にかかわらず、同じ食品を購入し続けようとする。これらは、おそらく、阻止の随伴性^{*6}が働いているものと考えられる。

*5 薬の有効性や安全性を確認するための臨床試験は、いっぽうに「ランダム化比較対照研究」により行われる。

*6 ある行動を継続していると、何らかの嫌悪的事象の出現が阻止できるという行動随伴性。行動を続けている限りは無変化であるため、この随伴性に基づく迷信行動は長期にわたり継続されやすい。

ダイエット商法ではしばしば、利用者の「Before」と「After」の写真が掲載され、その商品の購入によっていかにスリムな身体になったのかが強調されている*7。

3. 単一事例実験法の検証のロジック

(1) よく用いられる方法

代表的な方法としては、反転法と多重ベースライン法がある。

前者は、介入を行わないベースライン期（A）と実験的介入を行う介入期*8（B期）を一定期間ずつ反転させ、介入の効果を検証する方法である。主な方法としては、ABデザイン、ABAデザイン、ABABデザインなどがある。

いっぽう後者には、被験者間多重ベースライン、状況間多重ベースライン、行動間多重ベースラインなどがある（南風原, 2001）。

(2) ABデザイン

ABデザインは、ベースライン測定後に介入を行うだけのシンプルな実験デザインである。この方法では、介入開始後に想定外の出来事が影響を与えていたとしても確認できない。このほか、単純な時間経過による変化、成熟による変化などと区別できない。このことからみれば、ABデザインは、独立変数の効果を検証する手段としては欠陥があると指摘されている。

(3) ABAデザイン

ABAデザインは、上記ABデザインの実施に引き続いて、再びベースライン期を実施する（ここでは、区別のため、最初のベースラインを“A1”、介入後に再度実施するベースラインを“A2”と呼ぶことにする）。それぞれの時期の指標値（但し、指標値は高いほうが望ましいとする）の大小関係を不等号“<”で表したとき、

$$A1 < B > A2$$

となった場合、介入期の増加は単なる時間経過や成熟では説明できない。その分、介入の効果であったという可能性が確からしくなる。

なお、上記の不等号関係が

$$A1 < B < A2$$

となった場合でも、介入が無意味であったとは断定できない。例えば、ピアノの練習時間を指

*7 利用しても効果が無かった事例がどのくらいあるのかが公開されない。また中には、別人の写真を載せる悪徳商法もあると言われている。

*8 Interventionの訳。Treatment（処遇）と呼ばれることもあるが、本稿では、行動分析学で一般的に用いられている「介入」に統一した。介入といっても、本人の意志に反して、何かを強制するわけではない。第三者による系統的なサポート、という意味でとらえていただきたい。

*9 その行動に伴う好子（コウシ、「正の強化子」と同義）のうち、褒める、注目される、ポイントを増やす、といった第三者が提供する好子は「付加的好子」と呼ばれる。いっぽう、行動に自然に随伴する結果（ピアノ演奏であれば、メロディなど）が好子として働く場合、これを行動内在的好子と呼ぶ。

標とし、新しい練習法の効果を調べようとする場合など、B期における上達によってピアノを弾くことに伴う行動内的好子*⁹が随伴するようになり、後のA2期における自発的な練習量を増加させるという可能性も否定できない。

また、指標値が同じ程度の値であることを“=”で表した時に

$$A1 = B > A2$$

というような変化が起こる場合もありうるが、この場合も介入効果の可能性を考慮しておく必要がある。例えば認知症高齢者のように、自然状態では指標値が右下がり到低下する場合などでは、「A1 = B」という結果は、減少を食い止める効果があったと解釈することができる。

(4) ABABデザイン

介入は通常、本人にとって望ましいと考えられる方向で行動の改善がなされることを目的として実施される。望ましい変化を定着させるには、最後は再び、介入を実施し、改善が定着したことを見届ける必要がある。これが、再度の介入(B)を導入するABABデザインである(ここでは、区別のため、最初の介入を“B1”、A2の後に再度実施する介入を“B2”と呼ぶことにする)。但し、B2を半永久的に継続していたのでは自立できない。時折、介入を中止して様子を探り*¹⁰、自然の随伴性のみで改善がはかれるようになった時点では、当該のサポートを打ち切り、新たに設定される次のステップに移行することが望ましい。

(5) 多層ベースライン法

多層ベースラインは、上記のABデザインを、複数の被験者、複数の状況、異なる複数種類の行動などを対象に、介入の時期をずらして実施することを基本としている。それぞれに対応して、被験者間多重ベースライン、状況間多重ベースライン、行動間多重ベースラインなどと呼ばれる。

1人の被験者にABデザインを実施した場合、「A<B」という変化があっただけでは、時間経過や成熟と区別がつかないことについてはすでに述べた。しかし、複数の被験者や状況、行動などに対して時期をずらして実施し、いずれも、ずらした時点から変化が顕著に見られるようになれば、介入の効果を示しているとの確からしさが高くなる。

この方法には、まず、B2挿入による一時的な改善停滞を避けるというメリットがある。また、上述の「A1 < B < A2」や「A1 = B > A2」となるような変化においても、介入実施のズレ時期に応じた変化が観測されれば、介入の効果の確からしさを高めることができる。

なお、臨床研究では、グループ別の介入の時期をずらす待機リスト法による検証が行われることがあるが、グループ編成が無作為でないことからみて、この検証ロジックは単一事例に多層ベースライン法を加味した方法に近いものと見なすことができる。

*10 「プローブ」と呼ばれる。

4. 単一事例実験法の問題点

(1) ABAデザイン、ABABデザインの問題点

ABAデザインの問題点としては、特に、

- ①介入時（B段階）において単一の独立変数を操作しているように見えて、じつは「パッケージ」を導入している可能性。
- ②本当に第二ベースライン（A2）に戻す必要があるのかどうか、についての誤解

という2点を指摘しておくことにしたい。

まず、①に関して、

●シールを貼るという介入は漢字練習量の増加に効果があるか？

という、単一事例実験を考えてみよう*11。

この実験の介入期では、一定量の漢字練習達成ごとに、シールを貼るというご褒美が与えられる。いっぽう、ベースライン期では、教師からの励ましのみで、シールは与えられない。ABAデザインにより、

A1 < B > A2

という実験結果が得られた場合、シールを貼ることは強化効果があったと結論される。

しかし、この場合でも、「シールを貼る」というご褒美には

- a. シールに描かれたキャラクターの種類：別のキャラクターでは別の効果？
- b. シールの形態
- c. シール台紙の形態
- d. 台紙に貼られたシールを第三者に自慢する機会→シールが増えたことを褒められる効果
- e. 強化のタイミング（強化条件、強化スケジュールの違い）
- f. その他

などがセットで含まれており、介入の効果が確認されたとしても真に影響を及ぼしたのがa. からf. のどの部分なのかは確認することができない。その組み合わせは無限にあり、単独で操作できない独立変数も含まれている。

(2) パッケージ型の効果をどう扱うか

*11 南風原（2001）で、ABA、ABABデザインの説明の際にて取り上げられていた仮想事例。

上記の問題をさらに掘り下げて考えるため、2.(3)で取り上げた、

●ある子どもの勉学向上のために家庭教師Xさんを雇用する

という事例を考えてみることにしよう。

ここで、もしこの子どもの勉学行動が、

$A1 < B > A2$

という変化を示したとすると、Xさんを雇用することは効果があると判断される。

しかし、この場合、「Xさん雇用」というパッケージのうち、

- a. Xさんの使っていた教材が有効であった。
- b. Xさんの双方向的な教え方が有効であった。
- c. Xさんの人格的な魅力が有効であった。
- d. Xさんの容貌の魅力が有効であった。
- e. その他

というように、a.~e.のいずれの要因が有効であったのかは判断できない。

ここで「Xさん雇用」という包括的な効果の検証ではなく、Xさんの教え方のどの要素が有効であったのかという検証を行っておけば、Xさんが個人的事情で突然辞めた時にどう対処すればよいのか、慌てなくて済むであろう。上記a.~d.のうちのいずれかであることが事前に検証できていれば、次に別の家庭教師を雇う時に、同じ教材を使ってもらったほうがよいのか、同じタイプの教え方をする人がよいのか、それとも、人格や容貌を重視したほうがよいのか、といった正確な判断ができるようになる。

(3)「途中で止めてみる」ことは必要か

次に4.(1)で②として挙げた、

●本当に第二ベースライン(A2)に戻す必要があるのかどうか、についての誤解

について検討してみよう。

まず上記の家庭教師Xさん雇用の事例であるが、Xさんの長期間雇用が約束されていて、当面、特段の問題が生じていなければ、わざわざ途中で止めてみる(=一時期、Xさんに休んでもらう)必要はない。しかし、家庭教師雇用の本来の目的が「その子どもの勉学習慣向上」にあるとするならば、いつまでもXさんに頼っているわけにはいかない。家庭教師の効果の検証目的、というよりも、子どもの自立を促すために、Xさんの影響力を次第に減らしていくプログラムが求められる。さらに、上記(4)のa.~d.のような要因を検証するためには、教材を

取り替えてみたり、Xさんの教え方を双方向的以外に変えてみたり、Xさんの代わりに一時期、Yさんに来てもらって効果を検証するといった実験的手法をとることになる。こういう必要がある場合に限って、第二ベースラインの導入、もしくはさらに細かい実験計画の実施が必要になると考えられる。

もう1つ、

●花粉症に悩まされている人が、X町からY町に引っ越した。Y町の環境は、花粉症の改善に有効と言えるか？

という別の事例について考えてみることにしよう。

もし、X町からY町へ引っ越した後に花粉症にならず、生活面で何の不都合も無いのであれば、「試しに止めてみる」（この場合は、一時的にX町に戻ってみる）という期間をわざわざ挿入することは、お金もかかるし、無意味であろう。

いっぽう、花粉症にかかりやすいという1点を除けば、Y町よりX町のほうが遙かに住みやすいというのであれば、より多面的に花粉症の原因究明と対策を重視すべきである。単に、X町かY町かという包括的な比較ではなく、X町のどの場所から飛んでくるどういう花粉が影響を及ぼしているのかを分析する必要がある。その上で、X町に住んでいても花粉症にかからない方法が見出せれば、再びX町に戻ればよい。これは、実験計画上の操作というよりも、対象者のニーズに合わせた処置ということになる。

この事例に示されるように、何かの症状の原因が生理学的に同定されている場合は、わざわざ、反転実験計画で確かめる必要はない。有害な原因を取り除くだけで問題は解決する。

同じケースにおいて、花粉症ではなく、原因不明の不眠に陥るような場合、つまり、「X町に住んでいると不眠に悩まされるのでY町に引っ越したらよく眠れるようになった」というような場合は、真の原因の解明を目ざしつつ、X町とY町を比較するための実験計画を導入することにはそれなりの意義が出てくる。

X町からY町に引っ越した後に不眠が解消し、その後一時的にX町に戻ったところ再び不眠になってしまったとする。X町の環境のどういう要素が睡眠を妨げているのかは解明できないにしても、とにかく、X町に住み続けることは不眠の原因となっている可能性がきわめて高い。よってY町に永住する方向で住居を確保することは合理的な選択ということになる。

いっぽう、X町に戻ってももはや不眠にならなかったとすれば、不眠の原因はX町の環境以外の別のところ（本人自身の健康状態など）にあった可能性が高まる。よってそのまま、X町に住み続ければよい。

(4) 多層ベースライン法の問題点

3. (5) に述べたように、多層ベースラインは、ABデザインを、複数の被験者、複数の状

況、異なる複数種類の行動などを対象に、介入の時期をずらして実施することを基本としている。それぞれに対応して、被験者間多重ベースライン、状況間多重ベースライン、行動間多重ベースラインなどと呼ばれる。

このうち、被験者間多重ベースラインは、介入の効果の一般性、つまり、たった一人の被験者だけでなく、よく似た状況の多くの被験者に対しても効果があるという点を示す上では有力である。なお、現実場面では、グループ別の介入の時期をずらす待機リスト法による検証が行われることがある。

もっとも、被験者間多重ベースラインで確認できるのは、対象者に及ぼす効果がどの程度共通しているかという面に過ぎない。薬の効果に関して言えば、ある薬の服用を開始するとその時点から同じような改善が見られるという確認はできる。しかし、体質の異なる対象者に対して異なる作用をするような場合、効き目が出なかった対象者が過半数を占めたからといって、効き目のあった対象者への処方偽薬効果に過ぎなかったと結論づけるわけにはいかない。

いっぽう、南風原(2001, p.148)によれば、状況間多重ベースラインや行動間多重ベースラインは、「複数の状況下」および「複数の標的行動」について継続的に観察を行い、異なる時期に介入を開始してその効果を検討するという研究法であり、具体例として、

対人不安の傾向のある子どもについて、幼稚園と家庭という2つの異なる状況において調べたり、「名前をよばれたときの返事」と「集団遊びへの参加」という2つの標的行動について調べたりすることが考えられる。

と説明されている。しかし、「幼稚園と家庭」という条件間の検討は、単に、ある場所で介入を行ったことがどこまで般化したかという検討であるとも考えられる。また、例えば、シールを貼るというご褒美が「名前をよばれたときの返事」と「集団遊びへの参加」の両方の行動を強化できたとしても、それによって分かることは、その子どもにとってシール貼りは強化子として有効であった、というだけであり、それ以上に有用な情報が得られるかどうかは心もとない。

(5) 個体の同一性に関わる問題点

単一事例実験研究では、反復実験を重視しているが、そこには、実験的方法によって操作する独立変数以外の部分はすべて同じという暗黙の前提がある。この前提はどの程度妥当なものであろうか。

まず、遂行実験と習得実験について考えてみよう。

遂行実験というのは、異なる強化スケジュールのもとで、反応速度や反応パターンがどのように異なるのかを実験的に明らかにしようとするような実験のことを言う。例えば、平均30秒の間隔で強化されるような定時隔スケジュール(FI30sec)と変時隔スケジュール(VI30sec)における行動を比較したとする。1つの条件を5分間として、2つのスケジュールを
FI30sec→VI30sec→FI30sec→VI30sec→FI30sec→VI30sec

というように交替させた場合、それぞれのスケジュールの下での行動パターンは、どのセッションでも大差ないような遂行が観測されるだろう。この種の実験では、個体は常に同一状態であり、強化スケジュールの違いのみが遂行を変化させたと判断することが可能である。

これに対して、ピアノの練習を30分行うごとにシールを貼るという強化の効果を確認するという実験を行った場合は、実験が進めば進むほど、ピアノの演奏技術が上達していくと想定される。すでに述べたように、

$A1 \rightarrow B1 \rightarrow A2 \rightarrow B2$

というように、ベースラインと介入条件を交替させることは形式上は可能であるが、第一ベースライン期に比べて第二ベースライン期では演奏力が格段に上達しており、シール以外の好子、例えば、演奏によって流れるメロディそれ自身が自然の好子として演奏行動を強化するようになっている可能性がある。

外国語学習、計算力、漢字学習というように、習得を目的とした介入においても、多かれ少なかれ、個体側に右上がりの変化が生じていると考えてよいだろう。従って、ABA実験デザインで

●ある学習方法は成績向上に有効か？

を検討した場合には、

- ・ $A1 < B = A2$ の場合：その学習方法は成績を向上させた可能性あり。
- ・ $A1 < B2 < A2$ の場合：その学習方法は成績を向上させた可能性もあるが、単に、(方法によらない)学習の積み重ねで成績が向上した可能性も否定できない。
- ・ $A1 << B2 < A2$: (方法によらない)学習の積み重ねで成績が向上した可能性もあるが、その学習方法がそれを加速させる効果があった可能性あり。
- ・ $A1 > B2 = A2$: その学習方法は、学習に悪影響を及ぼし成績をかえって悪化させた。

という結論を出せるのが精一杯であろう(第一ベースライン、介入期、第二ベースラインのA1、B、A2を何らかの成績指標、変化の大小関係を不等号(<、>、極端な変化を<<、>>で表す)。

なお上記において、変化の大きさを比較できるのは、成績指標が実質的な内容を伴った間隔尺度であることを前提としている。

このほか、認知症高齢者を対象とするような場合は、もし何も介入しなければ時間経過とともに認知・行動指標が右下がりに低下していく可能性もある。その場合、介入により指標値が上昇しなくても、指標値の低下を食い止める効果があった、というように解釈できる可能性がある。

5. 単一事例実験研究の将来

(1) エビデンスに基づいた研究、実践

最近、心理学の一部の領域で「Evidence-Based (エビデンスに基づいた)」という考え方が注目されており、いくつかの関連学会年次大会や心理学関連シンポジウム等で、テーマとして取り上げられている*12。

2007年3月から9月までの6カ月間に限っても

●心理療法におけるエビデンスとナラティブ：招待講演とシンポジウム*13

(2007年3月21日、立命館大学衣笠キャンパス)

●エビデンスに基づいた発達障害支援の最先端*14

(2007年8月5日、日本行動分析学会第25回大会・学会企画シンポジウム、立教大学新座キャンパス)

●認知行動療法と実証(エビデンス)にもとづく臨床：クライアントにとって真に有効な実践は何か?*15

(2007年8月20日、日本心理学会第71回大会、東洋大学白山キャンパス)

というように、心理療法や発達障害児支援と関連づけた企画が相次いで開催された*16。

本稿の目的との関係で特に留意しなければならないのは、単一事例実験法が、エビデンス重視の流れの中でどう位置づけられているのかという点である。

Dryden & Rentoul (1991) に示された効果研究のステップでは、単一事例実験は、将来の研究への見通しを与えてくれる可能性があるが、結果の解釈が曖昧になるという限界があると指摘されている。

直近の日本心理学会第71回大会シンポジウム*17でも言及されたことであるが、エビデンスに基づいた療法が重視される1つの背景には、

*12 エビデンス重視の歴史的な流れについては、中野(2004)に詳しく解説されている。

*13 導入講演：サトウタツヤ

特別招待講演：John McLeod

シンポジウム話題提供：下山晴彦/能智正博氏/武藤崇/松見淳子

*14 企画・司会 浅野俊夫

話題提供 山本淳一/平澤紀子/井上雅彦

*15 企画：丹野義彦/石垣 琢磨

司会：丹野義彦

話題提供：市井雅哉/伊藤絵美/井上雅彦/津富宏

指定討論：石垣琢磨

*16 筆者(長谷川)自身の参加報告・見解を

<http://www.okayama-u.ac.jp/user/hasep/journal/psy-rec/index.html> に公開している。

*17 脚注15参照。

- エビデンスのある所にお金を出すことで、限られた資源を有効に活用する

という考え方がある。これは、

- 同じコストであるならば、効果検証実験において、効果量の平均値が高く、かつバラツキの小さい介入を採用したほうが効率的

という考え方にも繋がる。確かに、保健医療、公的な支援を考える上ではこうした視点は重要であるし、また、ユーザー（クライアント）自身が種々の療法のうちからどれを選ぶかという選択の手がかり、第三者にその療法を受け入れさせるための説得手段として有効であると言うことはできる。しかしそれらをもって、どの個人にも同一の療法を画一的に当てはめてよいということにはならない。

要因統制実験で検証されるエビデンスとはあくまで、万能の効果、もしくは、一定基準に基づいた集団の構成員全員に期待される効果についてのエビデンスであって、決して、個体本位のエビデンスではない。例えば、仲間と一緒に合唱をすることが認知症の進行を食い止める上で有効であるということが要因統制実験で検証されたからといって、認知症高齢者全員にとってそれが有効であるとは限らない*18。

心理療法におけるエビデンス重視という観点は、ともすれば、各種療法間の効果検証比較にとどまってしまう。しかし、実際に行われている各種療法の間には類似性もあり、明確な基準により定義されているわけでもない。むしろ大切なことは、対象者それぞれの特性に応じて、それぞれの療法のパーツの中から何を取り出し、組み合わせて実施していくかという、個人本位の有効性の検証であろう*19。

(2) 行動変容の実績ではなく、自然随伴性への移行こそ、最大の課題

最後に以下のようなセルフコントロールの事例を考えてみよう。

*18 音楽に興味のない対象者まで無理矢理借り出して歌わせたとしても「施設にはお世話になっているのでしようがないけれど、こんなことまでさせられるんですわ。」という義務感が生じるのみである。なお全人ケアの考え方については、長谷川（2004）を参照されたい。

*19 井上雅彦氏（兵庫教育大学）は、ご自身のブログ（2007.8.9付）「求むエビデンス脱藩同盟」（<http://aba.jugem.jp/?day=20070809>）の中で、日本行動分析学会第25会年次大会シンポジウムの話題提供において、「最先端の科学性を持った研究としてのエビデンスの追求と同時に、現時点で届けられるサービスの中での最良のエビデンスの構築の両者を進めていく必要がある」という観点から以下の2点を提議した、と記しておられる。

1. 効果研究の方法について知ること、自らの実践研究の効果研究としての質を厳しく自己評価すること、効果研究のレベルを段階的に引き上げる方向性を持つこと、RCT（無作為化比較試験）への研究ステップを作ること
2. またマニュアル化、地域資源への技術提供とシステム化、適切な評価の方法の提供という形で既存の地域システムをABAの客観的評価技術によってエンパワメントするという視点を導入すること

このうちの2については“我々自身が「療法」や「技法」にこだわることをやめ、他の支援方法を総合した現時点で我が国の地域で利用可能なベスト・プラクティスを提案するといったレベルでのエビデンスを作っていくということを意味しています。”としており、個体本位のエビデンス追究の必要性を説いておられる。

●体質改善のため、毎朝30分の散歩をすることに決めた。それを強化するために、「散歩をちゃんとした時に限って(その人の大好きな)アイスクリームを1個食べるという自己契約をした(散歩しなかった時はアイスクリームは決して食べない)。このセルフコントロールは有効と言えるか。

この効果をABAB実験計画で確認するためには、

A1: 第1ベースライン期。アイスクリームを食べるという自己契約をしない段階では、毎朝の散歩はちっとも長続きしない、という安定状態にあることを確認する。

B1: 実験期間。「散歩をする→アイスクリームを食べる」という形で、散歩行動を強化する。

A2: 第2ベースライン期。散歩をしてもアイスクリームを食べない、という条件のもとで散歩行動が維持できなくなることを確認する。

B2: 第2実験期間。再び「散歩をする→アイスクリームを食べる」という形で、散歩行動を維持する。

という実験計画により検討され、A1、B1、A2、B2という4期における散歩行動の出現頻度(達成率など)が、

$A1 < B1$ 、 $B1 > A2$ 、 $A2 < B2$ (←「<」あるいは「>」は、達成率などの大きさを比較する不等号)

という不等号関係になるようなデータが得られれば、「アイスクリームを食べる」ことは、散歩行動の強化にとって有効であったことが検証された、と結論できるであろう。

この実験は、ABAB実験デザインを忠実に実施しており、個人内において、エビデンスのあるセルフコントロールということになるが、はたして、その結論にどれだけ意味があるだろうか。この形で検証されたのは「アイスクリームが好子として有効に機能した」というだけであって、裏を返せば、

●この人にとっての散歩行動は、アイスクリームに頼らなければ、維持できない(つまり「アイスクリーム依存症」)。

を示したにすぎない。この人が、昼休みのウォーキングを始めようとした場合に、アイスクリームが同じように好子として使えるであろうという知識は増えるかもしれないが、それだけのことである。しかも、この散歩行動が体脂肪や悪玉コレステロール値減少を目的として始められたとしたら、せっかく散歩が習慣化しても、アイスクリームに含まれる乳脂肪分や糖分の過剰摂取によって、逆に、体質を悪化させてしまう恐れさえある。

セルフコントロールの最終目的が「散歩行動の習慣化、それによって結果的に達成される体質改善」というところにあるとするならば、重要なのは、「アイスクリームを食べる」はあくま

で暫定的、補助的、付加的な好子であると位置づけ、散歩行動が文字通り「自走」できるように、散歩行動に自然に伴うような強化子の種類や頻度をできるだけ増やす工夫をすることにある。

例えば小型のデジタルカメラを持参して、散歩途中で見つけた珍しい花、キノコ、野鳥、ネコ、昆虫などの写真を撮ることはそれ自体きわめて強化的となる。そうなるともはや、「アイスクリームを食べる」といった補助的、付加的な強化は要らない。特段の助けを借りなくても、散歩行動は自然に伴う好子によって継続できるようになるはずである。

行動が自然随伴性によって強化されるようになった場合、これまで同様にベースラインをA1、A2、介入条件をB1、B2という指標値で表すと、

A1 < B1、B1 = A2

となることが予想される。そしてその場合は、B2の導入はもはや不要と言ってよいだろう。

「自然」随伴性による強化と言っても、何も手を加えず、成り行きに任せればよいというものではない。B1期間においては、暫定的、補助的、付加的な好子に頼りつつ、常に、自然に伴う好子の種類と頻度をできるだけ増やすために精一杯工夫をすることが肝要である。例えば、当初設定した散歩コースが、変わりばえのしない家並みで、しかも車の騒音が嫌悪的になるというのであれば、コースを変更することを第一に考えるべきであろう。また、自然風景は必ずしも強化的でないという人の場合には、別途、自然観察の入門書などを読んで関心を高める工夫（自然事物が習得性好子になるような工夫）をするか、もしくは、全く別の強化子、例えば、散歩道で出会う人々との会話が楽しみになるような工夫をしていく必要がある。

以上述べてきたように、単一事例実験法は、

- a. 一般法則を検証する前段階としての発見的方法
- b. 多数において効果が検証されている療法や介入法を一個人に適用するにあたっての効果検証方法

という、異なる2つの手段として活用される。このうちb.の場合は、療法や介入によって効果が確認されたとしてもそれは過渡的な効果として位置づけるべきであり、改善された行動が自然の随伴性のもとで如何に「自走」できるかというフォローアップの段階まで責任を持って実施することが肝要であると考えられる。

引用文献

※外国語文献のうち出典を原文表記したものは、※印を除き、原書版を参照した上で引用しているが、読者の便宜をはかるため、翻訳書が刊行されている場合にはカギ括弧内にそれを付記した。

- Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs—Strategies for studying behavior change—Second Edition*. Pergamon Books. [バーロー・ハーセン著、高木俊一郎・佐久間徹(監訳)『一事例の実験デザイン 新装版—ケーススタディの基本と応用—』二瓶社、1993年]
- ※Dryden, W. & Rentoul, R. (Eds.) (1991). *Adult clinical problems: A Cognitive-behavioural approach*. Chapman & Hall. [ドライデン・レントウル(編著)丹野義彦(監訳)『認知臨床心理学入門—認知行動アプローチの実践的理解のために』東京大学出版会、1996年]
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
- 南風原朝和(2001). 第5章 準実験と単一事例実験. (南風原朝和・市川伸一・下山晴彦『心理学研究法入門 調査・実験から実践まで』東京大学出版会. pp.133-152.)
- 長谷川芳典(1994). 心理学研究法再考:(1)基礎的統計解析の誤用をなくすための30のチェック項目. 岡山大学文学部紀要, 21, 47-59.
- 長谷川芳典(2004). 第II部 ダイバーショナルセラピーの基礎理論 第1章 高齢者の理解 第1節 心と行動 [日本ダイバーショナルセラピー協会編『全人ケアの実践』朱鷺書房, pp58-65]
- 岩本隆茂・川俣甲子夫(1990). シングルケース研究法. 新しい実験計画法とその応用. 勤草書房.
- 中野良顕(2004). 行動倫理学の確立に向けて～EST時代の行動分析の倫理～. 行動分析学研究, 19, 18-51.
- O'Donohue, W. & Ferguson, K. E. (2001). *The psychology of B.F. Skinner*. Sage Publications, Inc. [オドノヒュー・ファーガソン(著)、佐久間徹(監訳)スキナーの心理学. 応用行動分析学(ABA)の誕生. 二瓶社. 2005]
- 大山正・岩脇三良・宮埜静夫(2005). 心理学研究法 データ収集・分析から論文作成まで. サイエンス社.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Prentice-Hall.
- 高野陽太郎・岡隆(2004). 心理学研究法 心を見つめる科学のまなざし. 有斐閣アルマ.