

総合的な探究の時間の評価に関する一考察 課題研究の評価についての校内標準化活動の実践報告

Review on Evaluation of Comprehensive Investigation Hours: Report on Challenges for Standardization of Evaluation of Task-Oriented Research in School

小林 廉 北岡和樹

要旨

本校では、5・6年生の全員を対象に実施する課題研究の評価体制の整備が問題となっている。そこで国際教養委員会では、2017年度から2018年度にかけて、まずは所産としての論文の評価の仕方に焦点を絞り、「校内での標準化」活動を実施した。本稿ではその実際を報告し、成果と課題について考察する。結果として、課題研究の評価規準に対する参加者たちの理解が深まったことが確認できた。一方で、研究の過程についての評価体制を整備していくことが急務である。

1. はじめに

我が国の高等学校において、課題研究を実施することの重要性がよりいっそう高まっている（例えば、文科省、2017、2018a）。本校では、本校独自の学習領域「国際教養」の一環として、後期課程での総合的な学習の時間（平成30年告示高等学校学習指導要領では「総合的な探究の時間」）に対応する「国際4」「国際5」「国際6」の時間において、以前から課題研究に該当する活動を行ってきた。その後スーパーサイエンススクール（SSH）やスーパーグローバルハイスクール（SGH）の指定を受けるなどして、2016年度からは課題研究の実施体制がより一層整備され、実施されるに至っている。2018年度現在、本校後期課程における課題研究の進行過程は以下の通りである。

- ① 4年次2学期下旬～3学期下旬…研究課題の設定と研究計画書の作成
- ② 5年次…研究の実施、3学期初旬に中間論文提出、その後発表、振り返り
- ③ 6年次…2学期中旬に最終論文提出

なお、4年生は2学期中旬まではPersonal Project（PP）と、課題研究への移行指導が行われる（詳しくは小林・廣瀬（2017）参照）。実際に研究が本格化するのは5年生からである。また、5年次・6年次における課題研究の対象生徒は一般コースのみであり、DPコースの生徒たちはこの時間TOKに取り組んでいる。

以上のように、課題研究の実施体制がより一層整備されてきている一方で、従来から問題とされてきたのが課題研究の「評価」である。研究の過程、そして所産としての論文をどう評価しフィードバックしていくかということについての体制がまだ整っていなかったのである。文科省（2018a）では、総合的な探究の時間の評価として、「具体的な生徒の姿を見取るに相応しい評価規準を設定し、評価方法や評価場面を適切に位置付けることが欠かせない」（p.139）ことや、「信頼される評価の方法であること」（同）などが挙げられている。これを、具体的には

課題研究を実施していくにあたってどう実現していくかが問題となる。おそらく、課題研究を実施している学校の中には、同じ点を問題としている学校が少なからずあると考えられる¹。

そこで、国際教養委員会²では、2017 年度から 2018 年度にかけて、まずは所産としての論文の評価の仕方に焦点を絞り、そのシステムの整備を試みてきた。とりわけ、「信頼される評価の方法であること」の実現のために設定したのが「校内での標準化」活動である。これは国際バカロレア（IB）のアイデアに基づく活動であり、「評価規準、到達度、およびそれらをどのように適用するかについての共通の理解を形成するための教師の話し合い」を指し、「教師は自分たちの判断の信頼性を高めることができ」とされているものである（国際バカロレア機構, 2016, p.96）。本稿では、課題研究の所産としての論文の評価についての「校内での標準化」活動の実際について、その成果と課題を含めた報告を行う。

2. 本校における課題研究の評価体制

まず、本校における課題研究の評価体制について確認しておく。

本校の課題研究の進行は、ISS チャレンジと呼んでいる校内課題研究コンテストと連動させている。このコンテストは SSH・SGH 指定校事業の一環として行われているものであり、1 年生から 6 年生の誰でも個人またはグループで任意で参加でき、参加者たちは指導教員（本校の教員）の指導を仰ぎながら約 1 年間かけて課題研究に取り組むことになる。参加する際は SSH 部門か SGH 部門を選ぶことになっており、最終的な所産である論文はそれぞれの部門の評価規準に照らして評価される。それぞれの部門で上位 4 チームは、本校 3 年生以上全員の前でプレゼンし、最終的に審査されることになっている。

5 年次における課題研究では、ISS チャレンジに参加しようがしていまいが、基本的には全員が同じ活動を行うことになる。すなわち、個人またはグループで SSH 部門か SGH 部門のいずれかを選択し、指導教員のもとで約 1 年間かけて課題研究に取り組み、中間論文という形で 3 学期初旬に論文を提出し、それぞれの部門の評価規準に照らして評価されることになる。ISS チャレンジとの違いには、それに参加すると金銭的な意味を含めた研究支援が受けられる場合があることや、論文の評価者が 2 名になることが挙げられる。参加しない場合は、指導教員 1 名のみが論文を評価することになる。いずれにせよそのフィードバックを受け、6 年次には、基本的には同じテーマで研究を続け、2 学期中旬に提出する最終論文が同じ評価規準で評価されることになる。

つまり、課題研究の所産である論文の評価は、指導教員が、ISS チャレンジ SSH 部門または SGH 部門の評価規準に照らして行うことになる。この規準の具体的中身については本稿末に資料として載せた。生徒には、この規準に照らした数値的評価と、指導教員によるコメントがフィードバックされることになる。

なお、4 年生から 6 年生までの総合的な学習（探究）の時間は週 1 時間、同一時間帯に設けており、ほぼ全校教員による指導体制を敷いて臨んでいる。5 年生・6 年生の課題研究を担当する教員間で指導する論文数なるべく同じになるように、またなるべく教員の専門性と生徒たちのテーマが合致するように、指導教員を決定している。

1 例えば 2017 年度に実施された第 59 回全国国立大学附属学校連盟高等学校部会教育研究大会の附属のあり方部会では、課題研究の評価が話題とされた。

2 本校独自の学習領域「国際教養」の計画立案および実施等を行う委員会。

3. 課題研究の評価についての校内標準化活動

3. 1 校内標準化活動の概要とその必要性

校内標準化活動とは、1章で述べたように、「評価規準、到達度、およびそれらをどのように適用するかについての共通の理解を形成するための教師の話し合い」（国際バカロレア機構、2016, p.96）を指す。課題研究の所産としての論文の評価においては、大きく2つの点でこの校内標準化活動が必要不可欠である。一つは、最終的に目指す生徒の具体的な姿を共有することである。それを事前につくっておくことが、普段の指導の充実につながる。もう一つが、評価の信頼性を高めることである。前章で確認したように、ISS チャレンジに参加しない限りは指導教員1名による評価となるため、その信頼性を高めておく必要性が大いにある。実際、今年度初めて本校で課題研究を指導することになる教員からは、評価についての不安の声が複数あがった。そうした実態も踏まえ、評価規準についての校内標準化活動を実施するに至ったのである。

3. 2 校内標準化活動の実施手順

では、校内標準化活動をどのように実施するか。国際バカロレア機構（2016）は、「自分たちの予想について議論し、さまざまな到達度を例示する生徒の成果物の実例を用いること」（p.96）を挙げている。これを踏まえ、次のように実施することにした。

表1 校内標準化活動の実施手順

- | |
|---|
| <p>① 生徒の成果物の実例として、過去のISS チャレンジにおける論文を用いる。当初は、2名の評価が高得点だったものと、割れたものの2つの論文を用いる予定だったが、時間の都合上、2名の評価が割れたもののみを用いることとなった。これらはSSH部門とSGH部門から1本ずつ選定した。なお、論文の選定については本校SSH委員およびSGH委員の協力を得た。</p> <p>② まず国際教養委員会のメンバーで標準化活動を行う。具体的には、各自が過去の評価を知らない状態で論文を評価規準に照らして評価し、国際教養委員会で話し合い、国際教養委員としての評価を一つ一つ定めていく。それを、校内標準化活動における一つのたたき台、すなわち「この論文ならこういう評価になると考えられる」という一つの姿として扱う。</p> <p>③ ISS チャレンジの指導教員を含む課題研究担当教員で標準化活動を行う。すなわち、事前にSSH部門かSGH部門を選んでもらい、それぞれに該当する論文を事前に評価してもらったうえで、部門ごとに集まり、グループおよび全体で話し合い、国際教養委員会としてのたたき台を改善し、「この論文ならこういう評価になると考えられる」という共通理解を図る。</p> |
|---|

さらに、上記の校内標準化活動について事後アンケートを実施し、その成果と課題を読み取ることとする。

3.3 校内標準化活動の実際

(1) 事前評価の実際

校内標準化活動の場には、国際教養委員会のメンバーを除くと、SSH 部門に 9 名の教員、SGH 部門に 15 名の教員が参加した。それぞれの事前評価（部門ごとに同じ論文を読んで各自が評価した結果）は次のとおりである。

表2 SSH 部門の事前評価の結果

	A	B	C-1	C-2	D-1	D-2	E-1	E-2	F	合計
教員SSH1	2	4	1	3	2	1	4	2	6	25
教員SSH2	2	2	1	1	1	1	2	2	2	14
教員SSH3	2	2	1	2	2	1	4	2	2	18
教員SSH4	2	0	1	4	1	0	2	2	2	14
教員SSH5	2	2	2	4	1	2	4	2	0	19
教員SSH6	2	2	1	3	1	1	2	1	2	15
教員SSH7	2	2	2	3	1	1	2	2	2	17
教員SSH8	2	2	1	3	2	1	4	2	4	21
教員SSH9	2	4	2	3	2	2	6	3	4	28

表3 SGH 部門の事前評価の結果

	A	B	C	D	E	F	G	合計
教員SGH1	2	1	1	1	1	1	2	9
教員SGH2	2	2	1	1	1	1	2	10
教員SGH3	3	2	1	1	2	2	2	13
教員SGH4	3	2	2	2	1	2	3	15
教員SGH5	3	3	1	1	1	1	2	12
教員SGH6	2	2	1	1	1	2	2	11
教員SGH7	2	1	1	1	1	1	2	9
教員SGH8	2	2	2	1	1	1	2	11
教員SGH9	2	3	1	1	1	2	2	12
教員SGH10	2	1	1	1	1	2	2	10
教員SGH11	2	1	2	1	1	1	2	10
教員SGH12	2	1	1	2	1	1	2	10
教員SGH13	2	2	1	1	1	1	2	10
教員SGH14	2	2	1	2	1	1	2	11
教員SGH15	2	2	1	2	1	1	2	11

1つの指標として、規準ごとの点数の合計に着目すると、SSH 部門で最大 14 点、SGH 部門で最大 6 点の開きがある。改めて、評価の信頼性を高めるために校内標準化活動を行うことの必要性が示されたといえる。なお、これら 2 つの評価規準は満点も間隔も異なるため比較できないが、あくまで参考として合計点でそれぞれの変動係数を求めてみると、SSH 部門が約 0.24,

SGH 部門が約 0.14 であり、SSH 部門の方が相対的にばらついているとみられる。それぞれの規準におけるばらつきについては後で言及する。

(2) 話し合いの実際

ここでは、筆者が実際に参加した SSH 部門の標準化活動に焦点を絞って、話し合いによって事前評価が大きく変わった事例を記述する。

SSH 部門の標準化活動における話し合いは、国際教養委員のメンバーも混じってまずは 3 名から 4 名程度のグループで行うこととし、それぞれのグループに評価規準を割り当て、グループとしての評価を決めるよう求めた。その後、全体で話し合うこととした。

ある 1 つのグループは、規準 F の「結論」について話し合っていた。この評価規準の説明は以下の通りである。

表 4 SSH 部門の評価規準 F「結論」の説明

到達度	レベルの説明
0	研究の目的に対して、適切な結論になっていない。
2	研究の目的に対し、得られた研究結果に基づいた結論を構築しているが、その論理には一貫性が欠けている。
4	研究の目的に対し、得られた研究結果に基づいて、結論を論理的に示している。さらに、これまでの研究方法の妥当性を評価し、今後の課題を明確に示している。
6	研究の目的に対し、得られた研究結果に基づいて、結論を論理的に示している。さらに、これまでの研究方法の妥当性を評価し、今後の課題や今後の展望（社会・学術面への貢献など）を具体的に示している。

グループは 4 名からなっており、事前評価では「2」が 2 名、「4」が 2 名であった。なお、国際教養委員会としての事前評価は「2」であった。ところが、この規準について話し合いを重ねていくうち、グループとしての結論は「0」に収束した。その理由としては、「レベルの説明」に、「研究の目的に対して」という文言が含まれていることに改めて着目したことが挙げられた。詳述すると、選定された論文では、一定の結論は書かれているものの、そもそも目的と方法がずれていて結論にギャップが生じてしまっているため、「研究の目的に対して」は適切な結論にはなっていないことが説明された。この説明が全体で共有されたとき、国際教養委員会も含めて他の教員もみな一様に納得した。確かにそのグループが説明した通りだったからである。規準 F の事前評価では、「0」としていた教員は 1 名のみであり、「6」とする教員もいるなどばらけた状態であったが、この話し合いを経て「0」という共通理解が図られたのである。

全体の話し合いでは、以上のようにして規準ごとの点数と説明が共有され、「この論文ならこういう評価になると考えられる」という最終的な共通理解が図られた。その点数は、国際教養委員会による事前評価より低いものとなった。

表 5 国際教養委員会による評価と最終的な評価との比較

	A	B	C-1	C-2	D-1	D-2	E-1	E-2	F	合計
国際教養委員会	2	2	2	3	2	1	4	1	2	19
最終決定	2	2	1	3	1	1	2	2	0	14

4. 校内標準化活動の成果と課題

校内標準化活動を実施してみたの成果と課題について整理する。その際、校内標準化活動の後に実施したアンケートの結果を参考にする。事後アンケートの回答数は、SSH 部門で 9 名中 4 名、SGH 部門で 15 名中 5 名であった。

第 1 の成果は、参加者たちの評価規準に対する理解が深まったと考えられることである。事後アンケートでは、「論文評価の校内標準化に参加されたことにより、評価基準についての理解は深まりましたか」という質問に対し、9 名全員が「かなり深まった」または「どちらかといえば深まった」という肯定的な回答をした。具体的な事例を挙げよう。SSH 部門に参加した教員の全体の話し合いの中で、基準 C-2「今年度の研究目的」の評価規準について質問が出された。

表 6 SSH 部門の評価規準 C-2「今年度の研究目的」の説明

到達度	レベルの説明
0	全く書かれていない。
1	研究目的が述べられているが、明確に説明されていない、または幅広すぎる。
2	今年度の研究活動において、何を明らかにするのか説明されている。
3	先行研究の調査または継続研究を踏まえ、今年度の研究活動において、何を明らかにするのか説明されている。
4	先行研究の調査または継続研究を踏まえ、今までに何が明らかになっていないのかということが明記されており、それをもとに今年度の研究活動において何を明らかにするのか説明されている。

質問は、「先行研究の調査または継続研究を踏まえ」とあるが、先行研究は何らか挙げられていればそれでよいのか、それとも本質的なものがしっかりと押さえられている必要があるのか、というものであった。これに対しては、評価者自身が評価対象となる研究テーマの先行研究を調べ、それが評価対象の論文でしっかりと挙げられているか見るべきであるという共通理解がなされた。つまり後者である。これは、生徒に求める姿を考えれば当然といえば当然だが、評価規準の文言だけみていると先行研究が何らか挙がっていればよいとも読めてしまうし、高校生としてこれが挙げられていれば十分といったように考える教員もいる。もちろん評価者である教員はその研究テーマの専門家ではないし、研究しているのは高校生であるから、どこまで先行研究を掘り下げられるのかという課題は残る。しかし、専門家ではない教員でも見つけられるような先行研究が挙げられていなければ、それは先行研究を踏まえたとは言えないだろうという共通理解が図られたのである。このような話し合いが他の評価規準についても行われた。このことが、参加者たちの評価規準に対する理解の深化を促したと考えられる。

第 2 に、評価規準の改善のきっかけとなる、評価基準自体の課題が表出したことである。上記のような話し合いは、評価者が評価規準を理解するにあたってどこに困っているのかの具体例を提供した。事前評価や話し合いの場でよく話題とされたのは、評価規準にある「わかりやすい」や「適切な」といった文言の「程度」であった。これらの判断はどうしても主観になるためブレが生じる。しかしそれは評価において避けられないことであり、だからこそこうした標準化活動がある。一方で、何をもって「わかりやすい」とするか、何をもって「適切」なの

かを評価規準においてより明確化することで、生徒と教員ともに目標とする姿の理解を促せる可能性もある。校内標準化活動の場は、そうした評価規準の改善に向けた生の声が出出る良い機会となる。

課題としては、所産としての論文の評価ではなく、過程の評価への必要性が高まったことが挙げられる。事後アンケートでは以下のような意見が寄せられた。

表7 事後アンケートで寄せられた意見

- ルーブリックは、生徒への指導を前提として作っているようで、それが「明確に」とか「適切な」などという表現に現れていると思います。この点については、生徒とのやり取りを踏まえないと、例えば「明確である」とか「明確でない」といった評価について、お互いが納得する状況になりそうもありません。だから、自分が論文を読んだだけでは、適正な評価になりそうもありません。
- 計画書や経過報告書をしっかりと作成し提出した、というプロセスの部分が、ある程度具体的に最終評価に反映される形がいいと思いました。
- メンター（注・指導教員のこと）が評価すると、論文に記載されていない研究過程のアレコレを知っているので、評価が甘くなってしまいそう。もしくは公平に評価ができないのでは？と思う部分があります。

これらはどれも、研究の所産ではなく過程へ言及した意見であるととらえられる。総合的な学習（探究）の時間に求められる本来的な評価（文科省、2018a）や、課題研究を通して育成できると期待できる資質・能力（例えば岡本、2017）を踏まえれば、研究の過程を適切に評価することが不可欠である。実際、ISS チャレンジのSGH 部門では、10月中旬に提出を義務付けている研究経過報告書の評価を指導教員が行い、フィードバックするシステムを確立させている。こうしたシステムを5年生6年生全員対象の課題研究に取り入れるなどして、研究の過程を評価する体制を整えていくことが急務である。

5. おわりに

本稿では、課題研究の所産である論文評価についての校内標準化活動の実際について報告した。こうした活動は一回行えばそれでよいという性格のものではなく、継続してこそ意味のあるものである。各学校における総合的な探究の時間の目標は、文科省（2018b）による総合的な探究の時間の目標とともに、各学校における教育目標を踏まえて定めるとされている（文科省、2018a）。総合的な探究の時間の評価規準について校内標準化活動を行うことは、評価規準の理解を深めて指導に反映させるだけでなく、学校として育成したい生徒の姿を共有することにつながると思われる。課題研究の過程の評価を含め、引き続き校内標準化活動を行っていきたい。

付記

課題研究の校内標準化活動は国際教養委員会が先導する形で行った。主としてSSH 部門を担当したのは北岡和樹であり、SGH 部門を担当したのが佐藤典幸である。他に、藤木正史、大橋典子、佐藤恵美、後藤葵が携わった。

引用・参考文献

岡本尚也（2017），『課題研究メソッド』，啓林館．

国際バカロレア機構（日本語版，2016），『MYP：原則から実践へ』．

小林廉，廣瀬充（2017），「Personal Project と課題研究に関する指導モデル試案－『前期課程での指導』と『PP および PP から課題研究への移行期の指導』に焦点化して」，『国際中等教育研究』，no.11，p.107 -121．

文科省（2017），「高大接続改革の実施方針等の策定について」

http://www.mext.go.jp/b_menu/houdou/29/07/1388131.htm

文科省（2018a），「高等学校学習指導要領解説 総合的な探究の時間編」

http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2018/07/13/1407196_21.pdf

文科省（2018b），「高等学校学習指導要領（平成 30 年 3 月告示）」

http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2018/07/11/1384661_6_1_2.pdf

（URL は全て平成 31 年 1 月 5 日最終確認）

Review on Evaluation of Comprehensive Investigation Hours: Report on Challenges for Standardization of Evaluation of Task-Oriented Research in School

Abstract

One of the pending matters of our school has been construction of an evaluation system of task-oriented research targeting fifth- and sixth-year students. In FY 2017 and 2018, the Global Liberal Studies Committee focused on how to evaluate theses as the deliverables at first and proceeded with “standardization in school.” Its actual progress is reported hereby for deliberation on its achievements and issues going forward. As a result of the challenges, we confirmed that the participants had deepened their understanding of the evaluation standard of task-oriented researches. On the other hand, it is urgently required to establish the evaluation structure of the research process.