（５）「課題研究」の評価の妥当性に関する考察

	生徒の自己評価					教員の評価			総合評価	総合評価
	(1) 研究 ノート	(2) 課題 発見力	(3) 知識習得 活用力	(4) 論理的 思考力	(5) 見通しを 立てる力	学びに 向かう力 人間性	思考力 判断力 表現力	知識 及び 技能	生徒評価 ５項目の 平均値	教員評価 ３項目の 平均値
全生徒	2.58	2.61	2.47	2.15	2.42	2.61	2.22	2.33	2.45	2.39
自己評価が低い生徒	1.29	1.14	1.07	1.00	1.00	1.71	1.57	1.57	1.10	1.62
自己評価が高い生徒	3.84	3.68	3.84	3.68	3.84	2.60	2.56	2.60	3.78	2.59
教員評価が低い生徒	2.19	2.33	1.96	1.93	1.89	1.56	1.11	1.19	2.06	1.28
教員評価が高い生徒	2.91	2.97	2.81	2.35	2.63	3.69	3.05	3.49	2.73	3.41

生徒には「自己評価シート」(P.○○を参照)を配付し、学期毎に自己評価を記入させた上で提出させた。一方教員は、ルーブリック（P.○○を参照）に基づき、サーバー上の入力シート（エクセルファイル）に直接生徒評価を入力していった。記入及び入力に際しては、生徒・教員共にＡ、Ｂ、Ｃの３段階の基準で評価を行った。上記表中の数値は、Ａ：４点、Ｂ：２点、Ｃ：１点として換算した数値を、項目毎に平均値をとったものである。従って、オールＣの評価では最低値となる 1.00 を、オールＡの評価では最高値となる 4.00 の値をとることになる。なお、未入力項目がある生徒のデータは予め省いた上で集計を行った。また、現在（２月）は、学年末の最終評価がまだ入力されていない段階であるため、ここでの考察は第２学期までの評価に基づいて行うものとする。

生徒と教員での評価項目が異なるため単純には比較できないが、上記表中の平均値を参考に項目毎のデータを比較してみることにする。まず、全生徒を対象に生徒の自己評価と教員の評価を比較すると、総合評価では緩やかな相関が見られた。つまり、自己評価の平均値が高い生徒は教員の評価も比較的高評価であり、低い生徒はその逆である傾向が読み取れた。教員の評価は生徒にもフィードバックされるため、ある程度生徒にとっても納得のいく評価となっているのではないだろうか。また、生徒と教員の平均値には有意な差は見られなかったが、標準偏差には有意な差が見られた。これは、教員側がＢ評価を標準としてＡやＣの評価を相対的に少なくしていたのに対して、生徒には任意に評価させたことが少なからず影響しているものと思われる。

次に、生徒の母集団を、「自己評価が低い（Ｃ評価が４項目以上）生徒」、「自己評価が高い（Ａ評価が４項目以上）生徒」、「教員評価が低い（Ｃ評価が２項目以上）生徒」、「教員評価が高い（Ａ評価が２項目以上）生徒」の４つのカテゴリーに分類した上で、先ほどと同様の比較を試みた。すると、「自己評価が低い生徒」のカテゴリーでのみ、生徒・教員間での総合評価での相関関係は強まったが、他の３つのカテゴリーではむしろ弱まった。特に、「自己評価が高い生徒」の一部に、教員の評価と著しく数値の異なる生徒が複数名含まれていた（生徒自己評価はオールＡであるのに、教員の評価は複数のＣ評価がある）。あくまでも「自己評価」であるということを勘案しても、評価の妥当性に疑義を抱かせるものである。逆に、「教員評価が高い生徒」の中には、自己評価が比較的低い生徒もかなりの数見られた。これは自己評価であるにも関わらず相対評価を拭いきれない生徒が多いためかもしれない。

続いて、生徒の自己評価項目（５項目）を比較すると、「研究ノート」や「課題発見力」

が比較的高い数値であったのに対して、「論理的思考力」や「見通しを立てる力」の項目の数値は相対的に低い傾向にあった。そして、この「論理的思考力」や「見通しを立てる力」の２項目は、生徒の総合評価との間の相関が高かった。研究ノートの記述や課題の発見までは自力で何とか実行できたが、そこから先伸び悩んだ生徒は自己評価が低くなったと考えられる。逆に、研究の見通しを立てられて、課題研究が考察まで思うように進んだ生徒は自己評価が高くなったものと考える。

一方、教員の評価項目（３項目）を比較すると、「学びに向かう力・人間性」の項目の数値が最も高く、「思考力・判断力・表現力」や「知識及び技能」の２項目の数値は相対的に低くなった。これは、課題研究に取り組む生徒の関心・意欲・態度等に関してはある程度達成していると評価しているが、それ以外の観点に関してはまだ不十分であると感じていることの表れであろう。ただし、今年度に関しては新型コロナウイルス感染症による休校期間の影響もあって、当初の予定通りには課題研究が進められていない。そのため、到達目標である評価規準の調整を途中（２学期）で行ったが、それが上手く機能していなかったのかもしれない。

生徒の自己評価と教員の評価を個別の項目毎に比較してみると、いくつか特徴的なことが見えてくる。例えば、「論理的思考力」や「見通しを立てる力」の自己評価が高い生徒の教員評価は必ずしも高くないという傾向が見られる。一方で、「研究ノート」や「知識習得活用力」の自己評価が高い生徒の教員評価は先の２項目の場合よりも比較的高い傾向にあるということが分かった。研究ノートは定期的に提出させたり、授業中に個別に面談をしたりしながら常に内容の点検を行っていた。また、生徒が集めた新聞記事や先行研究の資料なども、研究ノートと共に教員が確認する機会があった。そういった意味では客観的な評価がしやすい項目であったと言える。しかしながら、「論理的思考力」や「見通しを立てる力」の評価に関しては、生徒の主観と教員の主観との間には若干の開きがあったものと思われる。

他にも、「思考力・判断力・表現力」の教員評価の数値と生徒の総合評価の数値との間で緩やかな相関があるという結果が得られている。教員評価の３項目の中で、「学びに向かう力・人間性」の項目では約３分の１の生徒にＡ評価を付けているが、「知識及び技能」の項目では約２１％、「思考力・判断力・表現力」項目では１８％弱の生徒にしかＡ評価を与えていない。特に、Ａ評価を与えている生徒に限って言えば、生徒の総合評価の数値が高いという傾向が見られた。これに関しては、生徒の自己評価と教員の評価がある程度一致していると言えるだろう。

最後に、今後の課題をいくつか挙げておく。まず、教員評価の妥当性に関してであるが、今回は詳細な分析は行っていないが、個々の教員間で評価のバラツキがあったことは否めない。ルーブリックの評価規準に則って共通の評価を行っているはずであるが、評価基準の受け止め方に個人差が生じてしまったものと思われる。評価に当たっての共通理解が得られるような工夫が今後も必要となってくるであろう。また、今回の評価はあくまでも中間評価であって最終評価ではない。評価の妥当性については少なくとも年間を通して考える必要があり、今回の結果だけをもって語るのは早計である。次年度以降も継続して検討していくことが求められる。次に、生徒の自己評価に関してもいくつか改善の余地があると考える。数は少ないが適当に選んで自己評価シートに記入した者もいたように思う。何故その評価を選んだかの理由も書かせるようにすれば、もう少し慎重に回答するのではないだろうか。また、他人と比べる（相対評価）のではなく、自分自身がどうであったかを客観的に見つめさせ、絶対評価をさせる指導も必要であるように思う。いずれにしても、結果的に生徒の能力向上に資する評価にしていくことが望まれる。