

非認知能力と学級風土に焦点化したエビデンスにもとづく教育実践

—小学3年生の教室での試み—

* 田 端 健 人・**安倍 彰 人

要 旨

本研究は、教師と研究者との協働による「エビデンスにもとづく教育実践」の試みである。本研究実践では非認知能力と学級風土、またそれらと相関する対話・探究学習に焦点化した。第2執筆者が担任を務める小学3年生を対象に、まず年度はじめの4月に第2執筆者が質問紙調査を実施し、第1執筆者が分析し、結果を共有した。第2執筆者は、その分析結果を受けとめ、その後の自身の実践に生かした。7月に第2執筆者が同じ質問紙調査を実施し、4月と比較した。その結果、非認知「徳」と学級風土のスコアは高水準を維持し、対話・探究学習スコアに向上がみられた。対話・探究学習に積極的に取り組んだ成果が、非認知「徳」と学級風土のスコアの水準維持に寄与したと推測される。本稿では、こうしたエビデンスにもとづく教育実践の試みを、実践者の視点で評価し、その意義と課題を指摘した。最後に「実践知とデータ知の往還」「教室レベルと全国レベルとの往還」「観測者問題」の3つの視点から、本研究実践を省察した。

Key words : 教育データサイエンス、EBPM、可視化された学習、評価、質問紙調査

1. 問題背景とリサーチクエスチョン

数量的エビデンスは、使いようによっては、子どもや教師を委縮させ、差別や序列や競争の温床となる。一例をあげるなら、全国学力・学習状況調査の都道府県ランキングや全国平均正答率とのポイント差による評価である。都道府県の平均差に意味があるのか、何ポイント差なら有意なのか不明のまま、順位やポイント差が独り歩きし、現場の教育施策・実践に混乱をもたらしている。

一方、数量的エビデンスは、使いようによっては、子どもや教師に成長の手応えを与え、新しい課題への挑戦心に火をつけることもある。エビデンスのこうした利活用こそが、「エビデンスにもとづく政策決定 (Evidence Based Policy Making:EBPM)」と教育データサイエンスの歩むべき道であろう。

そのための必要条件は、EBPM にいたるまでのデータの収集・分析・評価が、特定の専門家に独占されな

いことである。それが一部の専門家の専売特許となるなら、教師や子どもは一方的にデータを搾取される情報源に貶められ、分析・評価される客体へと隷属化されてしまう。実践の当事者が主体となるデータサイエンスこそが、当事者を自由で自律的にする (cf., 田端 b, 2022, p.53)。非専門家にも意味がわかる統計による教育データサイエンス、「教師の、教師による、教師のためのデータサイエンス」が不可欠である。

では、教師が数量的エビデンスを利活用し、自身の指導に手応えを感じながら、指導力を向上させたり、指導を効果的にしたりする教育実践の具体とは、どのような姿だろうか。これが本研究のリサーチクエスチョンである。

2. 本実践研究の教育理論的位置づけ

このリサーチクエスチョンは、ハッティ (John Hattie) の「可視化された学習 (Visible Learning)」

* 宮城教育大学教職大学院・教授

** 宮城教育大学附属小学校・教諭

研究と連動している。この研究から導き出される「教育効果を高めるマインドフレーム」の第一は、「私〔教師やスクールリーダー〕は生徒の学習に及ぼす影響の評価者である」(ハッティほか, 2021, p.16,〔〕内引用者)である。このマインドフレームでは、教師が評価の主体になっている。教師が主体となり、子どもへの自分たちの働きかけの影響(インパクト)を、複数のエビデンスで評価するのが、教育効果を高める第一のマインドフレームである。その評価をもとに、効果的でない働きかけをやめたり、効果をいっそう高める働きかけにチャレンジしたりといった他のマインドフレームが続く。評価のためのエビデンスは、「データ、教師や生徒の声、経験、授業成果物や観察など」(ハッティほか, 2022, p.16)多種多様である。ただ、ハッティの「可視化された学習」研究の大黒柱が「効果量(Effect Size)」であることからすれば、教育効果を高めるマインドフレームには、数量的エビデンスが必須である。

そこで、本実践研究を「可視化された学習」研究の一環と位置づけ、第一のマインドフレームを出発点とする。そうすると次に検討すべきは、生徒に対する教師の様々な「インパクト(影響)」の何に注目し、それをどのように測定・評価するかである。

3. 教師のインパクトの焦点化

生徒の何に対するインパクトを測定・評価するかについては、実践者である第2執筆者の所属校の特色や課題から特定するのが良いと考えた。

第2執筆者の所属校は、「学力」に関しては、全国平均よりも高い成果を取っている。これについては、全国学力・学習状況調査や標準学力調査など数量的エビデンスが豊富である。ところが、それと対をなす「非認知能力」については、所属校には数量的エビデンスが欠けている。そこで本研究では、非認知能力に焦点を絞り、児童の非認知能力に与える教師のインパクトを測定・評価することにした。

また第2執筆者が教師としてその醸成に関心があり、学力や非認知能力とも相関するとされる「学級風土」も焦点化することにした。

4. 測定方法

測定方法は、一般によく利用される質問紙調査(Google フォームによる児童へのアンケート調査)とした。

次に質問項目の選定・作成である。

本研究の問題意識から、質問項目にはいくつかの前提条件がある。

- ① 質問項目が教師と子どもを傷つけないこと(いわゆる「非侵襲性」)。子どもが回答するとき、あるいは教師が回答結果を見るとき、双方に負の感情やストレスを与えないこと。
- ② 質問項目数が教師と子どもの負担にならないこと。項目数は必要最少限で、かつ実用的な精度を保つこと。
- ③ 一般的な目安を知るため、特定の年齢や学年で項目スコアの平均値と標準偏差を知ることができること。
- ④ 将来的な汎用性を見据え、小学生から中高生までをカバーできる項目であること。

本実践研究では、これらの条件をある程度クリアする大規模調査の質問項目を参照することにした。具体的には、全国学力・学習状況調査、埼玉県など自治体独自IRT学力・学習状況調査、国際学力調査PISAおよびTIMSSを参照した。なお、これらの中には、スコアの平均値と標準偏差が必ずしも公開されていないものもあるが、公開あるいは貸与データから独自に算出可能である。

5. 質問項目の作成

これらの調査項目を検討した結果、非認知能力については先行研究に依拠し、表1の8項目からなる「非認知『徳』尺度」(cf., 田端, 2023, p.114)を利用することにした。この合成変数については、小学6年生

表1：非認知「徳」尺度の質問項目

(1) 自分には、よいところがあると思いますか
(2) 将来の夢や目標をもっていますか
(3) 難しいことでも、失敗を恐れなくて挑戦していますか
(4) 人が困っているときは、進んで助けていますか
(5) いじめは、どんな理由があってもいけないことだと思いますか
(6) 人の役に立つ人間になりたいと思いますか
(7) 学校に行くのは楽しいと思いますか
(8) 地域や社会をよくするために何をすべきか考えることがありますか

(以下「小6」と中学3年生(以下「中3」)の全国平均と標準偏差が、先行研究で算出されている。

学級風土については、宮城教育大学教職大学院の教員と現職派遣院生の有志7名で協議し、主に埼玉県IRT学力・学習状況調査¹の学級風土に関わる項目群をもとに、独自の表現に改めたり、新たに追加したりした。それをさらに、第2執筆者との協議により、加除修正し、表2の15項目とした。

表2：学級風土尺度

(1) 学級での生活は楽しいですか。
(2) あなたの学級は落ち着いて学習できる様子ですか。
(3) あなたの学級は、いろいろな活動にまともに取り組んでいると思いますか。(運動会や遠足などの学校行事も入ります)
(4) 自分のよいところをみとめてくれる先生はいますか。
(5) 自分のよいところをみとめてくれる友達はいますか。
(6) 自分のなやみや相談にのってくれる先生はいますか。
(7) 学級での勉強は楽しいですか。
(8) あなたの学級は、わからないことを質問しやすいですか。
(9) あなたの学級は、安心してまちがったことも言える学級ですか。
(10) あなたの提案や意見は、学級で生かされそうですか。
(11) 自分ではできないことも、学級の友達や先生とならでできることがふえると思いますか。
(12) この学級では、だれとでも協力して活動できると思いますか。
(13) 周りの人はこまったときに助けてくれますか。
(14) 学校では自分の良いところを生かしていると思いますか。
(15) みんなと意見がちがっても自分の意見を出せますか。

さらに先行研究から「対話・探究学習尺度」(田端, 2023, p.117)の6項目も質問することにした(表3)。この合成変数は、非認知「徳」変数と強い相関があり、学力値とも弱い相関があることが既知だからである(cf., 田端, 2023, p.121)。

表3：対話・探究学習尺度

(1) 2年生のときに受けた授業で、自分の考えを発表する機会では、自分の考えがうまく伝わるよう、資料や文章、話の組立てなどを工夫して発表していましたか
(2) 2年生のときに受けた授業では、課題の解決に向けて、自分で考え、自分から取り組んでいましたか
(3) 学級の友達との間で話し合う活動を通じて、自分の考えを深めたり、広げたりすることができていますか
(4) 総合的な学習の時間では、自分で課題を立てて情報を集め整理して、調べたことを発表するなどの学習活動に取り組んでいますか
(5) あなたの学級では、学級生活をよりよくするために学級会で話し合い、互いの意見のよさを生かして解決方法を決めていますか
(6) 学級活動における学級の話し合いを生かして、今、自分が努力すべきことを決めて取り組んでいますか

なお7月調査では、「2年生のときに」を「3年生になって」に改めた。

全てを合わせるとかなりの項目数になるため、児童への負担や学習時間の圧迫を懸念したが、実践者である第2執筆者が許容範囲と判断し、実施することにした。

実施は学期はじめの4月と夏休み前の7月を計画した。

6. 4月調査の結果と評価

6-1. 非認知「徳」スコア

4月調査の回答で非認知「徳」スコアを算出すると、当該学級は平均27.04、標準偏差3.30であった。先行研究によると小6の全国平均は26.17、標準偏差3.60である(田端, 2023, p.116)。また先行研究によると、この合成変数の平均と標準偏差は、小6と中3でほとんど違いがない。とはいえ小3と小6で全国の傾向として差がないというエビデンスは今のところない。そこであくまで便宜的な目安として、小6全国平均と比較してみるなら、当該小3学級の非認知「徳」スコアはその目安より高いことになる。この平均値差を効果量(Cohenのd)で見積もると、 $d=0.25$ になる²。私たちはこの差を「それなりの差」と評価し、この学級の児童たちの前年度、つまり小2のときの担任たちが児童の非認知「徳」能力に与えた影響(インパクト)の大きさを再確認した。

またこの学級集団の非認知「徳」スコアの水準を小3でも維持することを目標とした。

6-2. 対話・探究学習スコア

当該小3の対話・探究学習スコアの平均は16.58、標準偏差2.89であった。全国小6平均は17.91、標準偏差3.65である(田端, 2013, p.117)。これも全国小6を便宜的な目安とするならば、当該小3クラス平均は目安より低いことになる。効果量で平均値差の大きさを見積もったところ、 $d=-0.40$ となり、小さいながら差がある(全国より低い)結果となった。

先行研究によると、令和4年度全国小6データで、対話・探究学習スコアと非認知「徳」スコアと

1 埼玉県学力・学習状況調査の質問紙調査項目は全て公開されている。例えば、令和5年度に関しては、次のウェブサイトの「質問調査」から、小学4年生～中学3年生までの質問調査項目にアクセスできる。

<https://www.pref.saitama.lg.jp/f2214/gakutyou/r5kenngakutyou.html> [2023.09.23最終閲覧]

2 効果量の計算には、下記ウェブサイトで公開されている、DS-EFA(田端健人・菅原敏)開発の「平均値差検定システム(AVES)」を利用した。<https://ds-efa.info/cohensd/> [2023.09.24最終閲覧]

の相関係数は0.63であり、中程度の相関がある(田端,2023,p.121)。そこで、今後対話・探究学習を活性化させることで、非認知「徳」スコアの高水準を維持することを、第2執筆者との共通認識とした。

6-3. 学級風土スコア

この項目群は独自開発のため、これら15項目を束ねて合成変数にできるか否かは未知である。そこで、当該学級29名の回答傾向から、信頼係数(クロンバックの α)を算出した。調査目的であるため $\alpha \geq 0.70$ を基準値とする。結果は $\alpha = 0.75$ となり、基準値を超えたため15項目の内的整合性が一定程度担保された。

当該学級の平均は45.90、標準偏差6.05であった。これについては目安となる基準値が未知なので、この平均値が高いのか低いのか、4月段階では評価できなかった。目標としてはこの値を下げないこととした。

この項目群でとりわけ意義深かったのは、項目6「自分のなやみや相談にのってくれる先生はいますか」であり、「いない」と回答した児童が5名いたことである。児童が悩んだり困ったりしたときに相談できる教師がいなくて児童自身が感じていることは、改善を要する教育課題である。この5名については特に配慮し、学級全員が教師に悩みを相談できるようになることを目標とした。また、5名のうち2名は、担任である第2執筆者にとって意外であった。回答時の気分が左右されたのかもしれないが、教師の実践感覚の見逃しの可能性も否定できない。この数量的エビデンスを知ること、担任の第2執筆者は、この2名については特に注意して働きかけることにした。

7. 4～7月に力を入れた取組

以上の数量的エビデンスの結果と評価を受け、担任である第2執筆者は、対話・探究学習にいつそう積極的に取り組んだ。具体的には、子どもの哲学(philosophy for/with children: p4c)を主に道徳科で取り入れた。p4cは児童にも好評で「やりたい」との声が多く、7月までに3回実施した。第1執筆者もファシリテーターとして同クラスで1回p4cを実施した。p4cは子どもたちが立てた問いを、子どもたちが話し合いによって探究する授業である。通常の授業は教師の発話が多くなりがちであるが、p4cでは子どもの発話が圧倒的に多くなり、教師の発話は少なくな

る。子どもの聞く対象も、教師の発話ではなく、子どもの発話にシフトする。

道徳科以外の教科学習でも、ペア学習やグループワークなど話し合い活動を積極的に実施した。

所属校は近年、探究型の学習にシフトしている。ただ、自分で課題を設定したりするのが苦手な児童もあり、一長一短との感触を第2執筆者は抱いている。

非認知面では、チャレンジを重視し、「失敗してもよい」と頻りに声をかけた。また「自分で問いをもつ」ような指導を心がけた。

8. 7月調査の結果と評価—4月調査との比較—

夏休みまでに1名の児童が転出した。そこで4月と7月調査との比較では、この1名を除外して分析する。

8-1. 非認知「徳」スコア

4月と7月の非認知「徳」スコアの結果は、表4の通りである。

表4：非認知「徳」スコア(4月と7月)

非認知徳スコア	4月	7月
平均値	26.82	26.32
標準偏差	3.15	4.30
中央値	27	27.5

平均値はわずかに下がったが、中央値はわずかに上がる結果となった。平均値差を効果量で見積もると $d = -0.13$ となり、さほどの下降ではないと判断し、学年はじめの高水準が維持されたと評価した。

データを箱ひげ図で可視化すると図1になる。左が

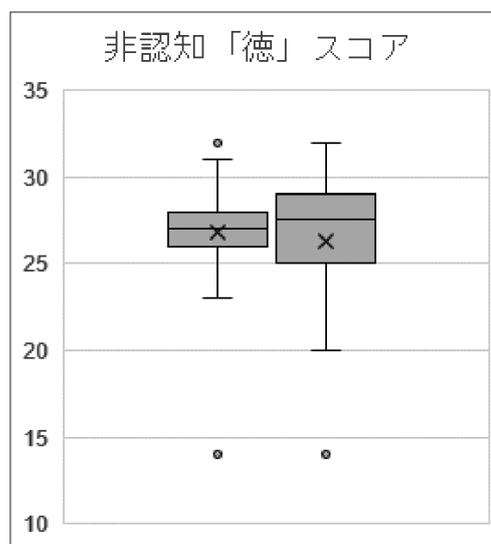


図1：非認知「徳」スコア(4月と7月)

4月、右が7月である（以下同様）。

図1からばらつきが大きくなったことがわかる。また4月では平均値（×印）が中央値（箱の分割線）とほぼ重なっていたのに対し、7月では中央値よりも下に移行したことから、平均以上の割合が増加したこともわかる。7月調査では、平均以下は全28名中8名（29%）に減少している。

ただ7月調査では外れ値（14ポイント）の児童が2名おり、今後のいっそうの配慮が必要であることを共通認識とした。

8-2. 学級風土スコア

学級風土スコアの平均値も4月と7月とでほとんど差がなかった。結果を箱ひげ図にすると図2になる。

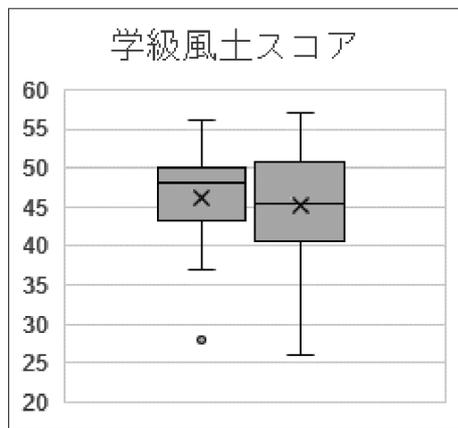


図2：学級風土スコア（4月と7月）

平均値と中央値の関係から、平均以下の児童割合が7月調査で増加していることが分かる。ばらつきも大きくなっている。

非認知「徳」スコアで外れ値となった児童2名（児童Aと児童B）と、学級風土スコアの最低ポイントの2名とは同じであった。この2名は特別な配慮を要する児童たちであり、担任の第2執筆者も配慮と働きかけに尽力してきたが、7月の調査ではその成果が数値には表れなかった。

第1執筆者の参与観察からすれば、学級の雰囲気や人間関係はきわめて良好であり、この平均値は比較的高いのではないかとの感触を得ている。

学級風土に関する一つの成果は、項目6「自分のなやみや相談のしてくれる先生はいますか」であり、「いない」と回答した児童が4月は5名だったのに対し、7月では2名（児童Aと児童C）に減少した。これは担任教師の日々の実践の成果と評価できる。

もちろんこの2名については、さらなる配慮が必要である。特に児童Aは非認知「徳」スコアと学級風土スコアの両方で低いスコアをつけている。児童Aへの支援や指導の難しさが数値にも表れている。

肯定的な割合が増加した項目は、表2の1、3、5、6、7、8、9、10、13であった。「学級での生活の楽しさ」（項目1）や「学級での勉強の楽しさ」（項目7）の質問で、肯定的な割合が増加したことは、うれしい結果である。「自分のよいところを認めてくれる友達がいる」（項目5）や「周りの人が困ったときに助けてくれる」（項目13）の肯定的回答割合の増加からも、学級全体でみるなら、7月までの実践がお互いを尊重し助け合う人間関係を醸成していることがうかがわれる。また「学級では安心して間違ったことも言える」の肯定的回答割合が6割強から8割強へと上昇し、「当てはまらない」の回答がゼロになったことも、大きな成果である。対話・探究学習の効果と推測される。

8-3. 対話・探究学習スコア

対話・探究学習スコアの平均値には上昇がみられた。箱ひげ図で可視化すると図3になる。

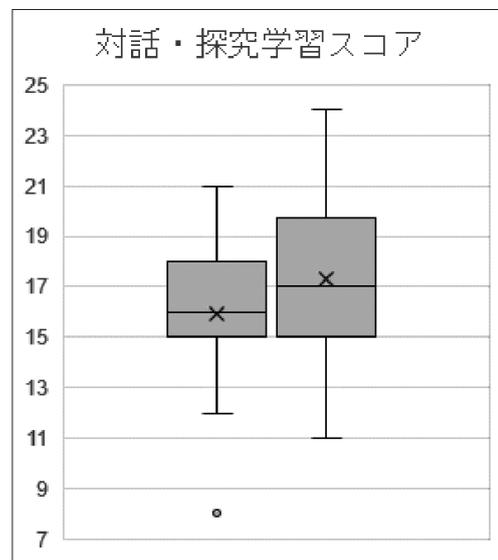


図3：対話・探究学習スコア（4月と7月）

平均値差を効果量で見積ると、 $d=0.40$ となり、確かな上昇があったと評価できる。担任教師が積極的に取り組んだ成果がそのまま数値に表れた。

ただ、この合成変数の7月平均は17.29であり、これでもまだ全国小6平均にはわずかに届かない。今後の継続ないしはさらなる充実を、私たちの共通目標とした。

対話・探究学習スコアと非認知「徳」スコアとの相関の強さの先行研究結果からすると、当初の予測通り、対話・探究学習に積極的に取り組んだことが、非認知「徳」スコアの高水準を下支えしたと考えられる。

スコアが最も低かった4名(11ポイントと13ポイント)のうち3名は、上記の児童Aと児童Bと児童Cである。この3名に充実感や満足感を与える対話・探究学習を工夫することが、これら3名の非認知「徳」スコアや学級風土スコアの向上につながることも推測できる。

9.3 合成変数の相関関係

この学級の児童の場合、3つの合成変数の相関関係はどの程度だろうか。転校生を除く28名のデータで相関係数を出してみた。統計ソフトRのpsychパッケージのpairs.panels関数を利用した。略記号として、非認知「徳」スコアに「NonC」、学級風土スコアに「CICl」、対話・探究学習スコアに「Dia」を用いた。

4月の結果は図4になる。

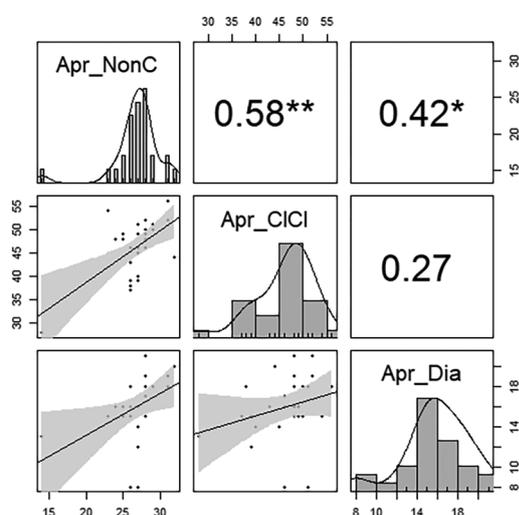


図4：4月の3合成変数の相関係数

先行研究によれば、令和4年度全国小6の非認知「徳」スコアと対話探究スコアの相関係数は、0.63である(田端,2023,p.121)。この学級では $r=0.42$ となり、全国小6より低い相関である。これは4月時点で、この学級の児童たちが対話・探究学習を前年度あまりしてこなかったと感じていたことと連動していると推測できる。つまりこの児童たちの場合、非認知「徳」スコアの高低には、対話・探究学習以外の要素が影響し

ていたと推測できる。

7月の結果は図5になる。

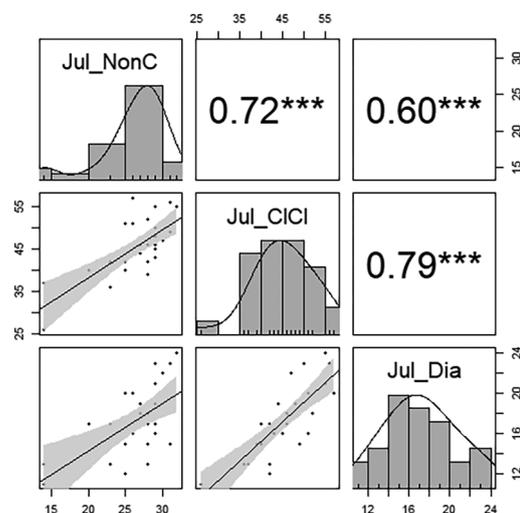


図5：7月の3合成変数の相関係数

3つの合成変数の相関係数がすべて4月よりも高くなっている。非認知「徳」スコアと対話・探究学習スコアとの相関係数も全国小6とほぼ同じになっている。対話・探究学習をしていると感じている児童ほど、非認知「徳」スコアが高いという結果である。同様に、対話・探究学習をしていると感じている児童ほど、学級風土がよいと感じている。さらに、学級風土が良いと感じている児童ほど、非認知「徳」スコアが高いという結果である。

4月では、対話・探究学習と学級風土との間に弱い相関($r=0.27$)しかなかったが、7月にはかなり強い相関($r=0.79$)になり、0.1%水準で有意(3つのアスタリスク)と判定されている。4月から7月にかけて、この学級では、対話・探究学習を中心として学級風土が作られ、非認知「徳」能力が育成されたと推察される。

10. 「エビデンスにもとづく教育実践」の評価—実践者の視点から—

本章では、「エビデンスにもとづく教育実践」に取り組んだ教師(第2執筆者)の視点から、データサイエンスを取り入れる負担感や意義を振り返りたい。

10-1. 質問紙調査の負担感

当該学級では、本稿で触れなかった項目群(7項目)を合わせた4つの調査を同時に実施した。4つの項目群につき、各項目を担当教師が児童に説明しながら実

施したため、1単位時間(45分)はかからなかったものの、それなりの時間が必要であった。ただ、2回目の時間は短縮された。第2執筆者としては、大きな負担感を感じなかった。だが以前勤めていた小学校の同僚はこうしたアンケート調査に負担感を抱いていたため、負担感は教師によって異なるだろう。

また負担感、調査にかかる手間暇と調査の意義・効果との綱引き、ないし両ベクトルの力の合成と考えられる。調査の意義が大きく、それに見合う効果があるなら、負担感も減少・相殺されるだろう。

実践した第2執筆者には、調査実施の頻度と時期については、妥当と感じられた。年度はじめ(4月)、夏休み前(7月)、1学期終わり(10月)、年度末(2月末頃)が適当ではないだろうか。

10-2. 質問紙調査の意義や効果

児童にとっては、自分が今どのように感じているかを振り返る機会になったと感じる。一種のメタ認知の育成にもつながるかもしれない。

今後の課題としては、回答した児童たちが数値の変容を知る必要はないかの検討が必要である。調査が教師のためだけでなく、児童のためにもなることが望ましい。

教師としては、児童たちがどのように感じているかを知ることには怖さも多少感じるが、児童たちの成長、ひいては教師自身の成長のためには必要であると考えられる。普段の教師の見とりと、児童の抱く感想とにどのような乖離があるかを知ることが、教師にとって非常に重要である。

昨今のいじめ問題を受け、すべての学校で「いじめアンケート」を実施しているが、今回の質問紙調査はいじめ予兆の早期発見にもつながる有効な調査と感じた。

4月調査の結果から、個別に配慮できた児童もいた。成長が数値に表れた児童もいれば、表れなかった児童もいる。支援の方法をさらに再検討し、いっそう効果的な指導と支援に努めたい。今回の調査は、特別支援コーディネーターを交えた支援の検討にもつなげやすくなるのではないだろうか。

10-3. 今後の調査のいっそうの充実に向けて

こうした調査を学年や学校全体での取組につなげるためには、項目の簡素化ないしさらなる焦点化が望ましい。ただ測定精度との兼ね合いもある。例えば、

4つの項目群を3つの項目群にするだけで、それなりの簡素化になる。独自開発の学級風土の質問項目は、さらに精選できるであろう。

また学年や学校全体に広げるためには、調査の意義、データやグラフの見方など教員研修が必要になる。その上で、調査実施後に結果を「どのように活用していくか」を教師個人と教師集団で明確化し、教師集団で共通理解していくことも重要である。

11. むすびにかえて

今回のエビデンスにもとづく教育実践を、次の3つの視点から省察することで、むすびにかえたい。

11-1. 実践知とデータ知との往還

実践知とデータ知を往還させることで、実践知の見逃しを補うことができる。その典型が学級風土アンケートの項目6である。この項目のおかげで、相談できる教師がいないと回答した5名を私たちは特定することができ、担任教師が特に注意することで、7月にはうち3名が相談できる教師がいると感じるようになった。同じアンケートを時間をおいて複数回実施することで、私たちは子どもたちの心の変化を把握でき、自分たちの働きかけの効果と課題を明確化できる。項目6のような子どもの「心」に関することを、観察によって見分けることは困難である。これは質問紙調査の強みであろう。

11-2. 教室レベルと全国レベルの往還

今回のデータ分析では一部、教室レベルと全国レベルを往還させた。非認知「徳」スコアや対話・探究学習スコアを、全国小6の平均値と標準偏差で見積もった部分である。教室の子どもたちを私たちは感知することができる。対して、全国の1学年約100万人の子どもたちを、私たちは感知することができない。両レベルを往還させることで、私たちは感知不可能な数値を、実感で肉付けすることができる。例えば、目の前の子どもたちの非認知「徳」スコアが全国平均よりも高ければ、全国平均値というのは、目の前の子どもほど育っていない状態であるとイメージできるようになる。逆に、全国平均値を目安とすることで、目の前の子どもたちの非認知「徳」能力の育ち具合も相対的にわかってくる。これも実践知とデータ知の豊かで創造的な編み合わせである。

今後、教室レベルと学校レベル、全国ひいては世界レベルとを往還させる実践研究がいつそう盛んになることが望ましい。例えばPISAやTIMSSといった国際学力調査では、子どもの学習意欲や学習環境やウェルビーイングなどに関する質問項目があり、私たちは国ごとの数値を知ることができる。こうしたデータを利活用すれば、私たちは日々の教室を、世界を目安として評価できる。

11-3. 観測者問題

観測や観察に関わる学術領域では「観測者問題」という現象が知られている (cf., 田端a, 2022, pp.i-ii, p.89)。観測される対象が、観測するツールや行為に影響を受ける、という現象である。観測対象が子どもであり、観測者が教師である場合には、なおさらである。「可視化された学習」に位置づく「エビデンスにもとづく教育実践」で鍵となるのは、観測ツールや観測行為の影響を、子どもと教師にとって好ましいものにする事である。そのためには本稿第1章で指摘したように、教師自身が観測の主体にならなければならない。自身と子どものためを思って、自身が観測したいと思う内容について、教師が積極的に観測しようと思えば、結果を自身の実践に生かそうとしてはじめて、エビデンスが教師と子どもの力になる。

観測ツール一つとっても、そこには観測者の仮説や思いや願いが込められている。例えば「自分には、よいところがあると思いませんか」の項目一つをとっても、そう思える子どもに育ててほしいという願いがこもっている。教師がしたい・知りたいと思わない質問はすべきではなく、教師がしたい・知りたいと思う質問をすべきである。本研究実践を参考にしていただく際には、質問項目を各実践者が吟味し、自身の思いや願いにいつそうふさわしいものにブラッシュアップしてほしい。

本稿は「可視化された学習」に位置づく「エビデンスにもとづく教育実践」の小さな試みである。こうした研究実践は、さほどの負担感なしにできる。本稿が4月と7月のわずか2回の調査だけに焦点化したのも、可視化された学習が日々の実践で容易に実現できることを示すためである。

引用文献

- ハッテイ, J., スミス, R., 2022『スクールリーダーのための教育効果を高めるマインドフレームー可視化された学校づくりの10の秘訣ー』原田信之訳者代表, 北大路書房.
- ハッテイ, J., ティーラー, C., 2021『教師のための教育効果を高めるマインドフレームー可視化された授業づくりの10の秘訣ー』原田信之訳者代表, 北大路書房.
- 田端健人, 2023『『教育の現象学』のデータサイエンス的転回ー全国学力・学習状況調査結果の分析からー』『学ぶと教えるの現象学研究』20号, パイディア出版, 64-130.
- 田端健人, 2022a『IRT分析ソフト EasyEstimation による全国学力・学習状況調査の検証と経年比較』パイディア出版.
- 田端健人, 2022b『教育データサイエンスと教師の自律』日本教育方法学会編『教育方法51 教師の自律性と教育方法』図書文化, 52-64.

【付記1】第1執筆者は「エビデンスにもとづく教育実践」のプランと測定方法を提案し、第2執筆者がそれを実践した。データ分析・可視化・評価は第1執筆者が行った。第1執筆者は第1～6、8、9、11章を、第2執筆者は第7、10章を執筆した。

【付記2】本研究は、科学研究費助成事業、基盤研究B「学力／非認知能力を効果的に育成する教育リーダーのデータサイエンス」(2023-2025年度、課題番号:23H00921、研究代表者:田端健人)の研究成果の一部である。

(令和6年1月29日受理)

Evidence Based Education Focused on Non-cognitive Skills and Classroom Climates

—The Trial in a 3rd Grade Elementary School Class—

TABATA Taketo and ABE Akito

Abstract:

This research is an attempt at “evidence-based educational practice” led by teachers. In this practice, we focused on noncognitive abilities and classroom climate, as well as dialogue and inquiry learning that correlate with these. First, a questionnaire survey was conducted in April at the beginning of the academic year, targeting third-grade elementary school students, whose class teacher was the second author. The second author recognized the results of the analysis and applied them to his own practice. The same questionnaire survey was conducted in July and compared with April. As a result, scores for noncognitive “virtues” and classroom climate remained at high levels, and scores for dialogue and inquiry learning improved. It is presumed that the results of active engagement in dialogue and inquiry learning contributed to maintaining the level of scores for non-cognitive “virtue” and classroom climate. In this paper, we evaluated these evidence-based educational practices from the teacher's perspective and pointed out their significance and challenges. Finally, we reflected on the implementation of this research from three perspectives: “the back and forth between practical knowledge and data knowledge”, “the back and forth between the classroom level and the national level”, and “the observer problem”.

Key Words : Data Science of Education, EBPM(Evidence Based Policy Making),
Visible Learning, Evaluation, Questionnaire Survey

