

予防教育科学に基づく「子どもの健康と適応」のためのユニバーサル 予防教育における評価のあり方

Evaluation in Universal Preventive Education for Children's Health and Adjustment Based on the Science of Preventive Education

山崎 勝之^{*,**}, 佐々木 恵^{**}, 内田香奈子^{**}, 勝間 理沙^{**}, 松本 有貴^{**}

^{*}鳴門教育大学大学院人間形成コース

^{**}鳴門教育大学予防教育科学教育研究センター

Katsuyuki Yamasaki^{*,**}, Megumi Sasaki^{**}, Kanako Uchida^{**}, Lisa Katsuma^{**}, Yuki Matsumoto^{**}

^{*} Department of Human Development, Naruto University of Education,

^{**} Center for Education and Research on the Science of Preventive Education, Naruto University of Education

抄録：TOP SELF と呼ばれる，子どもの健康と適応のための新しい革新的なユニバーサル予防教育が開発され，学校で実践されつつある。この教育は，学校のすべての子どもたちに恒常的に実施されることを目指している。本論文は，この予防教育における評価のあり方について考察し，TOP SELF における試行段階の評価を紹介することを目的とした。まず，評価の3つのタイプ，科学的評価，グループやクラス全体の評価，そして子ども個人への評価が詳細に説明された後，TOP SELF のオリジナルな評価方法が紹介された。TOP SELF では，教育による子どもの変化について正の側面を強調することが強調され，評価はあくまでも教育であることが示唆された。さらに，教育効果を評価する道具が高い信頼性と妥当性を持つべきであることが強調され，最後に，望ましい評価への今後の課題が，子どもを伸ばす効果的な評価が重要であることを強調しながら論議された。

キーワード：評価，ユニバーサル予防，予防教育科学，健康／適応，児童・生徒

Abstract : A new innovative universal preventive education for children's health and adjustment, named "TOP SELF (Trial Of Prevention School Education for Life and Friendship)", is being developed and implemented in schools. This education aims to be implemented on a regular basis in all children in schools. However, in TOP SELF, methods for evaluation have not been developed well. So, the purpose of the present paper was to discuss about how the evaluation for the preventive education should be and to introduce the tentative evaluation methods in TOP SELF. First, three types of evaluation were introduced: scientific evaluation for effectiveness of education, evaluation for groups or whole classes, and evaluation returned to each child. After it was underscored that scientific evaluation for effectiveness of education is difficult albeit important, TOP SELF's original evaluation methods were explained. Specifically, TOP SELF emphasizes the positive evaluation that focuses on positive phases of children's changes by education, suggesting that evaluation should be education for children. Moreover, it was indicated that tools to assess educational effects should have reliable and valid characteristics, although validity is almost impossible to be fully examined. Finally, future tasks for desirable evaluation are discussed, underscoring the importance of effective evaluation for children.

Keywords : evaluation, universal prevention, the science of preventive education, health/adjustment, school children

I. 予防教育科学に基づく「子どもの健康と適応」のためのユニバーサル教育

山崎他（山崎・内田，2010；山崎・佐々木・勝間・松本，印刷中）は，科学的根拠（エビデンス）に基づいて学校で実施する健康と適応のためのユニバーサル予防教

育を提唱している。その教育は，「いのちと友情」の学校予防教育（TOP SELF: Trial Of Prevention School Education for Life and Friendship）と呼ばれ，ベース総合教育（comprehensive base education）とオプショナル教育（partial optional education）が準備されている。

ベース総合教育は，小中学校の多くの学年で通年実施

され、健康と適応を全般的に予防する教育であり、学年進行にともなって次第に安定した強い効果をもたらす。一方、オプション教育は、特定の問題に特化して（いじめ問題など）短期的に実施される予防教育で、効果の持続性はベース総合教育に劣るが、高い速効性が期待される。

これらの教育は、目標構成や教育方法において、これまでの教育にはない多くの斬新な特徴を備えている。たとえば、エビデンスに基づいて導出された目標は、最終の大目標から、方法に直結する操作目標まで階層的に構成され、教育方法が各目標から乖離しないことに細心の注意が払われている。また教育方法も、目標達成に直接つながる内容に加えて、子どもたちの教育への参加意欲を最大限に引き出す工夫が随所になされている。たとえば、複数回にわたる授業が1つのドラマを構成し（たとえば、宝探しの旅に出る物語など）、そのドラマに参加しているような感覚をもって、各教育目標が達成されていく。

このように、教育目標や教育方法の構築において際だった特徴をもつ TOP SELF は、教育評価においても、これまでにはない新たな特徴をもつことが期待されている。なお本論文は、山崎・内田（2010）ならびに山崎他（印刷中）の続編とも言える内容をもっているため、TOP SELF に固有の概念や用語を理解するためには、これら先行論文を読んでいることが推奨される。

II. 予防教育における評価の特徴

1. 効果を確認する評価対象

TOP SELF の大きな目標の1つは、心身の健康の維持と向上である。この場合、教育に効果があつたかどうかの評価は、健康を目標とする場合、まず健康が維持されているか、高められたかという、直接的な健康側面が強調される。しかし、たとえば生活習慣病の予防になると、その成果は子どもたちが中年期以降になって生活習慣病に罹患するかどうかの長期的な視点を持たなければならないことになる。TOP SELF はユニバーサルな予防を目指すだけに、長期的な視点を欠くことができない。また、いじめ予防の教育でも、学校でいじめが減ったかどうかについてその後数年にわたる結果指標が最終的な評価対象になることが考えられる。

こう考えると、生活習慣病予防の教育等では、教育直後には直接的な評価結果を得られないことになり、教育効果の評価がおぼつかなくなる。しかし TOP SELF は、いかに生活習慣病予防のような中年期以降の疾病の予防を対象にしたとしても、その最終目標を構成する目標が階層的に組まれ、それぞれ下位の目標を達成することが上位の目標を達成することにつながり、ひいては最終目標（生活習慣病に罹患しない等）を達成することになる。

この点は、各目標がエビデンスと論理性をもって構成されていることから可能となる。つまり、このことは、最終目標を構成する下位の目標の達成度合いが教育効果の評価になることができることを示している。エビデンスをもって目標を構成することは教育の科学性を保証する1つの手だてであるが、この効果評価対象としての下位目標の構成も、エビデンスをもって目標を階層的に構成することの今ひとつの意義と言える。

例を挙げてみよう。表1は、TOP SELF オプション教育の精神健康系内のストレス予防プログラムの目標構成（佐々木・石垣・磯貝・山崎，2011）である。精神健康系の大目標は精神的な健康の維持・向上であるが、その直下の目標の1つがストレスの予防・低減になり、ここにストレス予防プログラムが展開される。その直下の目標は（構成）上位目標と呼ばれるが、その下には階層的に、（構成）中位目標、（構成）下位目標、操作目標が続き、操作目標の下に諸々の方法が展開されることになる。このプログラムの場合、上位目標が直接的に評価対象になることもできるが、操作目標に至るまでの目標も評価対象になることができる。特に、中位目標レベルでの各目標は、その水準と数の適切さから評価の中心になることができる。目下のところ、オプション教育の各教育プログラムでは、構成上位目標が直接的に評価対象になる場合は評価を行い、中位目標の各目標は評価対象の中心に据えることになる。中位目標ごとの評価は、目標のうちどの側面の達成度が低いか、高いかを示し、教育の改善点を明示してくれることになる。そもそも、何らかの教育を行った場合百点満点の評価結果はあり得ず、もしそうならば、それは評価に問題があることを意味している。つまり、評価は、教育の問題と発展の方向を示すものでなくてはならない。

近年、医学領域では、エビデンスをもった医学（evidence-based medicine）の必要性が叫ばれている。TOP SELF は、その必要性を学校教育、特に学校予防教育に適用している。ここでエビデンスに基づくことがどういふことなのか、学校予防教育を対象に考えておきたい。まず、行われた教育の効果を科学的に実証することで、この点については次節以降で詳述される。また、上述したように、教育目標が実証的科学研究におけるデータやそこから導出された理論に基づいて構成されていることもエビデンスに基づいた教育要素と言える。そして、教育方法が、可能な限り同様のデータや理論に基づいて教育効果が望めるものとして採用されることもこの教育要素といえる。ただ、教育方法に至っては、教育方法を導くエビデンスが少ないことと、実際の学校教育場面では細部に至り多様な教育方法が必要となり、この点では、教育効果の検証をもってその方法の是非を検討する方向でのエビデンスとなることが多い。ただ、教育方法の構

表1 オプショナル教育 精神健康系 「ストレス予防プログラム」の目標構成（佐々木他，2011より）

構成上位目標	構成中位目標	構成下位目標	操 作 目 標
ス ト レ ス の 低 減 ・ 予 防	I. ストレスが生じるメカニズムと自らのストレス過程の理解	1. ストレスが生じるメカニズムと、ストレス状況下における様々な反応の相互関係について理解する。	a. ストレスが生じるメカニズム（ストレッサー，認知的評価，コーピング，ストレス反応）について理解する。 b. ストレスが生じている時の状況・感情・認知・行動・身体反応の相互関係について理解する。
		2. ストレス過程の各要素について，日常生活の実際のデータから，自分の特徴を理解する。	c. 日常生活の中で，ストレス過程における各要素（ストレッサー，認知的評価，コーピング，ストレス反応）を自己観察する（セルフ・モニタリング）。 d. 得られたデータから自分のストレス過程の特徴を理解する。
	II. ストレス反応（身体的・心理的反応）を直接的に制御するスキルの獲得	3. リラクセーションの習得	e. 呼吸法の習得 f. 漸進的筋弛緩法の習得
		4. 情動焦点型コーピングのレパートリーの拡大	g. 感情表出の方法の習得 h. 情緒的ソーシャル・サポートの促進（サポートを求めるためのコミュニケーション・スキルの習得を含む）
	III. ストレス反応を引き起こしやすい認知（認知的評価）を制御するスキルの獲得	5. ストレス状況におけるネガティブな認知をとらえ，距離をおくスキルの獲得	i. ストレス反応（怒り・不安・抑うつ・身体反応など）をシグナルとして，ストレス状況における自分の認知（自動思考）をとらえるスキルの獲得（相手からの攻撃・侮辱→怒り，喪失→悲しみ，落胆，脅威→不安，緊張など） j. ネガティブな認知から距離をおくスキルの獲得（生じやすいネガティブな認知に名前をつける，TVの中の出来事のように実況中継するなど）
		6. ストレス状況におけるネガティブな認知を断ち切るスキルの獲得	k. セルフ・トークによる思考ストップ法の習得（「ストップ！」と頭の中で言う，電話をがしゅんと切るイメージを思い浮かべるなど） l. 道具を使った思考ストップ法の習得（輪ゴムばっちゃん法，指に印や字を書いておくなど）
		7. ストレス状況における認知のバリエーションの拡大	m. ネガティブな認知に反論するスキルの獲得 n. ネガティブなイメージを描き換えるスキルの獲得
	IV. ストレッサーの消失・軽減のための問題解決スキルの獲得	8. 問題となるストレッサーを特定し，解決方法を考えるスキルの獲得	o. 問題を明確化し，適切な目標を設定するスキルの獲得 p. 問題解決方法の選択肢を複数考案するスキルの獲得 q. 結果を予測しながら最良の問題解決方法を選択するスキルの獲得
		9. 計画に基づいた問題解決法を実行し，結果を評価するスキルの獲得	r. 選択した問題解決法を実行するスキルの獲得（問題解決過程に必要な場合にはソーシャル・スキルや自己主張スキルを含む） s. 実行した問題解決の結果について評価するスキルの獲得

築時には，既存のエビデンスは最大限に利用する構えが重要で，山崎他（印刷中）は，実際には参照すべき理論やデータがあるにもかかわらず，それを参照せず主観や経験に頼って教育を構築しがちな学校現場の姿勢に注意を促すことが肝要であると述べている。

2. 3種類の評価とその目的

教育を実施して効果を確認する場合，3種類の評価が存在する。1つは，実施した教育が，たとえば小学校高

学年に適用する教育であった場合，日本の小学校高学年全体を母集団と考え，その母集団全体に効果があるかどうかを検討する効果評価である。この評価は，今後ある種の教育方法を広く使用していくことの是非についての科学的な根拠を提供することになる。

2つ目は，教育を実施したクラス全体に効果があったと言えるのかどうかにかかわる効果評価である。ある教育を学校クラスを対象にして実施した場合，クラスのすべての子どもが何らかの恩恵を受けることはまれであり，

何ら恩恵を受けない子どもも当然出てくる。しかし、クラス全体として平均値等の変化で見た場合に見かけ上でも効果があったと言えるかどうかは、教育現場ではその教育を実施する是非を決定する重要な情報となる。もちろん、あるクラス全体のある特徴について、見かけ上の平均得点の上下の変化のみでは科学的に効果を評価することはできないが、実際にクラス全体の特徴の変化をとらえることは現場の教員にとっては意義がある。特にある時期や状況において、教育目標となる子どもの特徴の一般変化（または無変化）が予測される場合、その予測との比較において、想定以上の好ましい変化があれば教育効果が認められた可能性は大きい。後述するが、上記の1つ目の効果評価は、科学性が高いものの、その実施には相当な経費と労力が必要で、実施することは困難を極めるだけにこの2つ目の観点が強調される。

そして最後は、個々の子どもにおける教育効果の検討にかかわる評価である。学校教育では、個々の子どもの変化が独立した最小単位の評価となり、またもっとも重要な評価と言える。この場合も、見かけ上の特徴の改善や悪化は科学的評価とは異なる側面をもつが、それでもこの評価は学校における教育においては中核をなす。

こうして3つの効果評価が指摘され、第1の効果評価は、教育方法の開発段階では必要になり、科学的な効果の理想型を考えることができる。第2と第3の効果評価は、科学的な効果評価とは言えないものの、学校においては欠くことができない教育評価となる。

一般に、算数や国語等の教科の成績に比べて、TOP SELFが教育目標に置くことが多い心理的特徴の評価はむずかしい。それは、教科にある目標のように、達成を目指す内容に客観性を付与することがむずかしいからで、その達成目標は構成概念になることが多い。そもそも、心理的特徴の多くは、客観的に外から見えることはなく、我々が構成した概念になる（構成概念）。もっとも、後述するように、明確な得点で示される教科の成績評価は、その相対性が際だちやすいことから多くの弊害をもたらす可能性には注意したい。

3. 母集団への効果を想定した科学的な効果評価

心理学や多くの科学は、因果関係の究明がその重要な仕事になっている。予防教育の場合も、ある種の教育をして（原因）、目指した方向に子どもたちが変容したかどうか（結果）を検討することになる。つまり、教育が望ましい変化を導いたかどうかを調べる。しかし、この因果関係の科学的な究明は、きわめて困難な仕事となる。誰もが納得するかたちで、因果関係の有無に結論を出す方法があるかどうかもあるやしい。その困難さは山崎（2009）に詳しいが、科学的に欠点のない因果関係の究明法はないと言っても過言ではない。ただ、その欠点を

最小限にするために方法を工夫することが必要であることは強調される。

予防教育を実施した場合に限定して考えてみよう。まず、教育を実施するグループと、比較対象としてのグループが必要になる。この比較対象のグループの必要性は、教育グループでは教育実施以外にも多くの要因が働き、教育後に変化があったとしてもそれが教育による変化なのかどうかを明らかにすることが困難になる問題を克服することからくる。もし教育が施されなかったらどうなっていたのかを推定するために、教育を行っていない比較対象のグループ（統制グループ）が必要になる。たとえば、暴力性の低減をはかる教育を実施して、たとえば暴力性が高まったり、変化がなかったとしても、教育効果がなかったとは結論できない。つまり、同じ時期に、教育を施さない場合暴力性は大きな高まりを示すことがあり、教育はその高まりを抑えたことも考えられるからである。

この場合、教育グループと統制グループの設定にも困難な問題がある。すなわち、両グループは教育を行うか否か以外は、同質のグループである必要がある。この同質性を保証するためには、グループの特徴を規程する要因が無数にあるために、無作為に多数のグループ成員を集める必要がある。しかも各グループは、母集団を精度高く反映する必要がある。たとえば、日本における小学校高学年に適用して効果があるかどうかという仮説であれば、設定した両グループはその母集団を表現できるものでなくてはならない。他にも統制グループの設定には考慮すべき点は多い。教育グループでは、教育を実施するとすると、教育で設定した方法以外に教育と同時に動く要因が多数ある。このことは下にも別の観点から記述されるが、たとえば、教育を行って教育目標に関連する事態にさらされるだけでも何らかの効果がある可能性があり、その点では教育目標への注意等を同程度経験するだけの統制グループ（attention-placebo control group）が必要になるが、真にこのための統制になることはむずかしい。

次に、グループ間の比較は、何らかの統計的手法を使って実施される。それは、実際に教育を行うグループ、また統制グループに入るメンバーは、大きな母集団からすればその一部であるからで、確率的に教育が効果があったのかどうかを決定する必要がある。ここで統計分析が登場し、何らかのデータが利用されることになる。この場合、多くの統計的手法ではデータの独立性が前提条件にある。つまり、データが相互に影響を受けていないことが必要になる。そうすると、一人ひとりに子どもが統計分析時の素データになることがむずかしい。クラス全体の平均値も同様である。なぜなら、子ども同士や隣合ったクラスは相互に影響を受けるからである。完全に独立

したデータが実際にあるのかどうか自体も疑わしいが、その独立性がある程度は保たれ、またデータ数もある程度は集められたところで妥協することになり、この妥協点に学校を単位としたデータが推奨されることが多い。この場合でも、最低数十校の学校が教育と統制の各グループに必要になり、科学的効果評価の道は、経費や労力の必要からきわめて険しいと言わざるを得ない。もっとも、この種の効果評価を公表している論文は少なくとも（たとえば、Edmundson et al., 1996）、コクラン・コラボレーション（Cochrane Collaboration）のデータベース（The Cochrane Database of Systematic Reviews）などでは、厳しい基準でそのような効果評価研究を紹介している。残念ながら日本においては、学校での教育効果評価において、上述の基準を満たす研究は皆無である。

これが科学的効果評価のあるべき姿の原型と言えるが、上で示唆したように、この評価の科学性も完璧ではない。というのは、何らかの教育手法をもって教育を行った場合、その教育方法と交絡する要因が存在し、実際にはその教育方法ではなく、交絡した要因が効果に影響した可能性がある。簡単な例を挙げると、体重が糖尿病の発症に影響するかどうかを調べるため、高カロリー食を多く摂らせた場合、体重の増加以外に多くの要因が同時に動き、体重そのものの増加が糖尿病の発症に影響したかどうかを決定することがむずかしくなる。このような交絡要因は思いのほか多く、教育を実施する場合はこの要因に注意すること、という警告を発するばかりである。山崎（2009）は、この点を指摘し、目下のところ最適とも言える方法を示唆しているので参照されたい。

なお、科学的な効果評価の詳細な条件は Task force on evidence-based interventions in school psychology（2003）に詳しいので参照されたい。

4. 効果評価の測定方法

(1) 質問紙の利用

科学的効果評価のデザインは上に紹介された通りであるが、どれほど良いデザインでも、具体的な測定方法に問題あれば科学的評価はおぼつかない。特に、予防教育においては、測定対象が構成概念である場合が多いので、測定方法の問題は重要かつ複雑になる。

一般に、教育の前後を中心に、さらには、教育後一定期間を経たり、あるいは教育中に測定が行われるが、その場合、ある程度上の年齢であれば、もっともよく利用される方法が子ども自身が回答する自記式の質問紙である。自記式であるから幼い子どもへの適用は無理であるが、回答が可能な場合はその簡便さから頻繁に利用される。しかし、構成概念を測定する質問紙には多くの欠点がある。まず、標準化の精度の高い質問紙には信頼性と妥当性が高く備わる必要があるが、この妥当性の確認が

不十分である質問紙がほとんどである。質問紙の作成や是非の見分け方については山崎・内田（2005）に詳しいが、質問紙が測定しようとする対象の世界（ユニバース）を満遍なくとらえた項目をそろえた質問紙の準備とその妥当性の確認は困難を極める。似通った項目をそろえれば自ずと項目の等質性が高まりこの点での信頼性が高まる。また、一般に信頼性の確認は比較的容易であるが、妥当性の確認はむずかしい。妥当性が極めて高い尺度は同時に信頼性も高いが、信頼性がきわめて高い尺度の妥当性が高いとはかぎらないことから、妥当性の問題が重要になる。近年の妥当性の概念は単一の概念であるとされ（Messick, 1989）、その集約は構成概念的妥当性にあると言える。この構成概念的妥当性を証明する材料を揃えるのはかなりの労力が必要で（たとえば、Cizek, Rosenberg, & Koons, 2008 を参照）、また、上述した質問紙のユニバースの適切性を十分に検討できる現実的な方法は容易には見いだせない。そうすると、できる限り標準化の精度が高い質問紙を採用するという曖昧な妥協点しかないが、質問項目構築時に注意を払い、せめて専門家による内容的妥当性（content validity）は高めておきたい。

他にも、子どもが意識して自ら回答する質問紙は防衛が働くこともあるので、教師や親など多方面からの回答も得たい。ただ親、特に父親の評定は一般に精度が落ちることが指摘され（たとえば、Vega-Lahr & Field, 1986）、教師もその回答態度に個人差が大きい。この点から、現状では質問紙は、信頼性と妥当性に細心の注意を払い、その限界を知った上で子どもによる自記式を実施することが推奨され、回答態度のあり方や回答方法を十分に理解してもらった上で、担任教師がクラスの子どもの対象に質問紙を実施することが推奨される。

質問紙の他には、項目数を落として、クラスの仲間（特に同性仲間）全員からクラスメンバーの一人ひとりについて評定を得たり、また指名法（nomination method）で対象となる特徴が際立つ子どもを指名する方法もある。後者の指名法は、もっとも精度の高い測定方法の1つであるが、同じクラスの子ども同士が互いに評価をするという点に倫理上の問題を指摘する向きも多く、実施は容易ではない。

(2) TOP SELF の効果評価における測定方法

質問紙等は回顧的な測定方法であるが、実際に何らかの行動を観察によってとらえる方法もあり、観察法にはかなりの精度が期待される。ただ、クラス全体を運営する教員がこの観察をリアルタイムで実施することはむずかしく、またそのための補助者を置くことも現実的ではない。そこで、目標となる行動の意識を高めて、その日の放課後等に行動の有無をチェックすることが考えられ

る。後述するように、TOP SELF ではこの行動採取は望ましい行動に限定した採取になる。

また、教育によっては、欠席日数、保健室来室等、実態がある結果をとらえることもでき、この点が可能ならばかなり説得力のある評価となる。しかし、各教育でこの実態的な評価が可能になる場合は少ないものと予想される。

また効果評価は、子ども個人を対象にするばかりではなく、クラス全体の特性について個々の子どもが評価するという方法は、その簡便さと意味の大きさから取り入れたい。つまり、子どもたちが普段一緒に過ごす度合いの高い学校クラス全体をとらえる印象は、社会的な妥当性が高く、また実施に際して抵抗がないので、実施も容易である。この点では、TOP SELF で活動の中心となる小グループ全体への評価（構成メンバーによる）も考えられる。

こうして、目下の TOP SELF では、信頼性と妥当性を考慮した自記式質問紙と教師評定、クラス全体や小グループを対象とした評定、子どもと教師による行動観察を主要な評価方法としているが、これまでにはない観点からの評価も実施しており、この点については次に取りあげたい。

III. 教育としての観点を強調した評価のあり方

1. 評価の目的

個々の子どもにフィードバックされる評価は、どのような目的があって実施されるのであろうか。まず、子どもの実際の姿や結果を示す目的がある。実際の姿を示すことにより、改善や向上を期待する。通信簿やテスト得点はその類であり、学校における個人への評価の大半はこの評価となる。またその場合は、他の子どもとの比較により、相対的に結果を示す場合も少なくない。相対的に示す意図がなくても、点数等で明示された評価結果は、子ども同士の開示により自然と相対的になる。相対的に示すことで、自分の成績のもつ意味（優劣等）を示しやすいことにもなる。

この評価により子どもたちには、もっと頑張ろう、よく頑張った、まだまだダメだ等々、様々な心の動きが出る。もちろんこの評価は、「もっと頑張ろう」や「よし、この調子だ」という前向きな行動操作をねらうのであるが、現実にはねらい通りの結果をもたらさない場合が多い。極端な場合を想定してみよう。もしクラスの子どもたちがすべて成績が1番になりたい、なるべきだと考えているとしよう。すると、成績が順位で出されると1人をのぞいてすべての子どもにとってこの評価は罰になる。つまり、このクラスの子どもたちはほぼいつも罰を受けているということになる。このことが長引けば、遅かれ

速かれ子どもはやる気をなくしてしまう。学習性無力感（learned helplessness）に陥る例にあるように（たとえば、Overmier & Seligman, 1967; Seligman, 1975）、やっても無駄だということを学習していくことにもなりかねない。こうして、結果を客観的に知らせる、また相対的情報をもって知らせることには害となる要素が少なくない。これに対して絶対的評価として、たとえば、子どもが前回の自分自身の成績よりどれほど変化したかという情報はどうか。これは相対的な評価ではないので、クラスの子どもたちすべてにとって同様に賞となることもあり得る。しかし、この絶対的評価の場合にも、評価が向けられる側面やそのとらえ方によっては子どもの伸びている側面が強調されない方法となる。

近年のポジティブ心理学（positive psychology）は、人間のネガティブ（負）な側面を直すのではなく、ポジティブ（正）な側面を伸ばすことを重視している（Seligman, 2002; Seligman & Csikszentmihalyi, 2000）。この正の側面を伸ばすということは、多くの教科で新しい知識や考え方を教えていくように、学校教育にはふさわしい教育内容になる。また、上述したように、TOP SELF ベース総合教育の大目標の1つは自律性の育成である。そして、自律性の構成要素の1つが内発的動機づけの高さとなる。この内発的動機づけを損なう要因は多数あるが、コントロール力の高い賞罰や相対的な比較等、外的コントロール性があればその要因として挙げられる（Deci, Koestner, & Ryan, 1999）。こうして、この内発的動機づけの高さを保ちながら子どもたちを伸ばすには、他者との比較による相対的な評価や教師によるコントロール性の高い評価は避ける必要があり、また罰を与えることも問題となる。このことは、ほめること、賞なら何でもよいということではないことも示唆している。つまり、上からのコントロール性の高い賞の問題性が指摘される。子ども自らが自分の向上度合いを評価し、また、子どもの望ましい変化を認め、感動を分かち合うように寄り添う教師の評価が必要になろう。さらに補足すると、ここでは罰ではなく賞の重要性を強調しているが、それは、ベース総合教育で目標とする自律性の発達、自らの能動的な行動が自らの欲求を満たすという成功体験によるもので、決して失敗や罰によって形成されるものではないからである。

2. 教育としての評価

こうして、TOP SELF における評価の主要部分では、子どもの正の側面に焦点が当たることになる。この正の側面の評価の1つは、上にも少し紹介されたように、教師による、目標にかなった望ましい子どもの具体的な行動の採取である。この行動採取は子ども自身も実施し、教育の中間時点や終了時点に実際の行動を採取する時間を設けることが考えられる。これら実際に子どもが採取

表2 TOP SELF における評価の構成

既存の方法の精緻化

- ・質問紙（自己，教師評定）を信頼性と妥当性に配慮して使用する。
- ・プログラムによっては，実態のある側面（欠席日数や保健室来室頻度等）をとらえる。

TOP SELF のセミオリジナル

- ・教育目標にとって望ましい行動（正の行動）の観察・採取（教師，自己による）を行う。
- ・個人によるクラス等全体評価（簡易質問紙による）を行う。

TOP SELF オリジナル

「評価も教育である」という観点から，特別な方法を使用する。

- ・可能な限り，子どもによい側面を提示する。このため，教育目標に関連して向上した側面を強調した評価方法を用いる。

※こうして，上記の評価は，プログラム途中ならびに実施後に時系列的变化が明確になるかたちで子どもに提示され，また保護者にも提示される。

した教育目標に沿った望ましい行動を評価として戻すことは，実際に行われた具体的な行動であるだけに強化力が強い。とりわけ，ここでの教師からの評価はコントロール性はほとんどなく，子どもにとっては教師からの添うような自然な強化となり，目標となる行動をさらに促進

できる評価となる。

次に，教育の前後等の時期に子どもに実施する自記式の質問紙と同じ質問項目を利用し，それぞれの質問項目の内容における向上度を教育の中間時点や終了時点に子どもによる自己評価として実施することも考えられる。とくにこの評価は，伸びている側面に鋭敏な尺度とすることができる。評価の側面は，マイナス面あるいはプラス面を重視してとらえる方法，そして両者を客観的にとらえる方法があるが，ここではプラス面を重視してとらえることになる。

表2には，TOP SELF における目下の教育評価方法がまとめて提示されている。ベース総合教育とオプションル教育では，また各教育内のプログラムによっては細部が異なるが，およそ表2のような評価内容がTOP SELF の現状の評価であり，現在もなお，大きく改善されつつある。

IV. 教育としての評価を強調した TOP SELF の評価例

そこで，TOP SELF のオリジナル，セミオリジナルとなる評価方法の一部の実際を1つのオプションル教育を例に紹介したい。ここに挙げた評価は一例で，またその細部は今後も発展的に変更されていくことには留意されたい。

図1の(a)には，オプションル教育で使用された評価ノートの表ページが示されている。現在，この評価ノートは「お宝ノート」と呼ばれている。ページを開くと，

(a)

(b)

図1 TOP SELF オプションル教育評価例（お宝ノート）：（a）ノート表紙，（b）ノート最終ページ

あなたが見た「あなた」

めざすことⅠ

(自分の生活習慣について知っていて、なおした方がよいところがわかっている)

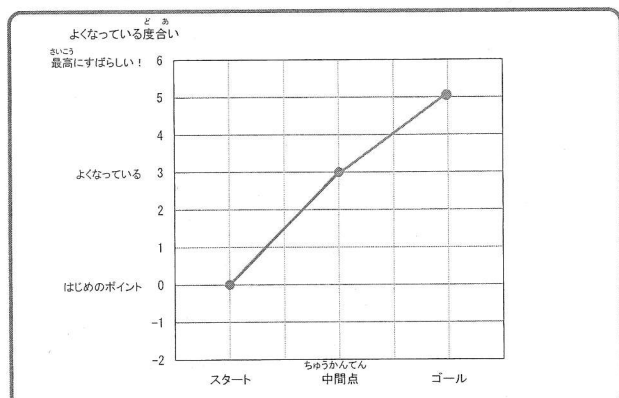


図2 TOP SELF オプショナル教育評価例 (お宝ノート) : 子どもによる向上度自己評価例

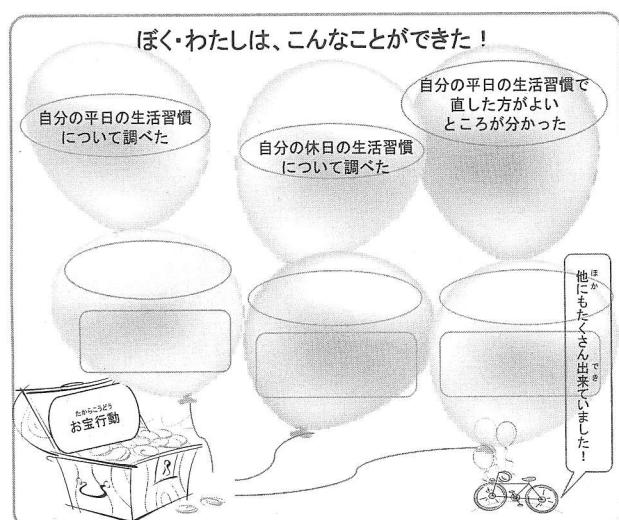


図3 TOP SELF オプショナル教育評価例 (お宝ノート) : 子どもによって採取された行動の提示例

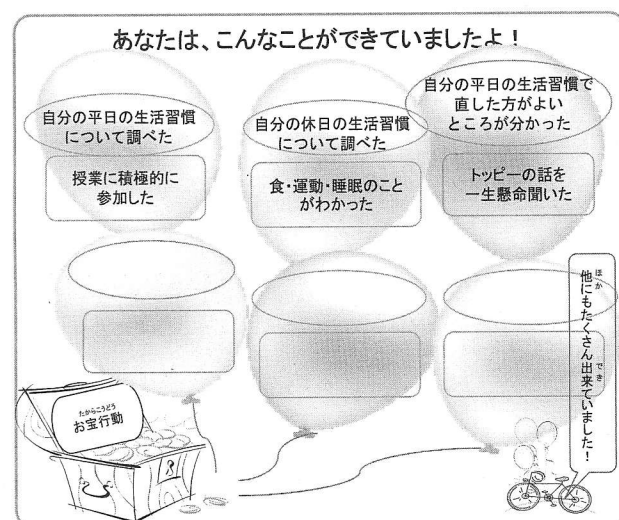


図4 TOP SELF オプショナル教育評価例 (お宝ノート) : プログラム実施者によって採取された行動の提示例

最初のページに、図2と図3が上下に並んでいる。図2は、児童・生徒自身がとらえた、1つの教育目標に対する向上度である。この評価では、オプショナル教育の(構成)中位目標がそれぞれ独立したページに示されている。図2において、各(構成)中位目標について問う3つの質問項目で、各項目に示されたことが、わるくなっている、よくなっている、とてもよくなっているの5段階で評価し、3項目の平均得点を示したものである。この選択肢からわかるように、よくなっている程度の細分化の程度が大きく、よくなっている方向での評価が強調されている。これらの向上度得点では、小グループやクラス全体の平均を出すが、これを個人の変化と対比できるようにして提示することは相対的な評価となるので、現在のところTOP SELFでは採用されていない。

図3では、図2のページの下半分に来る、教育の中間時点で子どもたちが採取した目標にそった行動が、具体的に示されている。これらの採取は、目標に合致した見本行動の生起(紙面に列挙されたもの)にチェックを入れたものや、子ども自身が自分で記述した行動が含まれている。この図2と図3にある評価は、担任教師(場合によってはプログラム実施者)によっても実施され、図2と図3の次ページに示される。図4は、教師による評価ページのうち下半分の評価、つまり教師による行動採取の結果が子どもに示されている。

こうして、これらの評価が目標ごとに示され、しかも、子どもの正の側面が強調された評価になっている。また、教師がコントロール力をもってほめるような評価は含まれていない。図1の(b)は、このお宝ノートの最終ページであるが、教師ならびに保護者から、子どもへの前向きな側面を強調したメッセージが入り、締めくくられる。

児童・生徒に返却開示される評価はこのような子どもの成長を示す評価に限定されるが、表2に示したように、個人に返却されない評価として、教員や子どもが別の観点からとらえる評価方法が他に用意されていることは言うまでもない。

V. 今後の評価の課題と発展

本論文で紹介してきたように、教育における評価は多面性があり、子どもを伸ばすという点では再考を余儀なくされる評価が少なくない。TOP SELFでの評価も、教育方法の進展とともに今後改善され、発展していく余地が多分に残されている。そこで、本論文の最後に、教育評価にまつわる課題と今後の発展についてふれておきたい。

現段階のTOP SELFはまだその域に達していないが、本論文で紹介した、統制群を設定した上での科学的な評

価は TOP SELF が広がりを見せる段階では欠くことができない。しかしながら、各オプション教育は別として、ベース総合教育のような規模の大きな教育になると、その評価をどの単位で実施するかはむずかしい。山崎他(印刷中)では、ベース総合教育を構成する4つの構成目標にかかわる教育が小学校3～6年生と中学校1年生の各学年でそれぞれ8回(週)ずつ実施されるというプランが示されているが、この各構成目標の8回が最小の効果評価対象としての教育になる可能性が高い。また、4つの教育の総合効果を検討することも各学年で実施することも可能であろう。もちろん、科学的な評価は、多数の教育参加校と統制条件参加校が必要になり、その実施は繰り返して行うことはむずかしいので、時期を選んで慎重かつ効果的に実施すべきであろう。

この点では TOP SELF は、目標構成と教育方法に科学性をもたせていて、この教育を行えば最終目標が達成される可能性を高めた上で教育実施をしている。このことから、TOP SELF の効果に科学的な根拠が付与された後に多くの学校に展開するというよりも、広域の学校で TOP SELF が実施されるようになった後に科学的評価を試みるという手順が現実的であろう。

次に、具体的測定方法の問題がある。客観的に正負対等にとらえるこれまでの評価に加えて、子どもの伸びている側面を強調した評価の組み合わせは TOP SELF の評価特徴になっていた。しかしいずれの場合も、評価として使用する質問紙の妥当性の問題が気になり、この点はとりわけ注意する必要がある。本論文でも妥当性の検証の困難さは指摘された。確かに、測定する構成概念のユニバースを満遍なくとらえた質問紙はむずかしい。また、その証明は現実的には不可能であるかもしれない。そうであれば、質問紙を構成する項目の内容的妥当性を最大限に高める必要があることにもふれた。この過程に科学性を付与することは極めてむずかしいが、少なくとも測定概念に精通した複数の研究者が共同で確認した上での質問項目ではありたい。

そして、最後は仲間同士の評価の可能性にふれたい。発達心理学における基礎研究では、仲間からの評価が重視される。特に、当該の特徴について際だったメンバーを指名する方法が推奨され、妥当性の高さが示唆されている。しかし、この種の評価方法が日本の学校においてはきわめて難しいことも事実である。マイナス面をもった特性への仲間評価はとりわけ敬遠されるが、プラス面に限定した評価もなかなか受け入れられない。この評価では、子どもたちは、否が応でも、クラス内の他のメンバーについての印象の意識化を形成してしまう。当然のことながら、そこでは仲間への評価が実行されるが、その評価を紙面に表すことが拒まれるわけで、この点についての学校の管理者や教員の心情は十分に理解できる。

しかし、そのことによる悪影響を示す客観的なデータがないことも事実である。プラス面に限り、仲間指名まで行かなくても仲間評価でもよいから、この評価方法を適用する可能性を今後模索したい。

教育には評価はつきものである。しかし、本論文で紹介してきたように、その内容は多面性があり、またその信頼性と妥当性を高めた評価はむずかしい。基本的には、評価は子どもを伸ばすためにある。また、評価内容によっては、適用した教育の是非と改善の方向を具体的に示すことができるものでもある。TOP SELF では、教育と評価の両方を尊重し、エビデンスをもった教育として、子どもたちの健康と適応を守りたい。

引用文献

- Cizek, G.J., Rosenberg, S.L., & Koons, H.H. (2008). Sources of validity evidence for educational and psychological tests. *Educational and Psychological Measurement*, 68, 397-412.
- Deci, E., Koestner, R., & Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125, 627-668.
- Edmundson, E., Parcel, G.S., Perry, C.L., Feldman, H.A., Smyth, M., Johnson, C.C., Layman, A., et al. (1996). The effects of the Child and Adolescent Trial for Cardiovascular Health Intervention on psycho-social determinants of cardiovascular disease risk behavior among third-grade students. *American Journal of Health Promotion*, 10, 217-225.
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp.13-103). New York: Macmillan.
- Overmier, J.B., & Seligman, M.E.P. (1967). Effects of inescapable shock upon subsequent escape and avoidance responding. *Journal of Comparative and Physiological Psychology*, 63, 28-33.
- 佐々木恵・石垣久美子・吉見摩耶・磯貝和裕・山崎勝之 (2011). 学校におけるストレスの低減・予防を目的としたユニバーサル教育－教育目標構成とそのエビデンスならびに教育方法の開発・実践－ 準備中の原稿
- Seligman, M.E.P. (1975). *Helplessness: On Depression, Development, and Death*. San Francisco: W.H. Freeman.
- Seligman, M.E.P. (2002). Positive psychology, positive prevention, and positive therapy. In C.R. Snyder & S.J. Lopez (Eds.), *Handbook of positive psychology* (pp.3-9). New York: Oxford University Press.
- Seligman, M.E.P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, 55, 5-14.

Task force on evidence-based interventions in school psychology (2003). *Procedural and coding manual*. (Sponsored by Division 16 of the American Psychological Association and the Society for the Study of School Psychology and endorsed by the National Association of School Psychologists)

Vega-Lahr, N., & Field, T.M. (1986). Type A behavior in preschool children. *Child Development*, 57, 1333-1348.

山崎勝之 (2009). 正感情と“Finding Positive Meaning”
コーピングが健康に及ぼす影響 ふくろう出版

山崎勝之・佐々木恵・内田香奈子・勝間理沙・松本有貴 (印刷中). 予防教育科学におけるベース総合教育とオプショナル教育 鳴門教育大学研究紀要

山崎勝之・内田香奈子 (2010). 学校における予防教育科学の展開 鳴門教育大学研究紀要, 25, 14-30.