

## 保健授業評価尺度の状況分析と課題

山合 洋人<sup>1) 2)</sup>

キーワード：保健授業、授業評価、レビュー、尺度構成項目

### I. 緒言

「健康」は人々の最大の関心事の一つであろう<sup>1)</sup>。その健康を保持増進するために、児童生徒がリテラシーを培う場のひとつとして学校教育が挙げられ、現代社会における多様化、深刻化した児童生徒の心身の健康課題への対応・解決に向けて、学校や教職員等の果たすべき役割は大きい<sup>2)</sup>。学校教育の中でも「健康」を対象や事象として真正面から捉え、授業として広く展開を図ろうとする保健授業<sup>注1)</sup>が重要であることは言うまでもない。しかし、かつては「保健」軽視の風潮がみられ、晴れた日は「体育」、雨が降った日は仕方なく「保健」という体育の代替的役割を担い、「アメフリ保健」(rainy days lessons) や「暗記保健」と揶揄されたりもした<sup>3)</sup>。そうした背景から現在もなお、学校現場においては体育授業に比べて保健授業への苦手意識を持つ教師が少なくなく<sup>4)5)</sup>、他教科に比べ授業研究、評価および教師教育等に関する研究においても、質・量ともに低調といわざるを得ない状況である<sup>3)6)7)</sup>。特に授業実践の質の低さは、児童生徒の保健授業に対する情意面に影響を及ぼしうる看過できない課題である<sup>8)</sup>。

授業実践の質を向上させたいと教師が考えた時、一般的には「実践の省察」に焦点があてられる。その際、教師が主観に基づき実践の省察を行うと、実践に関する問題点を正確に把握することができず、生徒の思考や関心から遊離したものになる可能性がある<sup>9)</sup>。そのため、省察の場面においては授業者の視点と生徒の視点の双方を往還することが必要であると指摘されている<sup>10)</sup>。実践の省察に生徒の視点を反映させる簡易的なツールの一つとして「授業評価アンケート(以下、授業評価)」が挙げられ、その機能について田口<sup>11)</sup>は以下の5つに要約している。

1. 意識改革の起爆剤
2. 授業改善の指針
3. 学生と教員間、あるいは教員相互のコミュニケーションツール

4. 優れた教育者を評価する、あるいは教員が自身の教育業績を示すための証拠資料
5. 教育の質保証などアカウントビリティのための根拠資料

授業改善を目的とした場合、上記の1、2、3の機能が寄与すると考えられ、実際に授業評価を通じて生徒の意見の一部が採り入れられ、授業が改善されていった事例がこれまでにいくつか報告されている<sup>12-14)</sup>。一方で、授業評価が授業改善の目的のために正しく機能しているか、授業評価によって授業の改善が促進されているのか<sup>15)</sup>、という疑問の声もあげられている。その背景には、授業評価の結果をどのように解釈し、どのような理論的枠組みに基づいて活用すればよいのかに関する戸惑いや混乱が存在する<sup>16)</sup>。安岡<sup>17)</sup>は授業評価に関する研究レビューを通じて、授業評価が授業改善に結びついているものは少ないと指摘しており、保健科教育における授業評価もその例外ではないだろう。しかしながら、保健科教育において授業改善のための授業評価に関する議論はこれまでほとんど行われておらず、レビュー論文も存在しない。このことは同研究領域が抱える大きな課題の1つであるといえる。そこで本研究では、国内学術雑誌に掲載された保健科教育に関する授業評価尺度開発を目的とした文献の体系的レビューを通して、既存の保健授業評価尺度の概要を明らかにすることを目的とした。また、そこからさらに一歩進めて、各保健授業評価尺度がどのような評価項目で構成され、保健授業の何を評価しようとしているのかといった点について構造的な視点で分析するために、各尺度構成項目の分類を行い、既存の保健授業評価尺度の課題や問題点を明らかにすることを目的とした。

### II. 方法

#### 1. 文献の検索方法

教育の在り方はその国の文化や社会情勢等に応じ

1) 東京学芸大学大学院連合学校教育学研究科

2) 筑波大学附属駒場中・高等学校

て変遷するものであり、授業評価尺度はその教育や教科の在り方等の影響を受けるため、本研究では国内の研究に焦点化して文献を検索した。文献検索の範囲は2020年8月までに日本の学術雑誌に掲載された保健授業評価尺度の開発に関する文献および、それらの文献における引用・参考文献リストとした。文献の検索方法には、国内の主要オンラインデータベースであるCiNii (Articles) と J-Stage を用いた<sup>18)</sup>。保健授業評価尺度の開発に関する文献を可能な限り網羅するため、検索キーワードとして「保健授業」について「保健授業、保健の授業、保健科教育、科目保健」の4種を用いた。同様に「評価尺度」について「評価尺度、評価票、評価シート、評価カード、授業評価」の5種を用い、「開発」については「開発、作成」の2種を用いた。検索式は(「保健授業」OR「保健の授業」OR「保健科教育」OR「科目保健」) AND (「評価尺度」OR「評価票」OR「評価シート」OR「評価カード」 OR「授業評価」) AND (「開発」OR「作成」) とし、計40通りの検索を実施した。この結果、重複も含め CiNii では13件、J-Stage では70件がヒットした。さらに、これらの文献における引用・参考文献リストから4つの文献を追加し、合計87文献をスクリーニングの対象とした。

## 2. 文献採択基準

文献採択基準には第一に「研究対象者が小学生から大学生のいずれかである」、第二に「1回の保健授業を形成的に評価するための授業評価尺度の開発を目的としている」の2点を設定した。尺度開発に関連のない文献および重複する文献や、他の研究において調査や分析の方法、開発過程、結果の信憑性に欠けると指摘されている文献、学会発表要旨、新聞等の記事は除外した。

## 3. 各尺度の信頼性及び妥当性の検証

先述の過程で採択された授業評価尺度について検討するために、信頼性と妥当性について本研究でどのように考えるのかを整理した。

### (1) 信頼性

信頼性とは測定値の安定性や一貫性を意味し<sup>19)</sup>、再現性、等価性、内的整合性といった側面を有している。Hogan ら<sup>20)</sup>は信頼性係数の報告状況レビューを通じて、 $\alpha$ 係数(内的整合性)を信頼性係数として報告した論文が最も多い(66.5%)ことや、1つの

信頼性係数によって信頼性を報告している論文が75%であったことなどを報告している。また、高本<sup>21)</sup>は2001~2013年度までに公刊された国内の心理尺度作成論文189本のうち、 $\alpha$ 係数を信頼性係数として報告した論文が168本であったことを明らかにし、国内においても信頼性係数の報告には $\alpha$ 係数が一般的に用いられていることを報告している。以上のことから、本研究における「信頼性の検証」は $\alpha$ 係数に拠るものと定義し、 $\alpha$ 係数を基に尺度の信頼性を検証しているか否かを表1に示した。なお、 $\alpha$ 係数に関する注意については南風原<sup>22)</sup>を参照されたい。

### (2) 妥当性の検証

妥当性とは尺度が測定したい対象を正確に測定できているかを意味しており、基準関連妥当性や構成概念妥当性、内容的妥当性などによって説明されてきた(妥当性の三位一体観)。しかし、1980年代以降の妥当性研究では、「構成概念妥当性」という考え方にすべての妥当性は収斂するという考え方が主流である(単一的な妥当性概念)。構成概念妥当性とは「測定結果の解釈の適切性の程度」であり、それを支えるための証拠は数も種類も測定目的や解釈によって異なることが指摘されている<sup>23)</sup>。これはすなわち、妥当性は一定の検証過程を経ることであるかないかを判定するような性質のものではないことを意味している<sup>24)</sup>。そこで本研究では、各尺度開発過程で用いられた統計解析の手法を表1に示した。なお、妥当性概念の詳細については村山<sup>23)</sup>を参照されたい。

## 4. 尺度構成項目の分析方法

各尺度の評価項目を内容により分類するために山合ら<sup>9)</sup>を参考に領域を設定した。山合らは授業構成要素である「教師・学習者・教材<sup>25)26)</sup>」の各要素をより詳細に評価するための下位概念を設定している。具体的に「教師」には「教授行為、学習目標の設定・明示、授業展開、学習者への対応や配慮」、「学習者」には「授業への意欲、授業態度、授業理解」、「教材」には「教科内容、教材の使い方、教材の工夫」を設定している。このような詳細な下位概念の設定を行っている研究は他に見当たらないことから、10下位概念を本研究の各分類領域とした。また、10下位概念に該当しない内容は「その他」とした。項目の内容が複数の領域に該当する場合は、最適と考える領域をメイン領域、その他の領域をサブ領域とした。

### III. 結果

#### 1. 保健授業評価尺度の概要 (表 1)

スクリーニング対象となった計 87 文献のうち、1 回の保健授業を形成的に評価するための授業評価票であると判断された文献は、岡崎ら<sup>27)</sup>、保健学習調査研究委員会<sup>28)</sup>、大津ら<sup>29)</sup>、文部省<sup>30)</sup>、白石ら<sup>31)</sup>、植田<sup>32)</sup>、七木田<sup>33)</sup>であった。そのうち、保健学習調査研究委員会および文部省はともに開発方法が明記されておらず、授業評価票の有用性が検証されていないため除外した。大津らについては、植田の「限られた時間内での使用を考えるとその簡便性はあまり高くなく、また情意以外の観点の評価は他で行わねばならないことなど、保健授業の評価票としては、限界がある」との指摘を支持したうえで、さらに、質問形式が問一答とは異なり、線結び評価シート<sup>29)</sup>の形式で評価項目の分類が困難であることから除外した。以上の過程を経て最終的に 4 件の文献を本研究の対象とし、各授業評価尺度の概要を表 1 に示した。

4 つの文献に示された評価尺度は 1992 年から 2002 年にかけて開発されていた。白石らを除く 3 尺度が複数の下位尺度 (4 から 8 下位尺度) で構成され、項目数は 4 項目から 24 項目であった。評価形式はいずれの尺度も児童生徒による自記式質問紙形式であり、白石らが 5 件法、その他の 3 尺度が 3 件法で回答を求めている。

評価項目の設定に関する根拠として、岡崎らと白石らは保健授業に対する生徒の感想を基にしつつ、体育授業評価に関する文献である小林<sup>34-37)</sup>から保健授業に援用できる項目を追加していた。植田は保健学習調査研究委員会及び大津ら<sup>29)38)39)</sup>、その他保健学習の評価研究<sup>40)41)</sup>の成果、体育授業研究<sup>42)</sup>で明らかとなった評価観点、ならびに文部省が示す観点別評価項目等を基に評価項目を設定していた。七木田は過去の授業評価研究<sup>43-47)</sup>及び体育科の一連の評価研究<sup>42)46-50)</sup>、保健科の評価研究<sup>27-29)32)39)</sup>等を基に評価項目を設定していた。以上のように、4 尺度すべてにおいて体育科教育研究分野の成果や授業評価尺度が援用されていることが判明した。

尺度を作成するための予備調査対象者数は、岡崎らが 141 名、白石らが 346 名、植田が 924 名 (有効回答数 770)、七木田が 4,621 名であった。

評価項目を抽出するための分析方法は、白石らが

クラスター分析、その他の 3 尺度は探索的因子分析を実施していた。岡崎ら、植田は因子分析における抽出法を明記しておらず、七木田は「主成分分析」と記していた。回転は 3 尺度いずれもバリマックス回転を指定しており、岡崎ら、植田は因子負荷量 0.4 未満、七木田は 0.6 未満の項目を削除の対象としていた。

信頼性の検証は、植田、七木田の 2 尺度で  $\alpha$  係数を基に実施されていた。妥当性について、岡崎らは項目分析 (上位下位分析法) および探索的因子分析を実施していた。白石らは項目分析 ( $\chi^2$  検定) およびクラスター分析を実施していた。植田は項目 - 全体相関および探索的因子分析、外部変数との相関係数の算出を実施していた。七木田は項目分析 (項目相互相関) および探索的因子分析、重回帰分析を実施していた。本研究では各尺度開発過程で用いられた統計解析の手法を示したにすぎず、これらの統計解析によって尺度の妥当性が検証できているか否かについて検討するものではない。なお、項目 - 全体相関や探索的因子分析は、妥当性を検証する手法というよりも「妥当性を高める手続き」とされている<sup>22)</sup>。

#### 2. 評価項目の分類結果 (表 2、表 3)

4 尺度合計 58 項目を山合らを基に設定した 10 領域に分類した。評価項目には、教師の教授行為や授業展開の仕方、学習者への対応や配慮を評価する項目や、学習者の授業への意欲、授業態度、授業理解、授業内容に対する認識について評価する項目があった。「その他」には、授業を総合的に評価する項目が分類された。各領域に分類された具体的な項目内容を表 2 に、評価項目の分類結果を表 3 に示した。その際、サブ領域は項目ごとに集計し、( ) 内に示した。各領域の分類結果の特徴は以下の通りである。

##### (1) 教授行為 (2 項目)

岡崎らに設定されており、教師の板書、話し方について評価する項目であった。

##### (2) 学習目標の設定・明示 (0 項目)

4 尺度すべてにおいて該当する項目は存在しなかった。

##### (3) 授業展開 (3 項目)

岡崎らに設定されており、「ディスカッションを多くやりたい」、「ひとつの内容に時間をかけてじっくり取り組むのがよい」といった授業展開に対する生徒の意見が反映されるような項目が分類された。

表1 各保健授業評価尺度の概要

タイトル・著者 (開発年)	対象者	尺度構成・ 評価形式	項目設定の根拠	開発方法	信頼性	信頼性・妥当性の検証 妥当性
保健授業における授業診断法の作成の試み - 高校生を対象に - (1992) 岡崎勝博ら	高校生	8下位尺度、 24項目、3件法	保健授業に対する自由記述アンケート及び小林 <sup>34-36)</sup>	上位下位分析法 (リカート方式) で各項目の適合性を確認後、因子分析。(抽出法不記載、固有値1基準、バリマックス回転、因子負荷量0.4以上の項目を採用)	×	項目分析(上位下位分析法) 探索的因子分析 <sup>※1)</sup>
保健授業の生徒による評価の研究 (1998) 白石龍生ら	高校生	1クラスター、 4項目、5件法	保健授業に対する感想文、小林 <sup>37)</sup>	27項目をクラスター分析し、保健授業の印象を4クラスターに分類 (生徒の学習過程、教師の教授過程、生徒の能動的な思考、生徒の受動的な思考)。「生徒の能動的な思考」に該当する項目から4項目を選定し、保健授業評価票としている。	×	項目分析 ( $\chi^2$ 検定) クラスター分析
小学校保健授業の教授 - 学習過程評価の開発 (1998) 植田誠治	小学生	4下位尺度、 16項目、3件法	先行研究や観点別評価項目	27の質問項目について項目得点、合計得点との相関係数を算出。その後、因子分析 (抽出法不記載、固有値1基準、バリマックス回転、因子負荷量0.4以上の項目を採用) を行い、より簡便性を高めるために各因子から4項目ずつを選定し、16項目からなる評価票を作成。	○	項目 - 全体相関 <sup>※1)</sup> 探索的因子分析 <sup>※1)</sup> 外部変数との相関
保健授業評価票作成の試み - 中学生の授業評価構造に着目して (2002) 七木文彦	中学生	5下位尺度、 14項目、3件法	先行研究	27の質問項目について項目得点、項目間の相関係数を算出し、3項目を削除。残る24項目について因子分析 (主成分分析、固有値1基準、バリマックス回転、因子負荷量0.6以上の項目を採用) を行い、5因子14項目からなる評価票を作成。	○	項目分析 (項目相互相関) 探索的因子分析 <sup>※1)</sup> 重回帰分析

※1: 妥当性を検証する手法というよりも「妥当性を高める手続き」とされている<sup>22)</sup>。

#### (4) 学習者への対応や配慮 (2 項目)

岡崎らに設定されており、学習者の意見や発言に対する教師の対応について評価する項目が設定されていた。

#### (5) 授業への意欲 (10 項目)

植田と七木田にそれぞれ 5 項目ずつ設定されていた。これらは主に学習者の授業に対する積極性や能動性を評価する項目で、項目内容は 2 つの尺度間で共通していた。

#### (6) 授業態度 (8 項目)

植田と七木田にそれぞれ設定されており、授業内における学習者間の協調性を評価する項目が分類された。植田の「友だちと助けあって学習できましたか。」と「友だちから教えてもらったり、助けてもらったりしましたか。」や、七木田の「友だちから教えてもらったり、助けてもらったりした。」と「友だちと助け合って、学習できた。」の項目のような、類似あるいは酷似した質問項目が同尺度内に存在することが判明した。

#### (7) 授業理解 (15 項目)

全領域の中で最も項目数が多く、4 尺度すべてに設定されていた。授業を通して学習した内容や、気付きや発見の有無、認識の変容等を確認する項目が分類された。

#### (8) 教科内容 (15 項目)

「授業理解」領域と同様に全領域の中で最も項目数が多く、そのすべてが岡崎らの尺度に設定されていた。保健授業と健康の関係性について評価する項目が多いことが特徴である。学習者の「授業内容」や「学習内容」に対する認識や、保健授業を通じた自身の変容等について評価する項目が分類された。

#### (9) 教材の使い方 (0 項目)

4 尺度すべてにおいて該当する項目は存在しなかった。

#### (10) 教材の工夫 (0 項目)

4 尺度すべてにおいて該当する項目は存在しなかった。

#### (11) その他 (3 項目)

授業を総合的に評価する項目が分類された。

明らかにすることを目的とした。本研究における基準を満たした 4 尺度の全評価項目のうち、94.8%の項目が山合らを基に設定した 10 領域のいずれかに分類された。しかし、各領域における分類結果を参照すると、分類された項目数に偏りが目立ち、項目が存在しない領域も散見された。尺度を用いた評価では、予め定められた項目のみを評価対象とするため、評価項目の設定は尺度の妥当性に大きく関わる<sup>51)</sup>。そのため、項目数の偏りは既存の保健授業評価尺度が抱える大きな課題の一つであるといえる。具体的に、「(1) 教授行為」に分類された項目は、「板書をきっちりまとめてほしい。」「もっと系統だてて話をしてほしい。」の 2 項目である。教授行為とは、準備した教材を用いて子供たちに分かりやすく説明をしたり、興味を引くような発問をしたり、明確な指示を出したり、評価を下したり、見やすい板書をしたりすること<sup>52)</sup>であり多岐にわたる。そのため、2 項目から得られる情報では「教授行為」を十分に評価することは難しく、より多くの項目設定が必要であろう。また、「(2) 学習目標の設定・明示」、「(9) 教材の使い方」、「(10) 教材の工夫」に分類された項目は存在せず、既存の保健授業評価尺度では評価することができない領域であることが明らかとなった。一方で、「(7) 授業理解」と「(8) 教科内容」に分類された項目はそれぞれ 15 項目、次いで「(5) 授業への意欲」が 10 項目、「(6) 授業態度」が 8 項目と多く、これらの領域を評価するための十分な項目数が設定されているといえるだろう。以上のことから、既存の保健授業評価尺度は学習者の授業内容に対する認識や、授業に対する意欲や態度、理解の状況を学習者自身あるいは教師が確認、判断するうえで有益な示唆を与えるが、教師の教授・指導のあり方や、教材における工夫やその使い方などに関する項目が少なく、教師の授業改善や授業力向上に寄与するフィードバック情報が十分に得られない現状であるといえる。このような評価項目の偏りが生じた要因の一つは各尺度の開発過程にあると推察される。尺度開発の過程で因子分析を行っている論文においては、分析の過程で脱落した項目や抽出されなかった因子が存在し、それらの中に「教授行為」や「学習目標の設定・明示」、「教材の使い方」、「教材の工夫」を評価することができる内容が含まれていた可能性がある。その場合、学習者の評価構造にそれらの評価項目が含まれない総合的な、または授業の一回性の

## IV. 考察

本研究では文献レビューと尺度構成項目の分類を通して、既存の保健授業評価尺度の概要と問題点を

表2 各領域に分類された具体的な項目内

領域	項目内容	著者
教授行為	保健の授業では、板書をきっちりまとめてほしい。 保健の授業ではもっと系統だてて話をしてほしい。	岡崎ら
授業展開	授業ではもっとディスカッションを多くやりたい。 保健の授業では、友達の見聞を聞くことができる。 ひとつの内容に時間をかけてじっくり取り組むのがよい。	岡崎ら
学習者への対応や配慮	保健の授業では自分の言いたいことが言える。(授業展開) 授業では生徒の見聞をもっと聞いてほしい。	岡崎ら
授業への意欲	むちゅうになって勉強することができましたか。 「もっと知りたい、もっとしらべたい」と思うことができましたか。 もっとつづけて勉強したいと思えましたか。 自分からすすんで勉強することができましたか。	植田
	「～を知りたい」、「～をはっきりさせたい」と思いながら勉強することができましたか。 せいっぱい、いっしょうけんめい勉強することができた。 むちゅうになって、勉強することができた。 自分からすすんで、勉強することができた。 「もっと知りたい」、「もっと調べたい」と思うことがあった。 「～を知りたい」、「～をはっきりさせたい」と思いながら、勉強することができた。	
授業態度	友だちと助けあって学習できましたか。 友だちの意見をきいて、いっしょに考えることができましたか。 友だちから教えてもらったり、助けてもらったりしましたか。 友だちと力をあわせてなかよく勉強することができましたか。	植田
	授業中、わからないことや、疑問に思うことがあるときは、自分で調べたり、質問したりしましたか。(授業への意欲) 今日の勉強に興味をもち、ほかの関係することについても、調べてみようと思った。(授業への意欲) 友だちから教えてもらったり、助けてもらったりした。 友だちと助け合って、学習できた。	
授業理解	保健の授業で「あつ、そうか」「あつ、わかった」と思うことがある。 保健の授業で、これから「自分もこうやってみよう」と考えることがある。	岡崎ら
	授業の中で新しい発見がありましたか。 授業の中で多角的な考え方ができましたか。(その他) 生活で役立つような内容がありましたか。(教科内容)	白石ら
	きょうの勉強でこれからの生活に役にたつことがあると思えましたか。(教科内容) 「アッ、ワカッタ」とか「アア、ソウカ」と思ったことがありましたか。	植田
	新しい発見やおどろきがありましたか。 自分の意見をもつことができましたか。 健康に役立つことを勉強した。(教科内容) 健康的に生活していくには、「こうすればいいのだな」と気づいたことがあった。 今日勉強したことは、これからの生活にいかすことができるだろう。(教科内容) 「知っていたこと」が、実はちがっていた。 「わかっている」と思っていたことが、実はわかっていなかった。 意外な事実を知った。	七木田
教科内容	保健の授業で取り上げる内容はおもしろい。 保健の授業は、「覚える」科目ではないのでおもしろい。 保健の授業では、健康についてまじめに考えられる。 保健の授業で、社会問題に興味をわいた。(授業理解) 保健の授業では、一般的な教養がついた。(授業理解) 保健の授業内容はためになることが多い。 保健の授業で、新しい価値観が育った。(授業理解) 保健の授業では、事実が掘り下げて示されるので知識が生きてくる。 保健の授業により、現在の健康問題に興味・関心が湧くようになった。(授業理解) 保健の授業で、自分の健康観が変わった。(授業理解) 保健の授業により健康問題を多角的な視点で見ることができる。(授業理解) 保健の授業では、もっと実践的なことを取り上げてほしい。 過去の健康問題ではなく、将来に生きる健康問題を教えてほしい。 保健の授業では、健康に生きていくための理論を学ぶことができる。 保健の授業で、民主主義の矛盾を健康という観点から見ることができる。(その他)	岡崎ら
その他	授業は、興味深かったですか。	白石ら
	きょうの勉強は、たいせつなことがらだと思えましたか。 きょうの勉強は、楽しかったですか。	植田

注) ( ) はサブ領域を示す。

表3 各尺度における評価項目の分類結果

著者名	各下位概念に該当する項目数												各尺度の 合計項目数	
	教師						学習者			教材				その他の 項目
	教授行為	学習目標の 設定・明示	授業展開	学習者への 対応や配慮	授業への 意欲	授業態度	授業理解	教科内容	教材の 使い方	教材の工夫				
岡崎ら	2	0	3	2	0	0	2 (6)	15	0	0	0	0	0	24
白石ら	0	0	0	0	0	0	3	0 (1)	0	0	0	0	1 (1)	4
植田	0	0	0	0	5 (1)	5	4	0 (1)	0	0	0	0	2	16
七木田	0	0	0	0	5 (1)	3	6	0 (2)	0	0	0	0	0	14
各領域の 合計項目数	2	0	3	2	10 (2)	8	15 (6)	15 (4)	0	0	0	0	3 (1)	58

注) ( ) はサブ領域数を示す。

評価が関係している可能性が考えられる。他方で、評価項目の収集段階においてリストアップされた項目に既に偏りが生じていた可能性も考えられる。中川<sup>5)</sup>が「教授学的にみて両教科(保健と体育)に大きな隔りがある」と指摘する一方で、4尺度の項目収集段階に着目すると、体育科教育研究分野の成果や授業評価尺度が多く援用されていることが判明した。このことにより、例えば「教授行為」に関連する項目については、因子分析を行っている岡崎、植田、七木田の論文のうち、植田、七木田では因子分析前の項目収集段階でリストに挙がっていない。学習者がどのような授業を望ましいと考えているかについては、これまでに多変量解析等を用いて授業評価尺度の因子構造について検討した研究や、総合評価と各設定項目との関連について検討した結果から、「興味ある授業内容」、「基本的な授業技術」、「学生とのコミュニケーション」など、ある程度一致した見解が示されており<sup>16)</sup>、授業改善を目的とした新たな保健授業評価尺度の開発を試みる際には評価項目の収集段階において偏りが生じないように、より広い視野で他研究領域の知見について目を向けていく必要があるだろう。

また、同尺度内に類似あるいは酷似した質問項目が存在する点も看過できない。なぜなら、類似項目の存在は尺度の妥当性を低下させる一要因となり得るためである。尺度の妥当化のためには、測定すべき構成概念と、それに含まれる要素や下位領域を明確に定め、それに沿って項目の作成・選択を行い、専門家の判断や、予備調査における予測すべき変数や弁別すべき変数との関連などをもとに項目を精選すること<sup>22)</sup>が必要である。

さらに、先述の通り、本研究の対象となった4つの評価尺度は1992年から2002年にかけて開発されている。教育モデルや専門家の合意は、その教育施設の特徴や社会、文化、時代の影響を受けやすく、また、何を評価するかは教育の目標と密接に結びついていることから<sup>51)</sup>、既存の保健授業評価尺度では現在の学校教育現場で行われる授業実践を適切に評価することができない可能性がある。そのため、対象に適した新たな保健授業評価尺度を開発し、授業実践を適切に評価・分析・改善していくことが求められるが、本研究の結果がその足掛かりとなることを期待する。

## V. 本研究の限界と課題

本研究の限界として研究対象の少なさが挙げられる。この点については本研究の文献採択基準が厳密であったことが要因であるというよりも、既存の保健授業評価尺度の総数が少ないことが要因であると結論付けることが妥当であろう。先述の通り、開発方法が不明で有用性が検証されていないために本研究の対象外となった保健授業評価尺度が存在するように、評価尺度の開発は信頼性や妥当性といった客観性を担保するためのプロセスを伴うために容易ではなく、このことが保健授業評価尺度の総数の少なさの一要因であるといえるだろう。一方で、本研究の対象となった保健授業評価尺度においては評価可能な項目に偏りが存在し、得られる情報が部分的であり、それを基に授業改善を図るには限界がある。これは客観性を担保するために行われた尺度開発の研究が、実践を評価する上で直面する構造的な課題の一つであるといえる。この課題解決に向けて新たな保健授業評価尺度の開発が必要であるのはもちろん、尺度開発だけに終始せず、それを活用あるいは他のビデオや発話分析等と併用して授業実践者にフィードバックしていく方法について検討していくことも今後は必要であろう。

本研究の課題として項目分類を筆者一人で実施している点が挙げられる。すなわち、本研究における項目分類は主観的分類であることを意味しており、今後の継続研究により項目分類の客観性を高めていく必要があると考える。

最後に、本研究では既存の保健授業評価尺度の概要と問題点を明らかにしたが、先述の「授業評価アンケートが授業改善の目的のために正しく機能しているのか<sup>15)</sup>、という疑問」に関する考察には至っていない。本件については、今後、保健授業評価尺度を適用した授業研究が蓄積されていく中で検討していくべき課題であると考えられる。

### 注

- 1) 小学校体育科「保健領域」、中学校保健体育科「保健分野」、高等学校保健体育科「科目保健」の授業を総称して「保健授業」と示した。

### 文 献

- 1) 厚生労働省, 健康日本 21 (総論),

[https://www.mhlw.go.jp/www1/topics/kenko21\\_11/s0.html](https://www.mhlw.go.jp/www1/topics/kenko21_11/s0.html), 2021.1.28.

- 2) 栗原淳 / 教員養成系大学保健協議会編: 学校保健ハンドブック(第7次改訂)「学校保健を担う教職員等の責任」. 22, ぎょうせい, 東京, 2019.
- 3) 長岡知 / 今村修ら編著: 保健科教育学の探究 - 研究の基礎と方法 - 「保健科の教師教育論② 教育実習の現状と課題」. 58-62, 大修館書店, 東京, 2020.
- 4) 赤田信一ら: 中学校保健体育教師を対象とした保健授業の実施に関する調査研究(第1報) - 保健授業をより円滑に実践するための教師のニーズと、そのサポート体制の構築に向けて -. 静岡大学教育学部研究報告(人文・社会科学篇) 第51号: 133-143, 2001.
- 5) 中川明ら: 保健体育の教師はなぜ保健の授業が苦手なのか? - 保健体育科専攻学生の教材・内容・授業観からの検討 -. 愛知教育大学保健体育講座研究紀要, 35: 63-73, 2010.
- 6) 野津有司 / 今村修ら編著: 保健科教育学の探究 - 研究の基礎と方法 - 「保健科教育研究の必要性」. 7-11, 大修館書店, 東京, 2020.
- 7) 山合洋人 / 今村修ら編著: 保健科教育学の探究 - 研究の基礎と方法 - 「実践研究の展開④ 公開授業研究」. 283-288, 大修館書店, 東京, 2020.
- 8) 日本学校保健会, 平成 28 年度保健学習推進委員会報告書 - 第3回全国調査の結果 -, <https://www.gakkohoken.jp/books/archives/201>, 2021.5.1.
- 9) 山合洋人ら: 学習者による保健授業の主観的評価尺度の開発 - 教師へのフィードバックに向けて -. 学校教育学研究論集, 44, 43-54, 2021.
- 10) 国立教育政策研究所, 主体的・対話的で深い学びを実現する授業改善の視点について, [https://www.nier.go.jp/05\\_kenkyu\\_seika/pdf\\_seika/r02/r020603-01.pdf](https://www.nier.go.jp/05_kenkyu_seika/pdf_seika/r02/r020603-01.pdf), 2020.11.26.
- 11) 田口真奈 / 山地弘起編著: 授業評価活用ハンドブック「授業評価の諸機能」. 31-51, 玉川大学出版部, 東京, 2007.
- 12) 飛内文代ら: 生徒の視点を生かした授業評価についての研究 - 高等学校における授業改善への活用の可能性 -. 青森県総合学校教育センター平成 15 年度研究紀要, 1-13, 2003.
- 13) 笹田茂樹: 「生徒による授業評価」に関する一考察 - 正則高等学校の事例を中心として -. 人間発達科学部紀要第4巻第1号, 21-28, 2009.
- 14) 新子泰夫: 高等学校における「生徒による授業アンケート」の授業改善への有用性. 奈良県教育研究所平成 22 年度研究収録 18 号, 1-10, 2010.
- 15) 松谷満ら: 全学共通教育の現状と課題 - 学生による授業評価アンケート調査の分析から -. 徳島大学高等教育研究センター大学教育研究ジャーナル, 2, 13-25, 2005.
- 16) 澤田忠幸: 学生による授業評価の課題と展望. 愛媛県立



- 医療技術大学紀要.7-1, 13-19, 2010.
- 17) 安岡高志：学生による授業評価の進展を探る. 京都大学高等教育研究.13, 73-87, 2007.
  - 18) 伊藤央二ら：国内スポーツツーリズム研究の系統的レビュー. 体育学研究, 62-2, 773-787, 2017.
  - 19) 村上宣寛：心理尺度のつくり方. 北大路書房, 京都, 2006.
  - 20) Hogan et al. :Reliability methods :A note on the frequency of use of various types. Educational and Psychological Measurement, 60, 523-531, 2000.
  - 21) 高本真寛ら：国内の心理尺度作成論文における信頼性係数の利用動向. 心理学評論, 58 (2), 220-235, 2015.
  - 22) 南風原朝和：尺度の作成・使用と妥当性の検討. 教育心理学年報, 51, 213-217, 2012.
  - 23) 村山航：妥当性 - 概念の歴史の変遷と心理測定の観点からの考察 -. 教育心理学年報, 51, 118-130, 2012.
  - 24) 小塩真司：心理尺度構成における再検査信頼性係数の評価 - 「心理学研究」に掲載された文献のメタ分析から -. 心理学評論, 59 巻1号, 68-83, 2016.
  - 25) 森脇健夫ら：授業づくりと学びの創造. 37-87, 学文社, 東京, 2011.
  - 26) 白石龍生ら：科目保健の生徒による授業評価についての研究. 大阪教育大学紀要第V部門教科教育, 62(1), 71-77, 2013.
  - 27) 岡崎勝博ら：保健授業における授業診断法の作成の試み - 高校生を対象に -. 筑波大学附属駒場中・高等学校研究報告, 32, 57-69, 1992.
  - 28) 保健学習調査研究委員会：小学校保健の授業の改善をめざして - 体育科の保健領域 -. 日本学校保健学会, 165, 1993.
  - 29) 大津一義ら：情意形成の評価シート作成の試み - 高校における保健の授業 -. 体育科教育, 10, 58-61, 1994.
  - 30) 文部省：新しい学力観に立つ体育科の授業の工夫. 108-164, 東洋館出版, 東京, 1995.
  - 31) 白石龍生ら：保健授業の生徒による評価の研究. 日本健康教育学会誌, 5-1, 15-21, 1998.
  - 32) 植田誠治：小学校保健授業の教授 - 学習過程評価票の開発. 学校保健研究, 40, 75-81, 1998.
  - 33) 七木田文彦：保健授業評価票作成の試み - 中学生の授業評価構造に着目して -. 学校保健研究, 44, 47-55, 2002.
  - 34) 小林篤：体育の原理と実践 - 体育科教育原論 -. 杏林書院, 東京, 1986.
  - 35) 小林篤：体育の授業分析. 大修館書店, 東京, 1983.
  - 36) 小林篤：体育の授業研究. 大修館書店, 東京, 1979.
  - 37) 小林篤：授業分析法入門. 45-48, 明治図書, 東京, 1975.
  - 38) 大津一義：保健学習の評価. 現代保健学習指導事典, 44 - 53, 大修館書店, 東京, 1984.
  - 39) 大津一義ら：情意に着目した保健の授業研究. 学校保健研究, 31, 258-265, 1989.
  - 40) 数見隆生：事実に出発し事実を深める保健科の指導. 教授学年報 I, 81-82, 1975.
  - 41) 友定保博：保健授業の評価 - その考え方・進め方. 体育科教育, 41(9), 35-38, 1993.
  - 42) 高橋健夫ら：体育授業の「形成的評価法」作成の試み - 子どもの授業評価の構造に着目して -. 体育学研究, 39, 29-37, 1994.
  - 43) 渡辺秀敏：楽しい授業と学習効果. 児童心理, 16(11), 24-31, 1961.
  - 44) 梶田叡一：授業・教師・学校の評価. 教育評価第2版, 217-270, 有斐閣, 東京, 1992.
  - 45) 辰野千寿：授業活動の評価. 授業の心理学, 157-163, 教育出版, 東京, 1992.
  - 46) David R. Krathwohl et al. : The basis for affective domain classification. Taxonomy of Educational Objectives, 24-44, Longman, New York, 1964.
  - 47) Benjamin S. Bloom et al. : Handbook on Formative and Summative Evaluation of Student Learning. 117-140, McGraw-Hill, New York, 1971.
  - 48) 長谷川悦示ら：小学校体育授業の形成的評価票及び診断基準作成の試み. スポーツ教育学研究, 14, 91-101, 1995.
  - 49) 小林篤：態度測定による体育の授業診断. 体育の授業研究, 170-176, 大修館書店, 東京, 1978.
  - 50) Wear C L : The Evaluation of Attitude Toward Physical Education as Activity Course. The Research Quarterly, 22, 114-126, 1951.
  - 51) 常住亜衣子ら：医療面接における医師・患者間コミュニケーションスキル評価尺度 - 文献レビューと尺度構成項目の分析. 医学教育, 44(5), 335-344, 2013.
  - 52) 今村修 / 教員養成系大学保健協議会編：学校保健ハンドブック〈第7次改訂〉「保健の担当教員として求められる力量形成のために」. 96 - 101, ぎょうせい, 東京, 2019.

(2022年5月9日投稿受付, 2022年7月19日受理)