

キャリア講演会への参加と意識変化

——回顧的プレテストを組み込んだ変化を把握するアプローチ——

A Pretest/Retrospective Pretest-Posttest Investigation of a One-Time Career Development Lecture Meeting

浦上昌則

Masanori URAKAMI

要 約

本研究は、対象者の変化を把握する方法として、回顧的プレテスト—ポストテストという調査デザインの概略を示すとともに、大学で開催された1度のキャリア関連講演会をとりあげ、その講演会への参加前後の意識変化について検討した。分析対象者は心理・教育系の学科の大学新入生96名であり、プレテスト—ポストテスト—回顧的プレテスト—フォローアップの3時点4種の測定を実施した。その結果、講演会企画趣旨に関する5つの意識は概ね高まったことが確認できたが、従来のプレテスト—ポストテストよりも、回顧的プレテスト—ポストテストの方がその変化が明確であった。ところが、変化した意識は持続せず、約1ヶ月後には、講演会前の状況とほぼ同じ様相となることも確認された。こういった結果を踏まえて、キャリア研究に回顧的プレテストという方法を導入する意義、またレスポンスシフト概念を導入する意義について議論した。

問題と目的

近年では、大学におけるキャリア教育も広く定着し、その結果、成果（アウトカム）を報告する研究も多く見られるようになった。こういった介入による変化を検討しようとする研究、特に量的データを指標とした研究では、その多くで1群のプレテスト—ポストテストという準実験デザインが採用されている。では、このデザインがキャリア教育における介入の効果／対象者の変化を把握する際に適した方法なのだろうか。海外の動向に目を向けると、介入の検討（プログラム評価）に回顧的なプレテストを用いている研究がいくつも存在する。そこで本研究では、この回顧的な方法の概略を示すとともに、大学で開催された1度のキャリア関連講演会をとりあげ、その講演会への参加前後の意識変化について検討することを目的とする。

まず、介入の効果検討を目的として回顧的な方法を用いている一例として、Vo et al. (2021) を

とりあげる。かれらは、キャリアにおける偶然の出来事への対応は、医学生の特化的な能力開発においても重要であるものの、ほとんど議論されることがないと指摘する。そして、ハプンスタンス学習理論 (happenstance learning theory; Krumboltz, 2009 など) に着目し、これが医学教育の文脈で言及されていないというギャップを解消するために、90 分間の対話型ワークショップを作成した。このワークショップは、教員インタビューや学生の証言動画、ハプンスタンス学習理論の解説やディスカッションなどから構成されている。

このワークショップの効果検討の際、回顧的プレテスト (retrospective pretest) とポストテストの得点差からの検討も行われている。回顧的プレテストは、ワークショップ後にポストテストと同時にされるものであり、ワークショップ前を思い出して (回顧して)、その時点の様子について回答を求めるものである。用いられた設問は次に示す 3 つであり、ワークショップ前を思い出しての回答 (回顧的プレテスト) と、ワークショップ後の回答 (ポストテスト) をそれぞれで 5 段階のリッカート法 (1 = strongly disagree, 5 = strongly agree) で評定させている。

その結果、「Indecision is not a problem to be fixed, but a playful, open-mindedness to future possibilities. (優柔不断は解決すべき問題ではなく、将来の可能性に対して計画的でオープンマインドな姿勢である)」は、回顧的プレテストの平均は 3.5、ポストテストでは 4.6 であった。「It is normal and desirable for unplanned events to influence interests, attitudes, and preferences. (想定外の出来事が興味や態度、嗜好に影響を与えるのは通常のことであり、望ましいことである)」と「It is worth the effort to promote and take advantage of happenstance situations. (偶然を促進し利用することは、努力する価値がある)」についてはいずれも同順で、4.1 と 4.7 であった (いずれも $n = 39$)。またこれらの差は統計的に有意であった。

この結果などから、ワークショップによって参加者はハプンスタンス学習理論にそった態度の重要性を認識するようになったと指摘するのである。ところが、このような知見に対しては疑義を呈する者も多いであろう。その根拠は、回顧的プレテストという研究デザイン、もしくは測定方法に対する疑念からと考えられる。

プレテストとポストテストの差から変化を把握できるということは、得点差が介入によって生じた違いを正確に、もしくは近似的に反映していると仮定していなければ成立しない (たとえば Geldhof et al., 2018)。ワークショップ後に回顧的プレテストを行うと、その得点差には、ワークショップに参加したことなど、評価者が望まない何からの影響が混入してしまい、回顧的プレテストとポストテストの差はワークショップによって生じた違いを正確に反映していない可能性が考えられる。それゆえ Vo et al. (2021) のような研究デザインによる結果では、ワークショップの効果を正しく評価できないのではないかと疑念が生じるのであろう。

この問題を回避しようとすれば、ワークショップの前後にテスト (プレテストーポストテスト) を行うという代案が提示されよう。現在、効果、変化を検討しようとする多くの研究で、このデザインが用いられている。

では、プレテストーポストテストというデザインを用いれば、その差は介入によって生じた違いを、回顧的プレテストーポストテストよりも正確に反映すると判断してよいのであろうか。我が国では、この点についての検討、議論が少ないように思われる。そしてプレテストーポストテストというデザインを用いれば、その差から変化を把握できるということについて、疑念を差し挟むことなく共有されているといえるだろう。

ところが海外の動向に目を向けると、たとえば Vo et al. (2021) のように、回顧的プレテストー

ポストテストというデザインへの注目、採用という動向が窺われる。これらは、主にプレテスト→ポストテストというデザインでは、その差から変化を正しく把握できない（しばしば、過小評価する）という知見や、それへの対応として回顧的プレテスト→ポストテストというデザインが提案されたことを背景としている。

この問題への注目の嚆矢とされる研究は、1980年前後に発表されたHowardらによるいくつかのものといえよう。そのひとつであるHoward, Ralph et al. (1979)は5つの研究から構成され、第1研究の空軍で実施した、独断性 (dogmatism) の減少を目的としたコミュニケーションスキル研修プログラムにおいて興味深い現象が観察されている。そこではプレテスト→ポストテストのデザインが採用され、プログラム参加の前後で自己報告式の独断性傾向を測定する尺度（得点が高いほど独断的）が実施された。プレテストとポストテストの得点差から、参加者の62% (434名)で独断性が高く、36% (253名)で低くなり、2% (17名)で変化がないことが確認された。平均値はプレテストで156.72、ポストテストで162.25であり、その差は統計的に有意であった。すなわち、研修プログラムによって、参加者はより独断的に変化したことが確認されたのである。

これは明らかに予測、仮説と矛盾する結果であるため、Howard, Ralph et al. は参加者に追加のインタビューを行った。そして、参加者はワークショップを通して、参加前の独断性に対する認識を変化させたことを見いだしている。インタビューにおける、特定の項目に対する回答の例として、以下が記されている (Howard, Ralph et al., 1979: 著者訳)。

項目：議論が白熱すると、自分の言いたいことに熱中してしまい、他の人の話を聞くのを忘れてしまう。

プレテスト時：私は人と話しているとき、他の人の言っていることをよく聞く。私は「-2 (全体としては同意できない)」だと思う。

ポストテスト時：グループでの経験のおかげで、人の話を聞いていないことに気がついた。最初に記入したときに「+3 (とても同意する)」と回答しておけばよかった。でも、グループのおかげで目が覚め、もっと積極的に人の話を聞こうと思うようになった。時々人は人の話を聞くのを忘れてしまうことがあるけれども、全体的にはそれほどひどい状態ではない。私は「-1 (少し同意できない)」をつける。

こういったインタビュー結果を踏まえ、Howard, Ralph et al. (1979)は、対象者が自分のレベルを判断する基準 (basis) を変化させることをレスポンスシフト (response-shift) とよんだ (Howard & Dailey (1979)では、プレテストと、回顧的プレテストにおける評価の違いと表現されている)。プレテスト→ポストテスト間での比較を行う場合、それらの判断の基準が一定していることが前提とされるが、レスポンスシフトというそれが崩れる現象が生じていたのである。Howard, Ralph et al. (1979)は、レスポンスシフトを回避するため、すなわちプレテストとポストテストでの自己報告を同じ内部基準によって測定する方法を模索し、Campbell & Stanley (1963)などの言及から、通常のプレテストに代えて、ポストテスト時に回顧的プレテストを用いることでレスポンスシフトの影響が排除されるだろうという仮説を立てている。そして、通常のプレテストと回顧的なプレテストを用いた第2研究以後の4つの研究を実施し、さまざまな重要な結果を得ている。主要な知見として、通常のプレテスト→ポストテストでは一貫して処遇効果が小さいこと、回顧的評価の回答は1年間という長期にわたって安定性していること、客観的な変化との関係は、自己報告によるプレテスト→ポストテストよりも回顧的プレテストを用いた場合の変化の方が強いことなどである。

そして、これらの相違は、回顧的プレテストとプレテストの違いに起因すると述べている。

このように、自己報告式のプレテスト→ポストテストというデザインを用いても、その差から変化を正しく把握できるとわけではないことが明らかになっている。レスポンスシフトは、その妥当性を脅かす代表的なものとして言及されるが (Hill & Betz, 2005 など)、Lam & Bengo (2003) はその他にもプレテスト効果などのバイアスの存在を指摘している。もちろん、どのような場合でもレスポンスシフトが観測されているわけではないが (Bhanji, et al. 2012; Hill & Betz, 2005 など)、レスポンスシフトによるバイアスが生じるならば、またプレテスト効果が生じるならば、回顧的プレテスト→ポストテストというデザインを採用することで、これらの問題を回避できると考えられる。

Howardらによる一連の研究以降、回顧的プレテスト→ポストテストを用いた研究が積み重ねられてきた。Hill (2020) は、Google Scholarを用いた「retrospective pretest and program evaluation」での検索の結果として、1991年から2000年までで41件、2001年から2010年までで326件、2011年から現在に至るまでで492件を確認できたと指摘する。検討されている領域は多岐にわたっており、Howard, Ralph et al. (1979) だけでも、コミュニケーションに関するワークショップ (第1, 2研究)、女性を対象としたアンドロジニーを促進するワークショップ (第3, 4研究)、大学の授業「援助面接におけるコミュニケーション」(第5研究) で検討されている。近年でも、大学等の授業 (Bhanji et al., 2012; Lane & Adhikari, 2019)、子育て教育プログラム (Geldhof et al., 2018)、教員など指導者に対するワークショップ (Chasteen & Chattergoon, 2019; Nielsen, 2011) など、さまざまな領域のプログラムで検討が行われている。

Hill (2020) は近年の研究の増加を、レスポンスシフトへの懸念の高まりに対応した回顧的プレテストへの注目の結果と指摘するが、同時に回顧的プレテストを支持する研究、研究者には、その場合に特に起こりやすい自己報告バイアスを軽視する、もしくは考慮していないことが多いとも指摘している。このHillに限らず多くの研究者が、回顧的プレテスト→ポストテストというデザインが抱える問題点についても理解しており、実施および解釈における留意点を指摘している (Geldhof et al., 2018; Hill & Betz, 2005; Nimon, 2014; Nimon & Allen, 2007; Pratt et al., 2000 など)。

このように、海外では着実に研究が蓄積されつつあるが、我が国の現状 (日本語での研究発表) に目を移すと、まったく状況が異なっているといえるだろう。論文検索サイトで「回顧的事前テスト」「回顧的プレテスト」といったワードで検索しても、岡田他 (2020) のものが該当する程度であり、ほとんど論文を見つけることができない。プログラム評価に関する多くの準実験デザインを紹介している安田・渡辺 (2008) でも、回顧的プレテスト→ポストテストは紹介されていない。他方で安田 (2011) によるプログラム評価の書籍はこれを取り上げており、岡田他 (2020) はこの安田 (2011) を参照している。

ただし、医療関係、特に Quality of Life (QOL) 研究の分野では「thentest/then test」という呼称で、回顧的プレテストと同様な評価法を用いた研究が我が国でも行われている (鈴嶋, 2015; 渡邊・隆島, 2014, 美崎他, 2013 など)。これらの研究は、客観的な病状は進行しているのに、主観的 QOL は保たれるといった現象に対して、Sprangers & Schwartz (1999) がレスポンスシフトの概念を適用したところに触発された研究群といえるだろう。レスポンスシフトとして患者の内部で何が起きているのかを探究しようとしており、キャリア支援の観点からも興味深い動向といえよう。

その他、「回顧的」で検索を行っても、面接調査など質的研究が多く認められるのに比して、量的な研究はほとんど該当しない。回顧的回答に関する小島・篠原 (2011) などが散見される程度である。

以上のように、我が国では、一部の領域をのぞき、介入の効果を検討する際には、プレテスト→

ポストテストを用いるという暗黙の了解があるように思われる。そしてこれは、キャリア研究の領域でも同様といえるだろう。しかしその了解のためか、プレテストーポストテストの差が本当の変化を示しているのか、レスポンスシフトは生じないのか、回顧的プレテストーポストテストというデザインは避けるべきなのかといったことに関する知見は蓄えられていない。多くのキャリア研究が最終的には人々に対して還元されるものだとすれば、キャリア教育、キャリア支援の効果を検討する方法に関する研究は一層注目されるべきといえよう。

そこで以下では、大学で開催された1度のキャリア関連講演会をとりあげ、その講演会の影響について、プレテストーポストテストー回顧的プレテストーフォローアップの3時点4種の測定を行いデータを収集する。そして、その分析から、講演会の影響と、回顧的プレテストという方法やレスポンスシフト概念を検討に導入する意義を検討する。

方法

講演会と調査対象、調査内容

講演会 2022年6月後半に行われた心理・教育系学科主催の、当該学科卒業生による講演会（対面形式、100分）を対象とする。この講演会は、主たる対象を1年生と想定し、「卒業後の進路を見据えながら有意義な大学での学びができるよう、学部・学科の特性を活かしながら学生が将来の夢や目標を持ち、社会で働き自立することの意義を考えるきっかけを提供する」ことを目的として企画された。登壇者は女性2名であり、事前に企画の趣旨は伝えてあったが、内容については特段の指示、指定は行われなかった。一般的な依頼の形式といえるであろう。話された内容は、学生時代の意識や生活、現在の職業を含む生活、およびその関連にかかわることが中心であった。また質疑応答の時間も含まれていた。なお、この講演会は学年を問わず、また全学の学生にもひらかれたものであり、オンラインで配信された。

対象者 この講演会は、主催学科の1年生を対象とした必修科目とも関連しており、受講生は参加が必須とされていた。本論では、当該学科の1年生参加者のみを調査、分析対象とする。

調査内容 質問項目は、上述の企画趣旨より5点（卒業後の進路を見据える、有意義な大学での学びができる、学部・学科の特性を活かす、将来の夢や目標を持つ、社会で働き自立することの意義を考える）を抽出し、それぞれを項目化した（Table 1参照）。なお、企画趣旨において「きっかけを提供する」と示されていることから、それぞれの点に到達目標を設定し、そこへの接近の程度を指標化するのではなく、そういった内容を生活の中でどの適度意識するかという側面について測定を行うこととした。

調査計画

調査は以下の3回の時点で行った。なお、いずれの回でも、他の調査内容とともに実施された。

1回目 講演会実施の4日前に、web調査の形式で実施した（以下プレテスト）。「あなたは、以下のようなことを日ごろからどの程度意識していますか」という教示を用い、「まったく意識しない」の1から「常に意識する」の5までの5段階で回答を求めた。なお選択肢は「まったく意識しない」、「常に意識する」の両極のみ示した。

2回目 講演会直後に、web調査の形式で実施した。この際には、講演後の今の様相と、講演を

聞く前の様相を回顧して回答を求めた。それぞれの教示は、「講演を聞いた今現在の気持ちを答えてください。あなたは、以下のようなことを日ごろからどの程度意識したいと思いますか」と、「今回の講演を聞く前の自分を思い返してください。その時点なら、『あなたは、以下のようなことを日ごろからどの程度意識したいと思いますか』と問われたらどう回答していたでしょうか」であった。回答方法は1回目と同様であった。以下、前者をポストテスト、後者を回想的プレテストとよぶ。なお実施順は、ポストテストを先、回顧的プレテストを後とした。

3回目 講演会から約1ヶ月後に、web調査の形式で実施した。教示や回答方法は1回目と同様であった（以下フォローアップテスト）。

倫理的配慮

以上の調査は当該学科の1年生を対象とした必修科目において実施された。毎回の調査時には、個人を特定して分析をするものでも、成績に反映されることもないことを説明した。なお、授業改善のための調査でもあったため、研究のためのデータとすることの可否についての意思表示を求め、本研究では同意が得られた対象者のみのデータを分析に用いている。また3回の調査間での回答者の同定は、回答者のみが知る6桁の数値の記入を求めることによって行った。

結果と考察

分析には、R (4.2.0) およびパッケージ `exactRankTests` (0.8-35), `psych` (2.2.5) を用いた。

3回の調査すべてで回答があり、また研究への同意が得られた対象者は96名（女性78名、男性18名、年齢は18から20歳（不明9名））であった。

4つの測定それぞれで、「まったく意識しない」を1、「常に意識する」を5とする1から5を付与し、5つの項目の分布の様相を確認した。平均値および標準偏差はTable 1に示す。

次に、4つの測定値間での差異について検討した。いずれの項目でも、「まったく意識しない」という回答は極めて少なく、高い得点の方へやや偏った分布の様相が確認できた。またShapiro-Wilk検定を行ったところ、すべての項目で正規分布の仮定が棄却された。そこで今回はWilcoxonの符号順位和検定を用い、項目ごとに測定値間について検定を行った。なおBonferroni法による補正を行い、全体の有意水準 α を0.05とし、それを6で割った $\alpha = 0.0083$ を基準としてそれぞれの有意性を判断した。さらに効果量として r とその95%の信頼区間を求めた。 r は、0.1程度が小さい効果量、0.3程度が中程度の効果量、0.5程度が大きな効果量と解釈される（水本・竹内、2008）。

その結果（Table 2参照）、回顧的プレテストーポストテスト間では、5項目すべてで有意な違いが認められた。またポストテストーフォローアップ間でも全項目で有意な違いが認められた。他方、プレテストーポストテスト間では、「『将来こういう生活をしたい』という、人生の満足につながる目標をもととすること」においてのみ有意な違いが認められなかった。

効果量の面からは、回顧的プレテストーポストテストにおいては、0.30から0.37の r が認められた。すなわち、すべての項目で中程度の効果が確認できたといえる。ポストテストーフォローアップ間では、0.23から0.36であった。プレテストーポストテストでは、有意とは判断できなかった『『将来こういう生活をしたい』という、人生の満足につながる目標をもととすること』の0.14から0.36

Table 1 4種の測定における平均値および標準偏差

	ブレテスト		回顧的ブレテスト		ポストテスト		フォローアップ	
	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差	平均値	標準偏差
大学での学びを有意義なものにしようとする事	3.96	0.83	3.95	0.94	4.34	0.75	3.97	0.85
卒業後の進路をイメージすること	3.54	1.12	3.59	1.10	4.15	0.95	3.65	1.13
所属している学部・学科での学びによって、どういう特徴的な力が身に付くのかを考えること	3.68	0.90	3.67	0.95	4.04	0.77	3.71	0.88
「将来こういう生活をしたい」という、人生の満足につながる目標をもとうとすること	3.89	1.07	3.75	0.97	4.11	0.89	3.76	1.04
働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること	3.54	1.06	3.54	0.99	3.98	0.87	3.36	1.01

Table 2 4種の測定間における検定結果と効果量、および信頼区間

	ブレテスト			回顧-ポスト			ブレ-回顧			ポスト-フォロー			ブレ-フォロー			回顧-フォロー		
	検定	r	95%CI	検定	r	95%CI	検定	r	95%CI	検定	r	95%CI	検定	r	95%CI	検定	r	95%CI
大学での学びを有意義なものにしようとする事	*	0.34	0.15 : 0.51	*	0.34	0.15 : 0.51	n.s.	0.00	0.20 : 0.20	*	0.29	0.10 : 0.47	n.s.	0.01	-0.19 : 0.21	n.s.	0.03	-0.18 : 0.22
卒業後の進路をイメージすること	*	0.36	0.17 : 0.52	*	0.36	0.18 : 0.53	n.s.	0.02	-0.18 : 0.22	*	0.29	0.09 : 0.46	n.s.	0.06	-0.14 : 0.26	n.s.	0.03	-0.17 : 0.23
所属している学部・学科での学びによって、どういう特徴的な力が身に付くのかを考えること	*	0.26	0.06 : 0.44	*	0.31	0.12 : 0.48	n.s.	0.00	-0.20 : 0.20	*	0.24	0.04 : 0.42	n.s.	0.03	-0.17 : 0.23	n.s.	0.04	-0.17 : 0.23
「将来こういう生活をしたい」という、人生の満足につながる目標をもとうとすること	n.s.	0.14	-0.07 : 0.33	*	0.30	0.11 : 0.47	n.s.	0.11	-0.10 : 0.30	*	0.23	0.03 : 0.41	n.s.	0.08	-0.12 : 0.28	n.s.	0.00	-0.20 : 0.20
働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること	*	0.29	0.10 : 0.47	*	0.37	0.19 : 0.53	n.s.	0.00	-0.20 : 0.20	*	0.36	0.17 : 0.52	n.s.	0.13	-0.07 : 0.33	n.s.	0.13	-0.07 : 0.32

検定はShapiro-Wilk検定の結果。* : Bonferroni法による補正を行い、 $\alpha = 0.0083$ を基準として有意と判断されたもの。

までの値が認められた。回顧的プレテスト→ポストテストの結果と比較すると、プレテスト→ポストテストの方が効果量大きい項目はなく、「『将来こういう生活をしたい』という、人生の満足につながる目標をもとうとすること」、「所属している学部・学科での学びによって、どういう特徴的な力が身に付くのかを考えること」、「働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること」で、小さい傾向にあったといえるだろう。

加えて、検定結果では有意とは認められなかったが、「『将来こういう生活をしたい』という、人生の満足につながる目標をもとうとすること」で、プレテスト→回顧的プレテスト間において、また「働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること」で、プレテスト→フォローアップ、回顧的プレテスト→フォローアップ間において0.1を超える小さな効果量が認められた。

これらの結果に示されるように、講演会によって、5つの企画趣旨に関する意識は概ね高まった。ここから、今回の講演会、すなわち社会に出ている卒業生（先輩）の、学生時代の意識や生活、さらに現在の職業を含む生活やそれらの関連にかかわる話を聞くことによって、今回項目を準備した5つの意識にポジティブな変化が生じたと判断してよからう。この点では、講演会は所期の目的に達することができており、効果的なものであったと評価できる。ところが、変化した意識は持続せず、約1ヶ月後には、講演会前の状況とほぼ同じ様相となることも確認された。そのため、意識変化だけでなく、それが継続するような講演会にするための工夫、もしくは変化した意識が継続するようなフォローアップ企画の追加などを検討する必要性が示唆されよう。

さて、今回の縦断的な調査から明らかになった諸点は以上のようなものであるが、本研究では回顧的プレテストやレスポンスシフト概念を導入する意義を検討することも目的としている。プレテスト→ポストテスト、回顧的プレテスト→ポストテストで、上述のような結果の差が認められた。2項目では両デザインでほぼ同様の変化が認められたが、残る3項目については、検定の結果、もしくは効果量の大きさが異なる傾向があった。実施者の立場からすれば、回顧的プレテスト→ポストテストの方が予測と整合的な結果といえる。またこれはプレテスト→ポストテストという計画は効果を過小評価しがちであるという従来の指摘（Howard, Ralph et al., 1979 など）と整合的な結果といえよう。ここから、本調査においてはレスポンスシフトが生じている可能性が考えられる。そこで次に、レスポンスシフトの観点から検討を試みる。

プレテスト→ポストテスト間、回顧的プレテスト→ポストテスト間で2つのテスト間の差（ポストテストからプレ／回顧的プレテストを減じた）を算出した。それらを差の動き（プラス方向、変化なし、マイナス方向）に要約したものと共に、Table 3に示す。またTable 3には、差の得点の平均値、標準偏差も示した。これらがレスポンスシフトの様相を示しているといえよう。

まず「大学での学びを有意義なものにしようとする」という項目においては、プレテスト→ポストテスト、回顧的プレテスト→ポストテストの両デザイン間でほとんど違いがない。この項目ではレスポンスシフトはほとんど生じないといえるだろう。

「卒業後の進路をイメージすること」は、最もプラス方向への移動者が多い項目である。これは効果量が両デザイン共に0.36と高めであったこととも関連していよう。差得点の平均はプレテスト→ポストテストの方が回顧的プレテスト→ポストテストよりもやや大きい。レスポンスシフトの観点からは、講演を通して自分を判断する基準が緩まったとも考えられる。講演前は自分を過小評価していた、すなわち自分は自分で思っている以上に将来をイメージしていた、などという変化が生じた者が含まれていたのかもしれない。

Table 3 プレテストと回顧的プレテストにおける差の様相

	得点差 (ポスト-プレ/回顧的プレ) の度数											得点差の平均, SD		
	-4	-3	-2	-1	0	1	2	3	4	マイナス	変化なし	プラス	平均	SD
大学の学びを有意義なものにしようとする事														
プレテスト	0	0	0	7	52	31	5	1	0	7	52	37	0.39	0.75
	(0.0)	(0.0)	(0.0)	(7.3)	(54.2)	(32.3)	(5.2)	(1.0)	(0.0)	(7.3)	(54.2)	(38.5)		
回顧的プレテスト	0	0	0	7	53	28	7	1	0	7	53	36	0.40	0.77
	(0.0)	(0.0)	(0.0)	(7.3)	(55.2)	(29.2)	(7.3)	(1.0)	(0.0)	(7.3)	(55.2)	(37.5)		
卒業後の進路をイメージすること														
プレテスト	0	1	0	9	38	29	16	3	0	10	38	48	0.60	1.04
	(0.0)	(1.0)	(0.0)	(9.4)	(39.6)	(30.2)	(16.7)	(3.1)	(0.0)	(10.4)	(39.6)	(50.0)		
回顧的プレテスト	0	0	1	7	42	32	12	2	0	8	42	46	0.55	0.92
	(0.0)	(0.0)	(1.0)	(7.3)	(43.8)	(33.3)	(12.5)	(2.1)	(0.0)	(8.3)	(43.8)	(47.9)		
所属している学部・学科での学びによって、どういう特徴的な力が身に付くのかを考えること														
プレテスト	0	0	0	16	43	24	12	1	0	16	43	37	0.36	0.94
	(0.0)	(0.0)	(0.0)	(16.7)	(44.8)	(25.0)	(12.5)	(1.0)	(0.0)	(16.7)	(44.8)	(38.5)		
回顧的プレテスト	0	0	0	10	50	26	10	0	0	10	50	36	0.38	0.81
	(0.0)	(0.0)	(0.0)	(10.4)	(52.1)	(27.1)	(10.4)	(0.0)	(0.0)	(10.4)	(52.1)	(37.5)		
「将来こういう生活をしたい」という、人生の満足につながる目標をもととすること														
プレテスト	0	0	4	14	48	21	5	3	1	18	48	30	0.23	1.07
	(0.0)	(0.0)	(4.2)	(14.6)	(50.0)	(21.9)	(5.2)	(3.1)	(1.0)	(18.8)	(50.0)	(31.3)		
回顧的プレテスト	0	0	1	7	53	27	7	1	0	8	53	35	0.36	0.81
	(0.0)	(0.0)	(1.0)	(7.3)	(55.2)	(28.1)	(7.3)	(1.0)	(0.0)	(8.3)	(55.2)	(36.5)		
働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること														
プレテスト	0	0	2	12	38	30	14	0	0	14	38	44	0.44	0.96
	(0.0)	(0.0)	(2.1)	(12.5)	(39.6)	(31.3)	(14.6)	(0.0)	(0.0)	(14.6)	(39.6)	(45.8)		
回顧的プレテスト	0	0	1	2	59	22	12	0	0	3	59	34	0.44	0.78
	(0.0)	(0.0)	(1.0)	(2.1)	(61.5)	(22.9)	(12.5)	(0.0)	(0.0)	(3.1)	(61.5)	(35.4)		

() 内の数値は割合を示す

「所属している学部・学科での学びによって、どういう特徴的な力が身に付くのかを考えること」は、今回の項目の中でマイナス方向への移動者がやや多めであることが特徴といえよう。全体で見ると、講演を通して意識が高まる項目と結論できるが、回顧的プレテスト―ポストテスト差においてもマイナス方向への移動者がある程度含まれるということは、中には講演を聞いた結果、思ったほど考える必要はないのではないかと考えた参加者が存在すると考えられる。

「『将来こういう生活をしたい』という、人生の満足につながる目標をもととすること」は、両調査デザイン間での違いが最も顕著であり、レスポンスシフトが生じている可能性を指摘できよう。この項目に対する判断基準はプレテスト時には過大評価気味であり、講演を通して調整され、厳しくなったと考えられる。この項目は、プレテスト―ポストテストでは有意な差が認められず、効果量も小さいものであった。今回の調査から、プレテスト―ポストテストと回顧的プレテスト―ポストテストのどちらが正確に変化を把握できているのかを判断することは難しいが、レスポンスシフトという概念を考慮に入れて検討する必要があることを顕著に示す結果といえるだろう。

最後に、「働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること」は、デザイン間で差の分布の様相は異なるものの、平均値で見るとほぼ等しいという特徴がある。Table 3 から、プレテスト―ポストテストでは、差の分布におけるバラつきが大きく、回顧的プレテスト―ポストテストよりも、マイナス方向、プラス方向のいずれでも変化者が多い。回顧的プレテスト―ポストテストでは変化なしが60%を超え、今回の項目の中で最も変化する者が少ないことが示されている。また効果量の点では、回顧的プレテスト―ポストテスト間の効果量は大きく、またプレテスト―ポストテストでの効果量との差もやや大きい。これは、講演前の判断基準は不安定で、対象全体にみられる過大もしくは過小といった評価の全体的特徴が認められないことを示すのかもしれない。この項目も、変化を正確に把握するためには、レスポンスシフトという概念を考慮に入れより詳細に検討すべきものといえるだろう。

回顧的プレテストのメリットとして（もしくは介入前にプレテストを行うことのデメリットとして）、しばしば指摘されるのは、介入前は情報不足などのため自分の様相を正しく回答することができないという点である（Pratt et al., 2000 など）。そのような場合、人は自分を過大評価しがちであり、それゆえプレテスト得点は高めになり、回顧的プレテストの得点が低めになる傾向をもつとされる。この説明に最も合致するであろう項目が、「『将来こういう生活をしたい』という、人生の満足につながる目標をもととすること」といえるだろう。

また、「卒業後の進路をイメージすること」や「働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること」でも、それぞれ特徴的なプレテストと回顧的プレテストの差異が生じていそうである。特に「働くことが、社会的に、また自分にとってどういう意義を持つのかを考えること」では講演前の判断基準は不安定ではないかと推察され、介入前は自分の様相を正しく回答できていないとも考えられる。またこれら3つの項目は、いずれも現在の大学生活ではなく将来のことについての内容である。このような、証拠や根拠に基づくものではなく、想像力やイメージに頼るような内容の項目は、介入前の回答について注意深く検討する必要があるのかもしれない。

なお、以上の検討はプレテストと回顧的プレテスト違いをレスポンスシフトのみで解釈した。ところが、回顧的プレテストに関しては、他にも多種のバイアス（たとえば想起バイアス、社会的望ましさ、変化に関する暗黙の理論（Norman, 2003）など）が影響している可能性があることが指摘されている。以上の解釈は、ひとつの例であることに留意しておく必要があるだろう。

まとめと今後への示唆

今回の調査では、1度のキャリア関連講演会とはいえ、参加者に企画意図にそった望ましい変化を生んでいることが概ね確認できた。しかしながらその効果は持続せず、約1ヶ月後には、講演会前の状況とほぼ同じ様相となっていることも確認された。また本研究は、プレテスト→ポストテスト→フォローアップというデザインに回顧的プレテストを組み入れて検討を行い、プレテスト→ポストテストと回顧的プレテスト→ポストテスト間で、項目によって結果に差が生じるという、レスポンスシフトが存在するであろうことも示唆された。こういった結果を踏まえて、キャリア研究に回顧的プレテストという方法やレスポンスシフト概念を導入する意義を検討する。

まず、回顧的プレテストを導入する意義について取り上げる。冒頭でも触れたように、我が国では、介入の効果を検討する際にはプレテスト→ポストテストを用いるという暗黙の了解があるように思われる。もしレスポンスシフトが存在するのであれば、このデザインでは変化の正しい様相を把握することはできない。また、一般的には統計的に有意な変化が認められた場合に論文などを通して公表され、認められなかった場合には公表されることはないといえるだろう。Howard, Ralph et al. (1979) の第1研究のように、プレテスト→ポストテストを用いることで予測とは逆の統計的検定結果になったり、変化を過小評価したりしているのであれば、本質的には有益である介入知見が公表されなかったり、逆の効果をもつ介入と誤認されてしまっているかもしれない。

たとえば、石黒他 (2022) は、医療系学生を対象とした起業家精神の涵養を目指すキャリア教育の効果を、プレテスト→ポストテストデザインで検討している。そして、自立性や創造性の面では得点が有意に高まったが、達成欲求やリスクテイキングについては有意に得点が下がったことを報告している。石黒他は、この低下という結果に対して積極的な解釈は行っていないが、これは予測や仮説と矛盾する結果であり、Howard, Ralph et al. (1979) が直面した状況と同じといえるかもしれない。すなわち起業家精神の涵養を目指した介入の結果、プレテスト時とポストテスト時では回答の際の判断基準が変化した、すなわちレスポンスシフトが生じたゆえの結果かもしれないと考えられる。

こういった問題への対応として提案された回顧的プレテストを採用すれば、介入によってどのような変化が生じているのかを一層明らかにできると期待できるだろう。介入による変化をより正確に把握するために、回顧的プレテストを導入する意義は大きいと考えられる。

なお、今回の結果は、プレテストを回顧的プレテストに置き換えることで変化が明確になる（大きくなる）ことを示唆するが、この解釈には慎重であるべきだろう。重要なのは、変化を正しく把握することであり、介入実施者にとって望ましい結果を得ることではない。以前から指摘されているように（たとえば Howard & Tinsley, 1979）、回顧的プレテストデータの収集をプレテスト→ポストテストのデザインに組み入れ、知見を積み重ねることで変化を正しく把握することに近づけるだろう。そして、いずれの測定方法でもバイアスが発生する可能性があるため、Hill (2020) が指摘するようにバランスのとれた議論をすることが重要であろう。

次に、レスポンスシフト概念の導入という点について検討する。Howard, Millham et al. (1981) は、対象となる概念そのものに対する理解や認識を変えるように計画された介入を評価する場合、介入がその目標を達成する限りにおいて、対象者が自己報告する理解や認識はプレテスト時とポストテスト時で異なると指摘する。このような理解・認識の変化がレスポンスシフトであり、この存在は

プレテストとポストテストの得点を比較すること自体を不適切なものにする。キャリア教育では、生き方や生活、職業といった概念に対する理解や認識自体にアプローチし、その変化を目的とすることも多いといえるだろう。こういった側面に介入し、またその変化をとらえるということは、レスポンスシフトを生じさせることを目的とし、またその変化をとらえるということにほかならない。そうであれば、プレテスト→ポストテストが示す変化は、その様相を正しく反映しないであろう。そのため、たとえば職業観や価値観の変容などを狙いとする場合、また職場体験やインターンシップなどによる変化をとらえようとする場合などでは、レスポンスシフトという概念を導入することが特に有益と考えられる（なお、回顧的プレテストやレスポンスシフトといった用語は用いられていないが、たとえば浦上（2016）のように回顧的な方法を用いた先行研究がないわけではない）。

また今回の結果は、講演会の影響は一時的なものに留まり、長期的にみると実質的には効果が認められないことを示唆するとも考えられる。ところが、このような回答の回帰的な現象が認められたことは、支援者にとっては驚くに値しないかもしれない。介入の効果、特に今回の講演会のような介入の場合、その効果が長続きしないことは経験的に共有されているのではないだろうか。論理的に考えても、1度の講演会で今回認められたような得点の変化が生じ、かつその状態が継続するならば、数回の講演会で尺度の測定範囲の上限に達してしまう。これも現実的ではない推測であろう。

しかしながら、この点は、極めて重要な検討ポイントであることは明らかである。すなわち、講演会による変化が一時的であり短期間で実施前の状況に戻るのならば、キャリア支援の一環としてこのような企画を実施する意味、意義は、少なくとも参加者側にはないということに等しい。たとえば安田・渡辺（2008）が指摘するように、プログラム評価は、それを継続するか否かの判断にもつながる。参加者の時間はもちろん、登壇者をはじめ支援者、実施者の時間、さらに労力も投入される企画を、効果はあるが、それは一時的で、通例としてすぐに実施前に戻るようなものとしか説明できないこと、さらにはその前提で企画を継続することは大きな問題といわざるをえない。

ところが、現在でも今回のような講演会は各所で行われている。これは、支援者は介入の効果が長続きしないことを実感しつつも、未解明ではあるが、長期的な効果があることを感じ取っているからかもしれない。ここから、介入の効果が長続きしないこと、測定上の数値が介入前の値に戻ることをどのようにとらえるべきなのかという課題が浮上しよう。介入の効果がすべて失われリセットされるという考え方もできるであろうが、レスポンスシフト概念を援用すれば、日常的に自分を判断する基準の再較正（recalibration）が行われており、それゆえに測定指標上の得点が介入以前の値に近づく結果とも考えられよう。Sprangers & Schwartz（1999）は、患者の客観的な悪化にもかかわらず、認識されたQOLに変化がないという現象からレスポンスシフトが認識されたとして、自らのQOLにおけるレスポンスシフトのモデルに、人は自身についてできるだけよく感じたいと願っているという前提をおいている。そのためレスポンスシフトは、恒常性を維持または回復することに向けられると指摘するが、これと類似したことが今回の測定上の数値の回帰にもいえるかもしれない。すなわち、各測定項目における回答には個人にとって適当な一定の位置があり、それを恒常的に維持するために再較正が行われているのかもしれないとも考えられよう。実際的にいえば、あるときに変化があったとしても（変わったと実感しても）、時間とともにそれが普通（普段の自分）になっていくようなものである。

学校などでのキャリア教育は、長期にわたってさまざまな支援、介入が行われる。この場合、個々の支援、介入の効果の総計が全体としての効果になるわけではないだろう。またキャリア教育では、

生き方や職業といった概念に対する理解や認識自体にアプローチし、その変化自体を目的とする場合もある。さらに、そうした介入がなくても、子どもは徐々に成長、成熟していく。こういったさまざまな要因が複雑にからみ合う変化については、レスポンスシフトという概念を導入することがその解明に有益と考えられる。

なおレスポンスシフトに関しては、Sprangers & Schwartz (1999) や Schwartz & Sprangers (1999), Norman (2003) などが QOL 研究に対して興味深い提言を行っている。また Schwartz & Sprangers (2010) は、より洗練された研究デザインについて議論を行っている。キャリア教育においてどのようにレスポンスシフト概念を参照すべきか、またどのような研究を蓄積していくかについての議論に、こういった先行研究は有益な示唆を与えてくれると期待できよう。

先にも述べたように、我が国では介入の効果を検討する際には、プレテスト―ポストテストを用いる場合が多く、キャリア研究の領域でも同様といえるだろう。そのため、プレテスト―ポストテストにおいて見いだされた差が正しく変化を表しているのかという点の検討が不足している。また、キャリア教育においては、その特徴からレスポンスシフトについての検討が不可欠とも考えられる。本研究で行った調査結果は、こういった点の検討が必要であることを示唆するものである。キャリア教育による対象者の変化を把握することの重要性は改めて指摘するまでもないだろうが、残念ながら「どうやって変化を把握するか」という研究デザイン自体についての議論は低調といえるだろう。キャリア研究において、今後、介入の効果を検討する方法に関する議論が活発になることが期待される。

引用文献

- Bhanji, F., Gottesman, R., de Grave, W., Steinert, Y., & Winer, L. R. (2012). The retrospective pre-post: A practical method to evaluate learning from an educational program. *Academic Emergency Medicine, 19*, 189-194.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171-246). Chicago: Rand McNally.
- Chasteen, S., & Chattergoon, R. (2019). *A comparison study of pre/post-test and retrospective pre-test for measuring faculty attitude change*. Physics Education Research Conference 2019, Provo, UT. 93-98. <https://doi.org/10.1119/perc.2019.pr.Chasteen>
- Geldhof, G. J., Warner, D. A., Finders, J. K., Thogmartin, A. A., Clark, A., & Longway, K. A. (2018). Revisiting the utility of retrospective pre-post designs: the need for mixed-method pilot data. *Evaluation and Program Planning, 70*, 83-89.
- Howard, G. S., & Dailey, P. R. (1979). Response-shift bias: A source of contamination of self-report measures. *Journal of Applied Psychology, 64*, 144-150.
- Howard, G. S., Millham, J., Slaten, S., & O'donnell, L. (1981). Influence of subject response style effects on retrospective measures. *Applied Psychological Measurement, 5*, 89-100.
- Howard, G. S., Ralph, K. M., Gulanick, N. A., Maxwell, S. E., Nance, D. W., & Gerber, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Applied Psychological Measurement, 3*, 1-23.
- Hill, L. G. (2020). Back to the future: Considerations in use and reporting of the retrospective pretest. *International Journal of Behavioral Development, 44*, 184-191.
- Hill, L. G., & Betz, D. L. (2005). Revisiting the retrospective pretest. *American Journal of Evaluation, 26*, 501-517.
- 石黒順子・村川修一・渡邊亜紀子 (2022). 医療系大学でのキャリア教育の効果測定 日本キャリア教育学会第 44 回研究大会 大会発表論文集, 90-91.

- 小島秀夫・篠原清夫 (2011). 回顧的解答の安定性・不安定性について 茨城大学教育学部紀要 (人文・社会科学, 芸術), 60, 85-98.
- Krumboltz, J. D. (2009). The happenstance learning theory. *Journal of Career Assessment*, 17, 135-154.
- Lane, W. B., & Adhikari, R. Y. (2019). *Evaluating the presence of response-shift bias in the CLASS with a two-pass survey*. Physics Education Research Conference 2018, Washington, DC. 219-222. <https://doi.org/10.1119/perc.2018.pr.Lane>
- 美崎定也・古谷英孝・廣幡健二・木原由希恵・田中友也・坂本雅光・三井博正・佐和田桂一・西野正洋・西法正・杉本和隆 (2013). 日本人の人工膝関節置換術後患者における患者立脚型アウトカムによる QOL 評価へのレスポンスシフト現象の影響 理学療法学 Supplement, 40(2) (第 48 回日本理学療法学会大会 抄録集), C-P-24. <https://doi.org/10.14900/cjpt.2012.0.48101785.0>
- 水本篤・竹内理 (2008). 研究論文における効果量の報告のために—基礎的概念と注意点— 英語教育研究, 31, 57-66.
- Nielsen, R. (2011). A retrospective pretest-posttest evaluation of a one-time personal finance training. *Journal of Extension*, 49(1), Article 1FEA4. <http://www.joe.org/joe/2011february/a4.php>
- Nimon, K. (2014). Explaining differences between retrospective and traditional pretest self-assessments: competing theories and empirical evidence. *International Journal of Research & Method in Education*, 37, 256-269.
- Nimon, K., & Allen, J. (2007). A review of the retrospective pretest: Implications for performance improvement evaluation and research. *Workforce Education Forum*, 44, 36-55.
- Norman, G. (2003). Hi! How are you? Response shift, implicit theories and differing epistemologies. *Quality of Life Research*, 12, 239-249.
- 岡田康孝・池田誠喜・芝山明義 (2020). いじめ防止のための道徳科・特別活動・総合的な学習の時間における実践研究—ロジックモデルに基づいた実践— 鳴門教育大学授業実践研究：授業改善をめざして, 19, 105-113.
- Pratt, C. C., McGuigan, W. M., & Katzev, A. R. (2000). Measuring program outcomes: Using retrospective pretest methodology. *American Journal of Evaluation*, 21, 341-349.
- Schwartz, C. E., & Sprangers, M. A. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48, 1531-1548.
- Schwartz, C. E., & Sprangers, M. A. (2010). Guidelines for improving the stringency of response shift research using the then-test. *Quality of Life Research*, 19, 455-464.
- Sprangers, M. A., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine*, 48, 1507-1515.
- 鈴鴨よしみ (2015). QOL 評価研究と行動医学—レスポンスシフトの視点から— 行動医学研究, 21, 12-16.
- 浦上昌則 (2016). 就職活動と職業観の変容 アカデミア 人文・自然科学編, 12, 53-65.
- Vo, M, Dallaghan, G.B., Borges, N., Gill, A. C., Good, B., Gollehon, N., Mehta, J.J., & Balmer, D. (2021). Planning for happenstance: Helping students optimize unexpected career developments. *MedEdPORTAL*, 17, 11087. https://doi.org/10.15766/mep_2374-8265.11087
- 渡邊宏樹・隆島研吾 (2014). 慢性進行性疾患における主観的 QOL 評価と Then Test の重要性 理学療法学 Supplement, 41(2) (第 49 回日本理学療法学会大会 抄録集), 1470. <https://doi.org/10.14900/cjpt.2013.1470>
- 安田節之 (2011). プログラム評価—対人・コミュニティ援助の質を高めるために— 新曜社
- 安田節之・渡辺直登 (2008). プログラム評価研究の方法 新曜社