

# 体育における学習評価に関する一考察

—学習者の自己評価活動を中心として—

武隈 晃\*・武隈 順子\*\*

(1997年10月15日 受理)

## A Consideration on the Evaluation in Physical Education Learning ; Focusing on the Activities of Self-Evaluation

Akira TAKEKUMA and Yoriko TAKEKUMA

### I. 緒 言

体育科教育において、「評価」と「指導」の相補性に関する認識は、少なくとも理念の上では大方の理解を得てきたと考えられる。しかしながら、「学習指導」に関わる活発な議論や実証的研究とは裏腹に、これまで「学習評価<sup>注1)</sup>」に関するそれは低調であった。内海は「教育評価議論のタブーが、未だ存在していると言わざるを得ない (1995, p.14)」と指摘する。

しかしここ数年、体育に限らず、あらゆる教科で学習評価に関する学校現場での関心は高まっている。事実、「評価」を主題に掲げた教員研修会等が数多く開催されているし、評価の、主に技術(ノウハウ)に焦点を定めた文献の出版が続いている。今回の「学習評価」の論議は、戦後の「相対評価」対「絶対評価」の議論と双壁を成すほどの高まりを見せている、という見方もできなくはないが、「相対評価」の議論の時とは様相が異なる。1989年の小学校学習指導要領(学習指導のスタンダード)の改訂に続く1991年の指導要録(学習評価のスタンダード)の改訂後、「新しい学力観」なる用語が定着し、これを受けた評価の在り方についての関心が、法的拘束力を持つ指導要録への記載という現実的なレベルでの必要によって喚起されるという図式を見ることができる。また、学力観が評価法を規定するという通常の流れに抗して、評価観が、逆に学力の内容を定立するという様相を呈しているところに特徴があり、かかる議論の創始が現場サイドというよりも、国レベルの教育体制にあるとする見解が示されている。

「新しい学力観」に相応する評価の枠組みは、知識や技能偏重のこれまでの評価観に代って、1) 関心・意欲・態度、2) 思考・判断、3) 技能、4) 知識・理解を大枠とし、しかもそれらの優先順位をも規定している。つまり、3) 技能、4) 知識・理解などの比較的「見えやすい」学力より

\*鹿児島大学教育学部

\*\*鹿児島県入来町立朝陽小学校

も、1) 関心・意欲・態度、2) 思考・判断などの「見えにくい」学力をより重要な内容として位置づけているのである。体育科教育において、「技能」は学習者の運動行動を教師が観察したり、客観的なテストを実施することによって、また、知識・理解（小学校の体育ではこの項目が除かれている）についても工夫されたテスト等によって一定程度把握することができた。しかし関心・意欲・態度や思考・判断についてはそれほど容易ではないという認識がある。

ところで、体育においては他の教科と異なり「態度」を教科目標・内容として古くから位置づけてきたという経緯がある。それは概ね次のような理由によるものであろう。体育の教科としての存在理由として、「子どもの発達」という「運動の主体」の側に焦点を当てる場合と、「文化としての運動やスポーツ」という「社会的事象」の側に焦点を当てる場合があり、両者の立場はしばしば鋭く対立し、また、それらの「総合性」に立った現実的な教科論も主張されてきた。いずれにせよ「態度」は必要不可欠な目標・内容として位置づけられたのである。すなわち、前者においては体育における態度の在り方やそれに関わる学習が子どもの「社会的発達」に貢献するものとして理解されてきたし、後者においては、体育学習の主要な対象となる「運動文化」の構成要因として「運動技術」や「ルール」と並んで、「マナー（行動のしかた）」が明示的に押さえられてきた。従って、態度はいずれの立場をとるにせよ、評価対象の計測（測定）可能性を厳しく問わない限りは、指導と評価の対象として常に視野に入れられてきたのである。にも拘らず、体育においても戸惑いは隠せない。それは「態度」が体育における伝統的な目標・内容であるとしても、それは協調・公正・安全など、どちらかといえば個々の運動固有の構成要因というよりも、運動全般に亘る「一般的態度」としての理解が普通であったこと、また、それは「新学力観」が示す、「関心・意欲・態度」の一部を構成するに過ぎないことに主要な理由がある。さらに、これまでは「運動技能」に評価の主軸を置き、「態度」は副次的な位置づけが成されることが（少なくとも現場レベルでは）多かったのに対して、それが評価の第一項目に位置づけられると、改めて特別の対応をしなければならないという意識が高まったことにもその理由がある。こうした状況において、先に指摘した研修会の興隆やノウハウ的な文献の普及という事態が生まれたものと考えられる。

体育において難しい問題はこれに止まらない。思考・判断が評価の第二項目に位置づけられたが、体育の授業においてどれだけ思考・判断の機会が保障されているのかという問題がそれである。ここに、体育や芸術系教科の課題がある。宇土は、学習評価の最低限のレベルとして「学習内容の習得状況をとらえる（1981, p.47）」ことの重要性を指摘している。仮に、体育の授業で「思考や判断」の必要な学習活動やそれに関わる指導が成されていなければ、それを評価の対象とするというのは奇異である。逆説的になるが、評価対象とするからには、それを成立させる内容が意図的に設定されていなければならない。かかる問題については、1)「課題解決型」の学習を組織すること、2) ボール運動や球技において、たとえば「戦術」等の学習上の財を内容として位置づけること、3) 器械運動・陸上運動（競技）・水泳等の個人的スポーツにおいて、技術認識に関わる学習が意図されること、等々が必要であると考えられるが、これらの点についての検討は他日を期したい。また、

指導要録が示す評価の枠組み自体に関する問題性についての論議も別稿に委ねることとする。

今回の状況において、もう一つ注目しなければならないことがある。それは学習者の「自己評価」への関心が急速に高まったことである。それは既に述べたことと関連している。すなわち、関心・意欲・態度や思考・判断という、教師からは「見えにくい」学力を評価の対象にしようとするれば、いきおい学習者の内省に頼らざるを得ないという判断がある。これを具体化するものが「学習者の自己評価（以下、単に『自己評価』と表記されたものは学習者の自己評価を意味する）」に他ならないという理解の仕方である。

ここで学習評価の基本的意義や機能について若干論及しておきたい。学習評価の最も重要な機能は「学習者の自己理解を助ける（宇土，1981，p. 9）」ことにある。とするならば、自己評価は教師による評価の補助的な手段としてあるのではなく、それ自体目的的な機能をもつものとして位置づけられる必要がある。指導要録や通知表への記載（以下、これらに関わる評価活動を「評定」とする）という二次的な目的に関わる必要から「自己評価」の意義を解釈するのは問題である。通常行われる質問紙調査や心理テストの形式を踏襲した複数の評価項目に尺度構成を与えた自己評価法（学習カードや学習ノートの形態をとることが多い）による評価データは、それが回答者たる学習者によって「評定」に反映されると認識された時点で、信頼性や妥当性を著しく低下させることになる。自己評価は、あくまで学習者自身による自己理解や自己認識を昂進させる意図的な「学習活動」の一環として組織される必要がある。しかしこのことは、教師が学習者個々人の学習状況を判断する手がかりとして「自己評価」による情報が役立たないことを言っているのではない。また、学習評価には、教師が自らの授業（指導活動）の成否や有効性を理解するためのフィードバック情報としての重要な意義がある。自己評価活動についてはその文脈においても理解される必要があろう。

自己評価の本質的な機能が、その副次的機能の有する利便性から曲解されるのは生産的ではない。本稿では、体育の授業における学習者による自己評価活動に焦点を当て、それが学習評価として有する機能を再検討することを目的に、若干の事例的研究の成果を題材としながら考察する。従って、自己評価および学習評価の全体像を整理するというよりも、自己評価活動の評価論的可能性について論及することに主眼が置かれる。

## Ⅱ. 自己評価の対象

体育では「何を評価するのか」、すなわち学習評価の対象についての議論がこれまで論理実証的な立場から十分に成されているとは言い難い。それは教科の目標論や内容論と離れて、評価について単独で議論することが困難であることに根源的な理由があるものと考えられる。目標論やそれを受けた内容論は多様で、その共通理解を得ないことが評価論の本格的な検討を抑止してきたという見方もできる。また、現場レベルでは目標・内容を暗黙裡に認め、意識するしないに関わらず問題の焦点を方法論（特に授業の方法論）に絞らざるを得ないという現実が背景にあるという理解のさ

れ方もある。ここではまず、評価対象論の前提となる体育の目標論・内容論にひとまず座標軸を置き換え、検討を進めることとする。

「身体の教育」、「身体を通した教育」、「運動による教育」、「運動の教育」……。教科としての体育の存在理由を問う「体育の目的論」や、体育の目指す方向性（場合によってはそれに加えて到達規準）を示す「体育の目標」に関わる議論は、それぞれの主張が、一方で、体育（教育）の置かれた社会や時代背景を伴いながら、他方で実践的研究の成果を取り込みながら多様に展開されてきた。それらの概要は、学習指導要領<sup>注2)</sup>や民間教育研究団体のプロダクトから看取することができよう。

内海（1995, pp.88-90）は目的・目標の領域について、1）運動文化領域、2）運動技術領域、3）認識形成領域、4）身体形成領域、5）社会行動領域、6）情意形成領域、7）美的表現領域に項目化できるとしている。また、1）運動技術学習、2）社会性学習、3）認知—反射学習、4）情意学習（Crum, 1992）、1）神経運動領域、2）認知領域、3）情意領域（Barrow & Brown, 1988）などの学習内容や目標領域の構成論にも論及している。また、武隈（1996a）は体育の内容について、1）技術的側面、2）認識的側面、3）関係的側面（関わる・伝える・つくる）で捉えられる「楽しみ方・学び方」を、「楽しさ・喜び」や「関心・意欲・態度」などの情意的「目標」と同次元で扱うことの問題性を指摘し、両者の構造的な把握の仕方について整理している。こうした、目標・内容論から演繹的に評価の対象を明らかにし、このうち学習者が判断可能な部分を自己評価の対象（内容）として設定するという方略が考えられる。

一方、高橋ら（1994）は体育授業の改善に役立つ客観的な授業評価法の開発を試みており、これを自己評価の立場から援用するという方途も考えられる。ただし、高橋らによる授業評価法もその確立過程を遡れば体育の目標・内容論と離れてパラダイムフリーに検討されてきたわけではない。従って前者と対立的なアプローチをとるものと理解するのは適当ではない。ただ、高橋らの特徴は学習者自身の内省によって測定可能な項目を厳選（信頼性や妥当性の高い項目を抽出）するという実証的研究の成果を反映させている点にあり、そのこと自体に評価対象（内容）の客観化を目指す上で重要な手続きを含んでいる。以下に述べられる事例的研究において採用された自己評価項目の設定には、高橋らの研究成果に負うところが大きい。

### Ⅲ. 学習者の自己評価と教師による評価の一致度

自己評価活動の評価論的可能性について検討する最初の題材として、学習者の自己評価と教師による評価がどの程度一致し得るのかについての実験的授業研究を取り上げることとする。

#### 1. 実験条件

##### (1) 被験者とクラスサイズ

K<sub>1</sub>県I町の小学校5年生女子6名（クラス構成員の全て）。

##### (2) 実験授業の内容と規模

「マット運動」の単元（7時間扱い）。

文部省小学校体育指導資料「新しい学力観に立つ体育科の授業の工夫」1995年，p.46に準拠。ただし，めあて①を「できる回り方で，ふつうのマットできれいに回る。くりかえしたり，組み合わせて回ったりする」，めあて②を「場所を工夫してできそうな回り方に挑戦する。できるようになったら，ふつうのマットで挑戦する」に置き換える。

### (3) 実験期間

1997年6月。

### (4) 自己評価の方法

予め作成した「自己評価カード」に被験者（学習者）自らが毎回の授業終了時（単元の1時間目はオリエンテーションや技の達成状況の確認に当てられたため，2～7時間目の計6回）に記入。評価項目は高橋（1994，pp.233-237）を参考に武隈（1996b）が作成した11項目からなる「マット運動の自己評価票」を採用。回答（自己評価）は最高を5点，最低を1点，中立点を3点とする5段階尺度。11項目は以下の通り。

- 1) 今日のマット運動の授業は楽しかったですか
- 2) 自分にあつためあてでしたか
- 3) 練習の進め方はよかったですか
- 4) 自分のめあてに向かって何回も練習できましたか
- 5) せいっぱい全力をつくせましたか
- 6) 「あっ，わかった」とか「ああ，そうか」と思うようなことがありましたか
- 7) 友達と見合ったり，教え合ったりできましたか
- 8) 学習資料を活用できましたか
- 9) 安全に気をつけて練習できましたか
- 10) めあて①のできばえ
- 11) めあて②のできばえ

### (5) 指導者（担任教師：教職歴13年目の34歳女性教師）による評価方法

全被験者（学習者）6名について授業終了直後に自己評価項目と同一の11項目に同じ5段階尺度で回答。1時間当たりの延べ評価項目数は11項目×6名の計66項目。これを毎時間繰り返した。実験の目的から，言うまでもなく指導者は自己評価の結果を全く見ずに評価を行った。

### (6) 被験者（学習者）と指導者（教師）のコミュニケーション

指導者は学習者の自己評価票と自身の評価結果を十分に把握した上で，授業終了後，当日中の任意の時間にその日の学習について話し合う時間をできるだけ確保した。時間は概ね3～5分であり，各評価項目に対する学習者の反応（回答）自体に対する討論は避け，授業中の学習活動全般に関する討論を行った。各評価項目レベルでの討論を避けたのは，そのことが，指導者による評価の方向にそれ以降の自己評価結果を引き寄せる教示となる可能性をできるだけ回避しようとしたためである。なお，授業後に学習者と直接的にコミュニケーションできなかった場

合も、自己評価票に指導者がコメントを書き、少なくとも間接的なコミュニケーションは保証されるようにした。

## 2. 分 析

- (1) 評価活動を行った最初の2時間(2,3時間目)と最後の2時間(6,7時間目)の各項目(11項目×6名×2時間=132項目)に対する学習者の自己評価と指導者による評価の完全一致率, および準一致率を算出した。「完全一致率」は延べ132項目に対する, 学習者と教師による5段階の回答(評価)が完全に一致した項目の比率, 「準一致率」は完全一致した項目数に, 両者の5段階の回答(評価)が±1の範囲内にあった項目数を加え, 延べ132項目に対するその比率を算出したものである。
- (2) 統計的分析は正規分布検定による比率の差の検定によった。

## 3. 結 果

- (1) 完全一致率は最初の2時間が34.8%であり, 最後の2時間では72.7%に上昇した。両者の間には0.1%水準で有意な差( $Z_0=6.676$ )が認められた。
- (2) 準一致率は最初の2時間が75.8%であり, 最後の2時間では93.2%に上昇した。両者の間には0.1%水準で有意な差( $Z_0=4.024$ )が認められた。
- (3) 11の項目別に完全一致率を見ると, 「自分にあっためあてだったか」, 「練習の進め方はよかったか」, 「『あっ, わかった』とか『ああ, そうか』と思うようなことがあったか」, 「学習資料を活用できたか」の4項目, すなわち思考・判断や認識に関わる項目(最後の2時間の平均完全一致率: 60.4%)は, 他の項目(最後の2時間の平均完全一致率: 79.8%)よりも5%水準で有意に低い( $Z_0=2.335$ )という結果が得られた。
- (4) 75%以上の項目(11項目中8項目以上)で学習者の自己評価と指導者による評価が完全一致したのは, 最初の2時間が0名であり, 最後の2時間では4名(66.7%)であった。

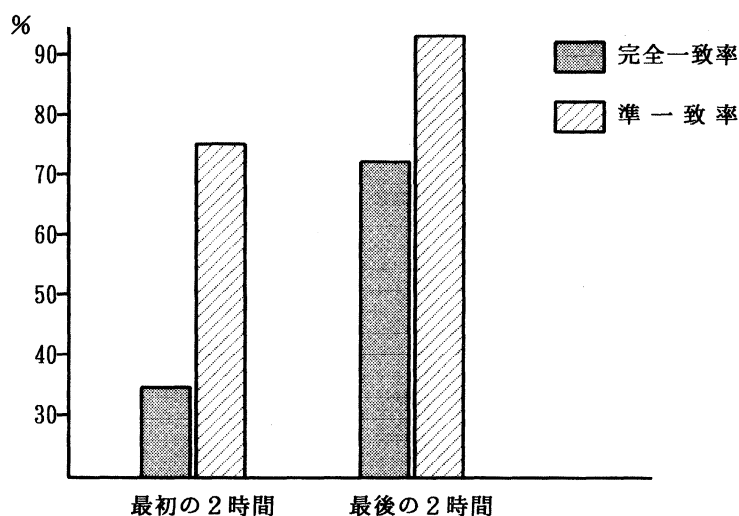


図1 学習者の自己評価と教師による評価の一致率

- (5) 90%以上の項目(11項目中10項目以上)で学習者の自己評価と指導者による評価が準一致したのは, 最初の2時間が2名(33.3%)であり, 最後の2時間では5名(83.3%)であった。

以上の実験的授業研究の結果は, 次のことを示唆している。

- (1) 授業後に学習者と教師が密度の濃いコミュニケーションを成立させることによって, 両者の

