

教員研修の活用報告に基づく行動変容度合の定量化

岡山県総合教育センター 上岡 伸, 中藤路子

研究の概要

本研究の目的は、教員研修の追跡調査から得られる、学びの活用報告に対する研修機関の客観的評価と、活用の有無に関する受講者のシンプルな自己申告とを比較して、どちらが行動変容の指標として適するかを検討することである。複数回の調査を通じ、中堅教諭等資質向上研修のうちの3つの共通研修を受講した教員から有効な 154 名分延べ 462 件のデータが得られた。マルチレベルロジスティック回帰モデルによる分析の結果、直後に主体的・具体的な行動計画を立てた研修では、追跡調査において研修で推奨された行動を報告する傾向にあった一方、行動計画の評価結果と活用有無の自己申告との間に統計的に有意な関係は見られなかった。Kirkpatrick のモデルに照らし、行動変容の指標としては、活用報告に対する研修機関の客観的評価の方が妥当であることが示唆された。

キーワード

教員研修、効果測定レベル、混合評価、マルチレベルモデル、キーワード法

1. 問題と目的

(1) レベル3測定の価値

中央教育審議会が示した「新たな教師の学びの姿」を踏まえ（文部科学省 2021）、教員研修の在り方は大きく変わろうとしている。文部科学省（2023a）は「教員研修プラットフォーム」という学習管理システムを構築し、令和6年度から運用を開始するほか、独立行政法人教職員支援機構（2023）は「研修観」（文部科学省 2022）の転換を掲げ、探究型の研修デザインに取り組んでいる。これらの事業展開と同時に、文部科学省（2023b）は「教員研修の高度化に資するモデル開発事業」を実施し、「教員研修の成果確認と評価モデルの確立」も図っている。これは裏を返せば、教員研修の評価モデルが未だ確立されていないことを意味している。

国内の教員研修機関では、研修評価の枠組みとして Kirkpatrick の4段階モデル（Kirkpatrick and Kirkpatrick 2006）が広く普及している（上岡 2022）。レベル1（反応）・2（学習）・3（行動）・4（業績）は段階的に達成される想定であるが、近年の研修評価研究においては4段階の中でも特にレベル3に位置付けられる行動変容度合の測定が重要であると考えられている。レベル4との間にレベル1・レベル2よりも高い相関を示し（Saks and Burke 2012）、なおかつレベル4よりも低コストで測定可能なレベル3は、研修評価の実務において重要な指標となる（中原ほか 2022）。さらに、レベル3測定のため実施される追跡調査は、研修で学んだことを実践するよう受講者に働きかけるリマインダーとしても作用し受講者の行動変容を促進する（Saks and Burke 2012; レビューとして、中原ほか 2022）。

(2) レベル3 測定の課題

ところが教員研修を担う教育センターにおけるレベル3 測定の実施状況はというと、平成18年度時点で一部の研修なりとも追跡調査を行う都道府県・政令指定都市 ($n=66$) の割合は62.1%であったものが(全国教育研究所連盟 2007)、令和2年度時点 ($n=63$) の調査では9.5%の教育センターが追跡調査を取り止めたと回答しており、実施率は49.2%まで後退していた(大阪府教育センター 2020)。

こうした事態の背景には、大きく二つの課題を指摘できる。一つは測定上のコストである。追跡調査への回答に当たって受講者側に生じる心理的・物理的コストは、低い回収率(大阪府教育センター 2020)に反映されることになる。そのため研修運営側は調査の企画・実施・集計のコストに加え、回答協力を個別に依頼するなど、回収率向上の取組にもコストを支払うことになる。追跡調査実施率の後退の主な理由は、双方に生じるこれらのコストが働き方改革の要請と衝突することにあると考えられる。

もう一つの課題は、測定結果の定量化の問題である。研修後に職場で何を実践したのか、その学びの活用内容は質的に把握される場合が多い。例えば日本教育工学会論文誌に報告されている先行研究を参照すると、徳村ほか(2006)は研修内で作成したアクションプランの実施状況について複数の自由記述項目で尋ねることでレベル3を評価している。小清水ほか(2014; 2016)はプランの実施状況を選択項目と自由記述項目を組み合わせで評価している。金子ほか(2021)と金子ほか(2023)は行動面・認識面の変化を自由記述で尋ねており、金子ほか(2021)では管理職にも受講者の変化を尋ねている。これらの先行研究は、受講者の記述から行動変容に関するリッチな情報が得られることを示唆しているが、いずれにおいても記述データをどう定量化するかは論じられていない。

しかし、教育センターが個別の研修改善を図る上では、改善の前後で評価結果を定量的に比較可能であることが求められる。また、評価結果を年度ごとのKPI(重要業績評価指標)として用いる上でも、質的データのままで運用が難しい。活用したかどうかを何らかの行動尺度によって測定すればたちまち量的評価が可能だが(e.g., 上岡 2023; 上岡・中藤 2024a)、尺度開発のコストが問題になるほか、学びの活用の際に行われた自己調整(Blume *et al.* 2019)に関する様々な文脈的情報を削ぎ落とすことにもなる。近年中原ほか(2022)が、数値とともに現場の生々しいエピソードを把握することによる研修の混合評価を提案しているように、評価結果のインパクトを高め、研修改善に向けたヒントを得る上では、職場で何が起きたのかを把握することが重要となる。

(3) レベル3 測定結果の定量化

あるいは小清水ほか(2014; 2016)が評価項目の一つとして採用したように、推奨行動実施の有無を1項目で尋ねれば一応の2値データが得られるが、その回答単独では実践の内容が推奨行動であるかどうかを確認できず、レベル3の指標とするには妥当性に疑問が残る。全国のエducationセンターの追跡調査にも、(1)最初に研修での学びを活用したかどうかを選択項目で尋ねた後、(2)次にどう活用したか(あるいはなぜ活用しなかったのか)を自由記述で問う構成が散見されるが(e.g., 兵庫県立教育研修所 2009; 上越教育大学 2008; 奈良県立教育研究所 2009)、同様に定量化のための手続きは示されていない。(1)で「活用した」と報告していても、(2)の記述が研修の推奨行動とは全く関係のない行動であるケースも生じるだろう。特定の推奨行動が実行されたかどうかを評価するためには、むしろ(2)の記述に着目し、その内容が推奨行動に該当するかどうかを評価して定量化することが妥当と考える。

(1)の自己申告と(2)の評価結果のどちらがレベル3として妥当であるかを判断する上では、Kirkpatrickの4段階モデルが参考になる。Kirkpatrickの4段階モデルはレベル2とレベル3の間に因果関係を想定しており(Alliger and Janak 1989)、Blume *et al.* (2010)のメタ分析では小程度の正の効果量

が示されている（研修後の知識： $\rho=.24$ 、研修後の自己効力感： $\rho=.20$ ）。そこで本研究では、レベル2測定結果との関連性を確かめることで、（１）の自己申告と（２）の評価結果とではレベル3の指標としてどちらが相応しいかを検討することにした。

同時に、本研究では一つ目の課題への対応も試みる。徳村ほか（2006）はオンライン調査の導入によって研修評価の取組が容易になることを示唆しており、上岡（2022）は教育センターにおける当日調査のオンライン化によって、自由記述項目によるレベル2測定が効率化した旨を報告している。そこで本研究の実践においてもオンライン調査を導入し、受講者と研修運営の双方に負担の少ない追跡調査の在り方を模索することにした。

2. 方法

（1）調査対象研修

研修評価のための調査は、岡山県総合教育センターにおいて行われた。追跡調査の対象となったのは、中堅教諭等資質向上研修（小学校・中学校・高等学校・特別支援学校について、それぞれ共通研修5日と選択研修2日）のうち、2022年5月～10月にかけて実施される12の研修（共通研修3日×4校種）である。各研修の概要を表1に示す。共通研修の内容は次のとおりである。

GIGA スクール対応基礎研修：授業内での有効な ICT 活用促進を目的として企画された。研修内容は、デジタル・シティズンシップ（内閣府 2022）についての講義・情報交換（70分）と、「授業での ICT 活用場面と主体的な学習活動の充実」についての講義及び授業改善の構想（80分）、そして、「自校の GIGA スクール構想実現」に向けた課題解決ワークシートの作成（30分）で構成された。小・中学校は同じ指導主事が同じ資料を用いて研修を行ったが、高等学校と特別支援学校ではそれぞれ異なる指導主事が異なる資料を用いて研修を行った。

B&S 研修：B&S 研修（岡山県総合教育センター 2015）は、若手教員への指導助言を促進することを目的として企画された。当日は初任教員と中堅教員との混成グループが編成され、研修内容は中堅教員による実践上の工夫に関する発表・協議（午前2時間45分）と、初任教員が中堅教員から指導助言を受ける演習（午後2時間45分）で構成された。午後の演習について、小学校では、全グループで4年生算数「簡単な場合についての割合」をベースに初任教員が模擬授業を行い、中堅教員が指導助言を行った。中学校・高等学校では教科別にグループが編成され、初任教員が持参した授業動画（中学校）や学習指導案

表1 調査対象となった研修の概要

研修の名称	推奨行動	日程	受講者	当日調査	回答者	追跡調査	回答者
共通研修1日目							
① GIGAスクール対応基礎研修(小)	授業におけるICT活用	半日	92	5月下旬	87	8月上旬	79
② GIGAスクール対応基礎研修(中)	授業におけるICT活用	半日	54	5月下旬	52	8月上旬	40
③ GIGAスクール対応基礎研修(高)	授業におけるICT活用	半日	88	5月下旬	75	7月下旬	66
④ GIGAスクール対応基礎研修(特)	授業におけるICT活用	半日	45	5月下旬	40	7月下旬	35
共通研修3日目							
⑤ B&S研修(小)	若手教員への指導助言	全日	92	8月上旬	79	10月下旬	84
⑥ B&S研修(中)	若手教員への指導助言	全日	53	8月上旬	40	10月下旬	48
⑦ B&S研修(高)	若手教員への指導助言	全日	88	7月下旬	66	10月上旬	85
⑧ B&S研修(特)	若手教員への指導助言	全日	45	7月下旬	35	10月下旬	40
共通研修4日目							
⑨ 異業種から学ぶ人材育成(小)	組織的な人材育成	半日	89	10月下旬	84	2月上旬	82
⑩ 異業種から学ぶ人材育成(中)	組織的な人材育成	半日	51	10月下旬	48	12月上旬	49
⑪ 異業種から学ぶ人材育成(高)	組織的な人材育成	半日	88	10月上旬	85	12月上旬	86
⑫ 異業種から学ぶ人材育成(特)	組織的な人材育成	半日	45	10月下旬	40	2月上旬	39

注1)括弧内の小・中・高・特は対象となった校種を示す(小学校・中学校・高等学校・特別支援学校)

注2)⑨～⑫の追跡調査は共通研修5日目の当日調査と併せて実施。

（高等学校）に対する指導助言が行われた。特別支援学校では、初任教員が持参した2学期（後期）の単元計画に対する指導助言と、学級経営上の諸課題についての相談会が行われた。

異業種から学ぶ人材育成：組織的な人材育成を促進することを目的として企画された。衣料の製造販売を行う県内企業の人材開発部門の職員が講師として招かれ、小集団活動の事例や心理的安全性（Edmondson 2018）を形成するためのノウハウ等に関する講義が半日行われた。全ての校種について同じ講師が同じ資料を用いて研修を行った。

（2）調査実施方法

当日調査：各研修の終了後、Google Forms を用いて「研修講座の振り返り」という名称の記名式オンライン調査が行われた。研修の運営担当者は、これら調査の目的が、研修の振り返りを通じて受講者に学びを深めてもらうことと、今後の事業改善に資することの二点であることを説明した上で、研修日から一週間以内の回答協力を依頼した。

追跡調査：追跡調査の項目は、次の研修日の「研修講座の振り返り」の末尾に連ねて設けられた。年度内に複数の研修日が設けられる初任者研修や中堅教諭等資質向上研修では、前の研修日のレベル3測定を当日調査と同時に行うことで調査機会を減らし、受講者と研修運営双方のコストを抑えることが可能だと考えられた。結果的に、当日調査と追跡調査の間には概ね2～3カ月の期間が空いた（表1）。

調査の結果、1時点以上の欠測がある受講者や錯誤回答者を除くと、全ての調査回について有効な回答協力者は計154名となった（小学校52名、中学校32名、高等学校48名、特別支援学校22名）。

（3）調査内容

研修効果の測定方法は次のとおりである。

レベル1：当日調査で測定した。上岡（2022）と同じ4項目（「本研修で用いた資料は、内容を理解するのに役立ちましたか。」、「本研修プログラムの時間配分は、内容の理解を深めるのに適切でしたか。」、「研修の内容は、自らの力量向上に有効であったと感じますか。」、「研修の内容は、あなたが期待していた内容でしたか。」）について、5段階で評定を求めた。これらの項目は、知識技能等を獲得する上で効果的な研修内容であったかを把握するために、岡山県総合教育センターの研修企画担当部署の指導主事らが先行研究（Alliger *et al.* 1997; 兵庫県立教育研修所 2009; 奈良県立教育研究所 2008）を参考にして作成した項目である。

レベル2：当日調査で測定した。上岡（2022）の改訂版キーワード法により研修のキーワードを活用するための行動計画について記述を求めた。その際、計画内容を受講者が忘失する事態に備え、Google Forms の「回答のコピーを送信する」機能を用いることで、行動計画の控えが受講者のメールアドレスに送信される仕組みとした。なお、改訂版キーワード法は平松（2001）のキーワード法を改訂したものであり、受講者に研修のキーワードとキーワードに関する行動計画の記述を求め、内容の具体性等から知識・態度の獲得度合を評価する手法である。初任者研修受講者を対象とした調査では、測定された理解度は良好な評価者間信頼性（Cohen の $\kappa=.72$ ）を示し、また、その理解度は研修が推奨する行動の行動意図（Ajzen 1991）の研修前後の変化を説明したことから、測定結果について一定の妥当性が確認されている（上岡 2022）。

レベル3：追跡調査で測定した。上越教育大学（2008）の実践を参考に、表2のように項目の分岐を設けた。「現場の実践で生きて働く知識は、研修直後の理解ではなく、その後定着した知識」であることか

表2 追跡調査の項目構成

●●月●●日に実施した「●●●●●研修」についておたずねします。	
1. 現在でも印象に残っているキーワードを簡単にお書きください。(5つまで)	
2. 研修で学んだことを、所属校で活用しましたか。	
→活用した	
3. 活用した場面やその方法について、具体的に書いてください。	
4. 上記の活用を、今後も継続しようと思いますか。【1. 継続しない～5. 継続する】	
→活用していない	
3. 今後、機会があれば活用しますか。	
→活用するつもりである	
4. どのように活用する予定であるか、考えられる活用場面や方法について具体的に書いてください。	
→活用するつもりはない	
4. その理由を具体的に書いてください。	

ら（上越教育大学 2008）、問1で再度キーワードを問うことで、当日調査で挙げられたキーワードとの比較から研修内容が実践的であったかを把握できると考えられた。続く問2では、追跡調査時点で活用したかどうかについて選択項目を設け、問3以降では選択に応じた分岐を設けた。「活用した」者に対しては、活用内容について記述を求めた上で、その活用を継続するつもりかどうかを質問した。そして、「活用していない」者に対しては、その時点での行動計画または活用しない理由についての記述を求めた。

この質問構成について、岡山県総合教育センター内の事業評価担当部署から研修運営担当部署を経由して所属長までの稟議が行われ、内容的に妥当と判断された。

3. 結果

(1) 研修効果の定量化

154名から得られた3つの共通研修に関する調査回答である延べ462件のデータを、以下の手続きにより研修効果として定量化した。各研修効果の記述統計を表3に示す。

レベル1：4項目の信頼性係数を算出したところ、 $\alpha=.93$ と高い値を示した。そこで上岡（2022）に倣って4項目を単純加算し、レベル1の指標として用いることにした。

レベル2：教育センターの事業評価業務担当者2名が行動計画の評価を行った。まず、各研修で想定される推奨行動（表1）について研修運営担当部署への聞き取りが行われた。次に、当日の研修資料等を基に、各研修の行動計画に求められる具体性の基準について評価者間で協議が持たれた後、別々に行動計画の評価が行われた（Cohen の $\kappa=.69$ ）。そして、評価結果に差のあった箇所について、二者間の協議を経て最終的な評価が決定された（A～C 評価）。

表3 各研修効果の記述統計

研修	レベル1		レベル2			レベル3(自己申告)		レベル3(記述評価)	
	Mean	SD	A	B	C	活用した	活用していない	実行済	未実行
①	15.58	3.13	17	16	19	40	12	24	28
②	16.75	3.31	7	13	12	17	15	13	19
③	16.46	3.37	12	27	9	36	12	25	23
④	15.14	3.72	6	10	6	14	8	8	14
⑤	15.94	2.80	22	15	15	41	11	26	26
⑥	17.63	2.94	3	2	27	25	7	11	21
⑦	17.73	2.56	6	11	31	41	7	13	35
⑧	16.05	3.23	2	10	10	16	6	6	16
⑨	17.48	2.57	10	28	14	46	6	23	29
⑩	17.78	2.88	5	24	3	24	8	13	19
⑪	17.94	2.76	6	32	10	32	16	18	30
⑫	17.59	2.52	5	13	4	13	9	5	17

評価基準は次のとおりである（上岡 2022）。

A 評価：研修で得た知識を基に、推奨行動を実行するための主体的かつ具体的な行動計画を記述している。

B 評価：A 評価には該当しないが、推奨行動に対する肯定的態度の形成ないし強化が記述から読み取れる。

C 評価：A 評価・B 評価に該当しない。

A 評価の行動計画は、例えば次のような内容である。下線部分が主体的かつ具体的な行動計画であると判断された内容である。

・推奨行動：授業における ICT 活用

「個別最適な学びと協働的な学びを実現していくために、授業の中に ICT 端末をうまく活用しながら授業作りをしていきたいと感じました。自分自身、まだ指導の個別化と学習の個性化について勉強していかなければならないところがあるので、そこを意識しつつ、協働的な学びも大切にしていけるようにしたいと感じました。社会科で、今までは学習のまとめとして紙の新聞を作っていたこともありましたが、そこをロイロノートでまとめて共有するなど、新たな方法を模索していきたいと感じました。」（小学校教諭）

・推奨行動：若手教員への指導助言

「初任者の授業や指導案を拝見して、自分の授業を振り返る契機となった。特に「その学習活動は生徒にどのような資質能力をつけさせたいのか」、を明確にしておくことの重要性を改めて感じた。そのためには単元や授業を組み立てる際の生徒観・教材観をきちんと分析することが不可欠である。＜中略＞また経験年数を重ねてくると、自分の授業のスタイルがよくも悪くも確立してくる。勤務校にも若手教員が複数おられるが、経験や年代の異なる教員の授業を見学し、意見交換することで、新たな視点を導入することができると思われるし、お互いの授業を学び合うことで同僚性の構築にも繋がっていくとも思われる。2 学期以降、このことをより一層積極的に行っていこうと考える。」（高等学校教諭）

・推奨行動：組織的な人材育成

「心理的安全性。多様な個を生み出すのも、協働できる職場環境を生み出すのも、まずは心理的安全性が何より大切だと考えます。OJT 研修だけでなく、研究授業検討会や反省会、各行事の運営などで、職員一人一人が自分の考えている事をちゃんと言える風通しの良さを大事にしていきたい。そうすることで、一人一人の中にある個性が発揮され、多様な個が確保される豊かな集団になると思います。そんな多様な個が揃った集団は、変化に対応できる組織だと思います。なので、それぞれが自分の仕事を抱え込んで進めるのではなくて、普段からちょっとしたことでもコミュニケーションを取ることを心掛けることで、仕事も人間関係も円滑になるようにしていきたい。」（小学校教諭）

レベル 3（自己申告）：追跡調査において「活用した」を選択した者に 1、「活用していない」を選択した者に 0 を割り当てて 2 値データ化した。

レベル 3（記述評価）：追跡調査において「活用した」者から報告された活用内容の記述が当該研修での推奨行動に該当するかどうかについて、教育センターの事業評価業務担当者 2 名が別々に評価を行った（Cohen の $\kappa=0.72$ ）。そして、評価結果に差のあった箇所について、二者間の協議を経て最終的な評価が決定された。実行済と評価された者には 1、未実行と評価された者と「活用していない」者には 0 を割り当てて 2 値データ化した。

なお評価の際、記述内容がレベル 2 の行動計画と対応しているかどうかは考慮されなかった。研修での学びを職場へ転移するための自己調整過程で、実践内容は個人差要因・職場環境要因の影響を受けつつ動的に修正されてゆくことから（Blume *et al.* 2019）、元々の行動計画とは異なる実践となることも多分

に想定されるためである。評価基準は次のとおりである。

実行済：推奨行動を実行したことが記述から明確に読み取れる。
未実行：問2で「活用していない」を選択している、または、問2で「活用した」を選択しているものの実行済に該当しない。

(2) 級内相関係数の算出

以降の分析には R (4.2.2) を用いた。462件のプールドデータは154名の受講者から得られた3回分の測定結果であると同時に(受講者レベル)、12日分の研修について得られたデータでもあり(研修レベル)、2つのクラスタレベルで交差的にネストされた階層構造となっている。このようなマルチレベルデータには、交差分類モデル(cross-classified model)を適用することで推定結果の偏りを制御できると考えられている(Goldstein 1994)。そこでまず、上岡・中藤(2024b)の分析事例と川端(2019)を参考に、lme4パッケージの glmer 関数を用い両変量効果のみを投入した帰無モデルから各クラスタレベルで生じている級内相関係数を算出したところ、表4のようになった。

自己申告のレベル3についてはいずれの変量効果も $ICC > .05$ と清水(2014)の示した第2基準を満たしたが、記述評価のレベル3については研修レベルで $ICC = .037$ と基準を下回った($AIC = 624.7$)。そこで研修レベルを除いた帰無モデルを再度推定したところ適合度の改善が見られたため($AIC = 622.9$)、記述評価のレベル3について分析する際は受講者レベルのみを変量効果として採用することにした。

表4 レベル3の級内相関係数		
クラスタレベル	自己申告	記述評価
受講者	.231	.146
研修	.080	.037

(3) レベル2とレベル3の関連性

最後に、各レベル3を従属変数、レベル2を独立変数とした分析を行った。レベル2はダミー変数化して投入し(A評価ダミー・B評価ダミー)、4段階モデルにおいて正の相関が想定されるレベル1と各校種・各推奨行動のダミー変数を統制変数として投入した。その際、受講者内の変動を探るため、レベル1とレベル2については受講者レベルのグループ内平均で中心化した。

表5 レベル3に対するレベル2の効果						
固定効果	自己申告モデル		記述評価モデル			
	推定値	SE	推定値	SE	推定値	SE
切片	1.92 ***	0.38	-0.15	0.26	-0.19	0.24
中学校ダミー	-0.86 +	0.44	-0.41	0.33	-0.37	0.30
高等学校ダミー	-0.41	0.41	-0.38	0.29	-0.35	0.27
特別支援学校ダミー	-1.08 *	0.49	-0.91 *	0.39	-0.84 *	0.36
GIGAダミー	-0.37	0.32	0.17	0.27	0.28	0.26
B&Sダミー	0.27	0.35	-0.15	0.28	-0.11	0.25
レベル1	0.04	0.06	-0.01	0.06	-0.03	0.05
レベル2(A評価ダミー)	0.58	0.45	1.74 ***	0.38	—	—
レベル2(B評価ダミー)	-0.31	0.36	0.14	0.31	—	—
変量効果	分散	SD	分散	SD	分散	SD
受講者(Intercept)	1.27	1.12	0.50	0.71	0.32	0.57
研修(Intercept)	0.01	0.11	—	—	—	—
Marginal R^2	.07		.10		.03	
Conditional R^2	.33		.22		.12	

*** $p < .001$, * $p < .05$, + $p < .10$

分析の結果、統計的に有意なレベル2の効果は記述評価モデルのA評価ダミーのみであった（推定値=1.74, $p < .001$ ）。一方、自己申告モデルではレベル2からレベル3に統計的に有意な効果は見られなかった（表5）。

次に、記述評価モデルにおけるレベル2の効果量を把握するため、レベル2の両ダミーを除いて推定したところ、固定効果の効果量の差分は $\Delta \text{Marginal } R^2 = .07$ と小程度であった（ $\text{Marginal } R^2 = .10 \rightarrow .03$ ）。

4. 考察

(1) 記述評価の妥当性

教育センターの追跡調査において(1)・(2)のような項目構成がよく見られる理由は、異なる研修間で調査項目をそのまま使い回せることにありと考えられる。都道府県の教育センターは少なくとも年間100日以上研修を行っており、1日の中で複数のテーマが扱われることも稀ではない（上岡 2022）。テーマごとの調査項目のカスタマイズが不要であれば物理的コストを抑え、ヒューマンエラーも防止できる。

レベル2と自己申告のレベル3との間には有意な正の関係が見られなかった一方、予想どおり記述評価のレベル3との間では有意な正の関係が見られた。この結果から、レベル2測定に改訂版キーワード法を用いる場合において、レベル3の指標として妥当であるのは活用内容の評価結果であることが確認された。

同時に、レベル3に対してA評価ダミーが有意な効果を示したことは、研修転移研究の分野で注目されている具体的な行動計画（実行意図: Gollwitzer 1999）の効果を裏付けてもいる（e.g., Friedman and Ronen 2015; Greenan 2023; Quesada-Pallarès *et al.* 2022）。また、改訂版キーワード法において態度変容が読み取れる（B評価）だけでは、その後の行動への影響は大きく限定されることも新たに明らかとなった（推定値=.14, $n.s.$ ）。

今回の実践ではオンライン調査を導入し、レベル3測定上のコストにも対応を試みた。追跡調査は後日行われる研修の当日調査と抱き合わせで実施され、研修中の協力依頼の外には特段回収率向上の取組は行われなかったが、回収率は最も低い②でも74.1%であった。調査機会を減らし、なおかつ効率よく回収できたと言える。

(2) 今後の課題

今回調査対象となった中堅教諭等資質向上研修は、年度内に複数回実施される性質があったため、次の研修日の「研修講座の振り返り」と同時に自然な流れで追跡調査の機会を設けることができた。しかし、研修講座の最終回や1回限りの単発的な研修では、やはりレベル3を測定するだけのために別の調査機会を設ける必要がある。教員が日常使用する校務支援システムの機能に調査フォームを一体化するなど、より少ないコストで追跡調査を行うための仕組みを今後も検討していく必要があるだろう。

今回の評価実践の中では、記述データの評価に要するコストが課題として浮き彫りとなった。オンライン調査につきデータ入力作業が不要であったとはいえ、事業評価業務担当者2名は研修12日分の記述データ（当日・追跡）の評価のためにそれぞれ少なくとも通算1営業日（7時間45分）以上を要した。今後、レベル3測定の対象拡大を図る上では評価作業の効率化が必須と考えられ、AI技術によるアシストは有効な手段の一つになると見込まれる。

本研究で得られた知見は、一般化可能性の面で課題が残る。中堅教諭等資質向上研修は教育公務員特例法が定める法定研修であるため、個別の教育センターが独自に企画する研修と比べれば一般化の余地は

大きいと考えるが、他地域・他職種・異なる経験年数・異なる研修内容に一般化できるかどうかは十分検討する必要がある。また、本研究ではレベル2の測定に改訂版キーワード法を用いたが、その他の手法で測定したレベル2とも同様に正の関係が見られるかについて、今後実践報告が増加することに期待したい。

謝辞

調査に御協力いただいた研修受講者の皆様と、各研修で講師を担当いただいた皆様に厚く御礼申し上げます。

参考文献

- Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, **50**(2): 179-211
- Alliger, G. M., and Janak, E. A. (1989) Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology*, **42**: 331-341
- Alliger, G. M., Tannenbaum, S. I., Bennett, W., Jr., Traver, H., and Shotland, A. (1997) A meta-analysis of the relations among training criteria. *Personnel Psychology*, **50**: 341-358
- Blume, B. D., Ford, J. K., Baldwin, T. T., and Huang, J. L. (2010) Transfer of training: A meta-analytic review. *Journal of Management*, **36**(4): 1065-1105
- Blume, B. D., Ford, J. K., Surface, E. A., and Olenick, J. (2019) A dynamic model of training transfer. *Human Resource Management Review*, **29**(2): 270-283
- 独立行政法人教職員支援機構 (2023) 「研修観の転換」に向けて—NITS と教育委員会等による「新たな教職員研修」の協働開発—. <https://www.youtube.com/watch?v=5uB37bRuQfM> (参照日 2024.01.08)
- Edmondson, A. C. (2018) /野津智子 (訳) (2021) 恐れのない組織—「心理的安全性」が学習・イノベーション・成長をもたらす—. 英治出版
- Friedman, S., and Ronen, S. (2015) The effect of implementation intentions on transfer of training. *European Journal of Social Psychology*, **45**(4): 409-416
- Goldstein, H. (1994) Multilevel cross-classified models. *Sociological Methods & Research*, **22**(3): 364-375
- Gollwitzer, P. M. (1999) Implementation intentions: Strong effects of simple plans. *American Psychologist*, **54**(7): 493-503
- Greenan, P. (2023) The impact of implementation intentions on the transfer of training from a management development program, *Human Resource Development International*, **26**(5): 577-602
- 平松陽一 (2001) 教育研修の効果測定と評価のしかた. インターワーク出版
- 兵庫県立教育研修所 (2009) 研修の事業評価システムの構築に関する研究—研修の有効性を評価する—. 兵庫県立教育研修所研究紀要, 118 https://www.hyogo-c.ed.jp/~kenshu-p-lib/research/119_01.pdf (参照日 2024.01.08)
- 上越教育大学 (2008) 平成19年度文部科学省委託事業「教員研修評価・改善システム開発事業」—新潟県立教育センターにおける Check-Act システムの研究開発—.
- 金子智栄子, 金子智昭, 植草一世, 清水優菜 (2023) ベテラン保育者研修におけるマイクロティーチングの実践事例—混合研究法による園内研修の有効性の検討—. 日本教育工学会論文誌, **47** (1): 63-77
- 金子智昭, 金子智栄子, 植草一世, 金子功一, 清水優菜 (2021) 保育現場におけるマイクロティーチング活用の効果—園内研修を通じた保育者相互の力量形成—. 日本教育工学会論文誌, **44** (4): 483-496
- 川端一光 (2019) マルチレベル一般化線形モデル. 尾崎幸謙, 川端一光, 山田剛史 (編著) Rで学ぶマルチレベルモデル [実践編]—Mplusによる発展的分析—. 朝倉書店, pp.2-22
- Kirkpatrick, D. L. and Kirkpatrick, J. D. (2006) *Evaluating Training Programs: the four levels (3rd ed.)*. Berrett-Koehler
- 小清水貴子, 藤木卓, 室田真男 (2014) 校内における ICT 活用推進を促す教員研修の評価方法の提案と効果の検証. 日本教育工学会論文誌, **38** (2): 135-144

- 小清水貴子, 藤木卓, 室田真男 (2016) ICT 活用推進リーダーを対象にした集合研修の改善と評価. 日本教育工学会論文誌, **40** (2): 113-126
- 文部科学省 (2021) 「令和の日本型学校教育」を担う新たな教師の学びの姿の実現に向けて—審議まとめ—. https://www.mext.go.jp/content/20211124-mxt_kyoikujinzai02-000019122_1.pdf (参照日 2024.01.08)
- 文部科学省 (2022) 「令和の日本型学校教育」を担う教師の養成・採用・研修等の在り方について—「新たな教師の学びの姿」の実現と, 多様な専門性を有する質の高い教職員集団の形成 (答申) —. https://www.mext.go.jp/content/20221219-mxt_kyoikujinzai01-1412985_00004-1.pdf (参照日 2024.03.04)
- 文部科学省 (2023a) 「研修受講履歴記録システム及び教員研修プラットフォーム」の構築について. https://www.mext.go.jp/kaigisiryō/content/20231204-mxt_kyoikujinzai01-000032933_03.pdf (参照日 2024.01.08)
- 文部科学省 (2023b) 教員研修高度化支援—教員研修の高度化に資するモデル開発事業実施要項—. https://www.mext.go.jp/content/20230303-mxt_kyoikujinzai01_000027902-12.pdf (参照日 2024.01.08)
- 内閣府 (2022) Society 5.0の実現に向けた教育・人材育成に関する政策パッケージ (案) —CSTI 教育・人材育成 WG 最終とりまとめ—. https://www8.cao.go.jp/cstp/tougosenryaku/11kai/siryō3_3_1.pdf (参照日 2024.01.08)
- 中原淳, 関根雅泰, 島村公俊, 林博之 (2022) 研修開発入門—「研修評価」の教科書—. ダイヤモンド社
- 奈良県立教育研究所 (2009) 研修の効果測定—研修講座の工夫・改善に向けて—. 平成21年度研究紀要: 1-11
- 岡山県総合教育センター (2015) B&S 研修とは—初任者研修と10年経験者研修の合同実施による意図的・計画的な「出会い」—. 中等教育資料, **64**(11): 78-81
- 大阪府教育センター (2020) 研修評価についてのアンケート結果概要. https://www.osaka-c.ed.jp/category/forteacher/investigate/publication/r02/R02_kikaku_summary.pdf (参照日 2024.01.08)
- Quesada-Pallarès, C., González-Ortiz-de-Zárate, A., Pineda-Herrero, P., and Cascallar, E. (2022) Intention to transfer and transfer following eLearning in Spain. *Vocations and Learning*, **15**(2): 359-385
- Saks, A. M., and Burke, L. A. (2012) An investigation into the relationship between training evaluation and the transfer of training. *International Journal of Training and Development*, **16**(2): 118-127
- 清水裕士 (2014) 個人と集団のマルチレベル分析. ナカニシヤ出版
- 徳村朝昭, 前川朝康, 鈴木克明 (2006) ICT スキル研修改善のためのアクションプラン導入手法. 日本教育工学会論文誌, **30**(3): 203-212
- 上岡伸 (2022) 改訂版キーワード法による研修効果測定の妥当性に関する検討. 日本教育工学会論文誌, **46**(2): 289-302
- 上岡伸 (2023) 初任教員の研修転移プロセスにおける行動意図の役割. 日本教育工学会論文誌, **47**(2): 229-236
- 上岡伸, 中藤路子 (2024a) 学校事務職員の自己調整学習を促進する研修プログラムの開発. 日本教育工学会論文誌, 早期公開版, DOI: <https://doi.org/10.15077/jjet.47087>
- 上岡伸, 中藤路子 (2024b) ハイフレックス型研修の学習効果は受講形態によって差があるか. 日本教育工学会論文誌, 早期公開版, DOI: <https://doi.org/10.15077/jjet.47090>
- 全国教育研究所連盟 (2007) 平成18年度教育課題調査結果. <https://www.nier.go.jp/zenkyou/zenkyou/chousa/06chousahokoku.pdf> (参照日 2024.01.08)

令和6年3月発行

岡山県総合教育センター 研究紀要 第14号 研究番号 23-04

「教員研修の活用報告に基づく行動変容度合の定量化」