

学生の評価能力に関する考察

白井 靖敏

Discussion about the Assessment Capabilities of Students

Yasutoshi SHIRAI

目的

初等中等教育における「総合的な学習の時間」の本格実施にともない、学習内容や方法、評価など、様々な議論がなされている。その中で学習の評価に関して、興味・感心・態度や技能・表現など、知識・理解とは異なりペーパーテストでは測りにくい観点についての評価手法が検討されている^{1・4)}。成果蓄積型の評価法としてのポートフォリオ評価や、他者による相互評価を取り入れているケースがある^{5・6)}。

そこで、本研究では、評価方法のうち相互評価について、本学の学生を対象に、他者による評価の信頼性や学生の客観的な評価能力について考察する。具体的に、自己評価と他者による評価とのズレを分析検討することによって、それぞれの信頼性を検討し、他者による評価が客観的な指標となり得るのかを探る。また同時に、自己評価と他者による評価との相関関係があるとしたら、数値的な変換処理を施すことによって、たとえば他者による評価だけに頼った場合の補正に役立つとともに、今、大学等で行われようとしている学生による授業評価や教員相互による評価などにも応用できる可能性があるだろうと考える。

方法

本学で筆者が担当している授業、「教育支援技法」(家政学部・教職3年)と「教育の方法と技術」(文学部・児童教育学科・幼児保育専攻4年)において、前者は学生の模擬授業、後者はITを活用した教育方法のうち電子紙芝居の作品と実演についての学習成果の評価を次の方法で行った。

家政学部の教職科目「教育支援技法」における模擬授業では、指導案を含むパワーポイントによる

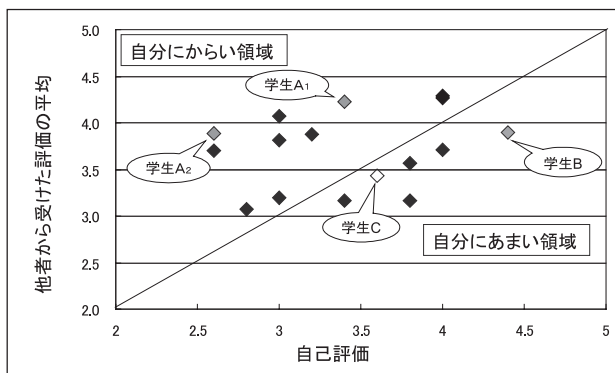


図1 自己評価と他者から受けた評価との関係 (n = 15)
線より上の領域は自分からく(厳しく)、線より下の領域は自分にあまい。

プレゼンテーション教材を作成するとともに、それを使った1人20分程度の模擬授業を、電子教材の分かりやすさ、内容、構成、提示の仕方、話し方の観点別に6段階評価シート(5:すごくよい、4:よい、3:ややよい、2:ややよくない、1:よくない、0:まったくよくない)を配布して学生一人一人が自己評価および他者全員を評価した。

文学部の「教育の方法と技術」では、パワーポイントを使った動きのある電子紙芝居の作成と実演を行い、画面構成、画面の工夫、見やすさ、分かりやすさ、話し方の観点別に6段階評価シート(前述と同様)を配布して学生一人一人が自己評価および他者全員を評価した。

結果

1. 自己評価と他者による評価

図1に「教育支援技法」の学習成果のまとめとして行った模擬授業の評価結果を、縦軸に自己評価、横軸に他者全員から受けた評価を示す。評価点は評価シートに記載されたそれぞれの観点別に6段階評価されたものの平均である。他者から受けた評価より自己評価が低い学生を、自分に厳しい、ここでは「からい」とし、その逆を自分に「あまい」と定義する。図中の直線は自己評価と他者から受けた評価が一致していることを示し、これより下の領域に位置する学生は自分に「あまい」、これより上の領域に位置する学生は自分に「からい」ことをおよそ表すものである。ただし、直線からの偏差がどの程度であれば有意かについては検定していない。

たとえば学生A₁は自己評価が3.4、他者から受けた評価の平均が4.2となっており自分に「からい」と言える。逆に、学生Bは自己評価が4.4、他者から受けた評価の平均が3.9であり自分に「あまい」、学生Cは自己評価が3.6、他者から受けた評価の平均が3.4で、ほぼ一致している。この授業を選択している学生の約7割が自分を「からく」評価していることも分かった。

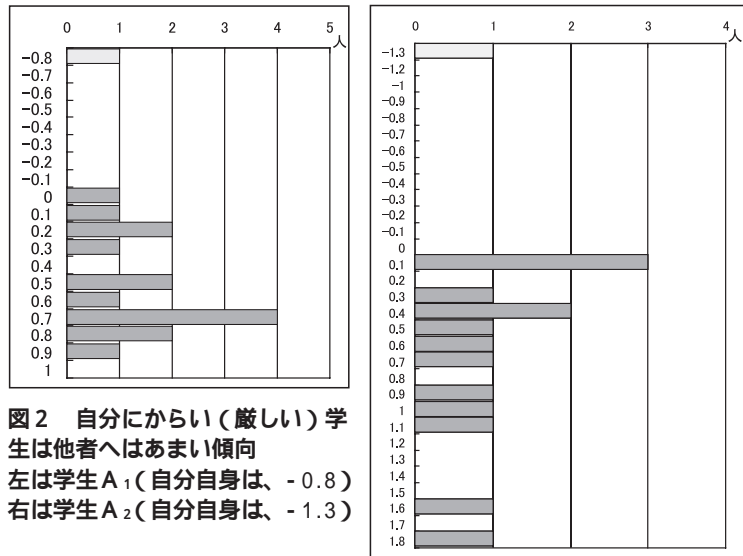


図2 自分にからい(厳しい)学生は他者へはあまい傾向
左は学生A₁(自分自身は、-0.8)
右は学生A₂(自分自身は、-1.3)

2. 自分に「からい」学生の他者への評価傾向

他者から受けた評価の平均が4.2の自分に「からい」学生A₁は、他の学生に対してどのような評価をしているかを図2に示す。図の縦軸は、学生A₁については、他者から受けた評価から自己評価を引いた値を表し、他の学生については、学生A₁が他者を評価した点数からそれぞれの学生が他者全員から受けた評価の平均点を引いた値を表している。結果を見ると、たとえば、自分に「からい」学生A₁は他者全員に対してあまく評価していることが分かる。最大が+0.9、

平均が+0.5であった。また、自分に「からい」学生A₂も最大+1.8、平均+0.7であった。

3. 自分に「あまい」学生の他者への評価傾向

他者から受けた評価の平均が3.9で自己評価が4.4の自分に「あまい」学生Bについて、他の

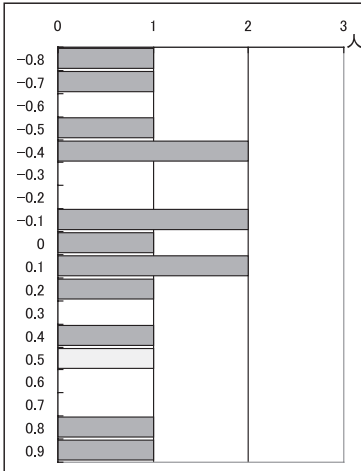


図3 自分にあまい学生は他者へはからい(厳しい)傾向があるが、バラツキが大きい
学生はBは、0.5

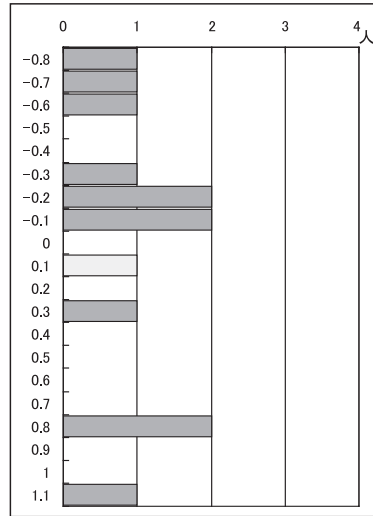


図4 自己と他者から受けた評価とほぼ一致している学生はバラツキが大きい
学生Cは、0.1

学生に対してどのような評価をしているかを図3に示す。学生Bについて、自己評価から、他者から受けた評価点を引いた値が0.5であるとして他者への評価を見ると、ばらつき(標準偏差0.5)があり、最小で-0.8、最大で+1.8となり、平均は0であった。つまり、この学生は他者に対して安定した評価ができていないことを示している。自分に「あまい」他の学生についても同様の傾向がみられた。

4. 自己評価と他者からの評価がほぼ一致している学生の他者への評価傾向

自己評価と他者からの評価がほぼ一致している学生Cについて、他の学生に対してどのような評価をしているかを図4に示す。学生Cは自分では4と評価しているので他者から受けた評価との差は0.1である。他者への評価にはばらつき(標準偏差0.6)があり、最小で-0.8、最大で+1.1となり、平均は0となった。つまり、他者に対する評価が安定していないことがわかる。自己評価と他者からの評価がほぼ一致している学生は少ないが、他の学生についても同様の傾向がみられた。

まとめると、自分に「からい」学生は他者へは「あまい」傾向があり、他者に対する評価基準がある程度しっかりしていると考えられる。逆に、自分に「あまい」学

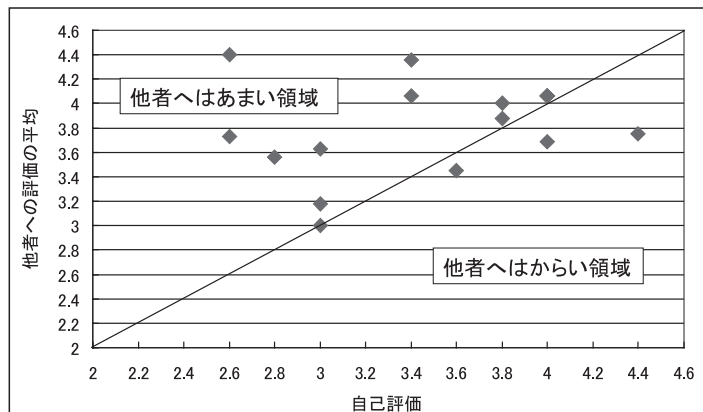


図5 自己評価と他者への評価の平均(家政学部・家政学科 n=14)
一般に他者へは「あまい」評価 $y = 0.1322x + 3.3342$, $R = 0.0367$

生は他者に対する評価にばらつきがあり安定した評価ができていないと予想される。自己評価と他者から受けた評価がほぼ一致している学生は自分に対する評価を客観的に見ることができると考えられるが他者への評価については自分に「あまい」学生と同様、安定した評価ができていないことがうかがわれる。

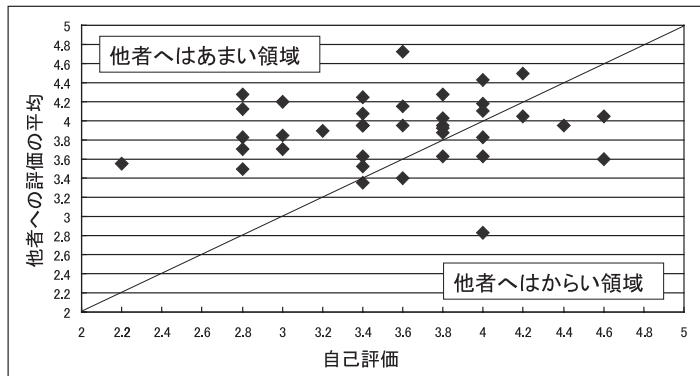


図6 自己評価と他者への評価の平均(文学部児童教育学科幼児保育学専攻 n=37)
一般的に他者へは「あまい」評価 $y = 0.0781x + 3.6268$, $R = 0.0145$

5. クラス内の相対的な評価傾向

1つのクラス内でみたとき、自分に「からく」他者へは「あまい」傾向のある学生は74%、逆に自分に「あまく」他者へ「からい」学生は13%、ほぼ一致している学生、つまり自分を客観視できるだろうと考えられる学生は13%であった(図5)。児童教育学科の電子紙芝居では、自分に「からく」、他者へは「あまい」傾向のある学生は75%、逆に自分に「あまく」他者へ「からい」学生は23%、ほぼ一致している学生、つまり自他ともに安定した評価ができていられる学生は2%であった(図6)。

また、相関はほとんどみられないことから、客観的で正しい評価ができるような評価能力を育成する必要があると考える。

考 察

家政学部の「教育支援技法」と文学部児童教育学科幼児保育専攻の「教育の方法と技術」の2つのクラスだけの分析であることと、サンプル数が少ないので統計的に意味のある分析の信頼性がやや乏しくなるので、一般解を求めることはできないが、対象クラスのおおまかな傾向は見て取れる。自分に「からく」、他者へは「あまい」評価傾向のある学生が2クラスとも7割を越えていることから、対象としたクラスの学生は自分に厳しい傾向があるといえるだろう。この結果は理解度と自信度を調べた下村らの研究にも見られる^{7・8)}。

自分に「からい」学生の他者への評価は、家政学部では平均+0.5、標準偏差0.5、文学部では平均+0.7、標準偏差0.5となった。およそ(自己評価+0.5)±0.5が他者全員から受ける評価の平均評価に近くなると思われる。ただし、偏差が大きいのでひとつの傾向としてのみ意味をもつ。

対象学生の評価能力の点でみると、自分に「からい」傾向のある学生は、他者に「あまい」という、ひとつの評価基準を自分なりに持っていると考えられるので、他者への評価基準をやや「からく」しながら自己評価との調整訓練により比較的短期間で客観的な評価能力が身に付くと考えられる。

自己評価と他者から受けた評価がほぼ一致している学生は約1割と少ないが、客観的な評価がある程度できるものと見てよいだろう。ただし、他者への評価にばらつきがあるため、他者

への評価訓練を積むことによって安定した評価ができていくと考えられる。

しかし、自分に「あまい」学生は、他者への評価のばらつきが大きく、他者への安定した評価ができていないため、自己評価や他者への評価訓練をかなり積まないと、客観的な評価ができる評価能力が身に付かないだろう。

一方、自己評価と他者から受ける評価との間にはほとんど相関がない。自分に「からい」場合と「あまい」場合によって、他者を見る評価基準が異なるので、一般的に見て他者評価の信頼性は乏しい。そのため、他者による評価は統計的に意味をもつ複数者によってなされ、その平均を採用するのが望ましい。あるいは、評価基準にゆらぎの少ない訓練された評価者でなければならないだろう。特に、学生同士による他者評価については変動が大きく、すぐにそのまま一般的な指標とはなりにくいことが分かった。

たとえば、学生による授業評価について、自己の学習についての自己評価と他者、いわゆる授業者への評価との関係で「あまい」のか「からい」のかによって、学生の評価基準が推し測れるとともに、評価能力もある程度推定できると考えられる。より信頼のある評価をおこなうためには、自己評価と他者への評価の差を考慮した補正が必要となるだろう。また、教員相互による授業評価についても同様に、自己評価と他者への評価について分析し、かつ安定した評価基準をもつ訓練された評価者によって評価がなされることが必要ではないだろうか。

謝辞

本研究は、平成16年度本学特別研究助成（個人）によるものである。

参考文献

- 1) 田中忠芳 総合的学習を対象とした大学演習科目における学習グループ人数と自己評価・自由感想文の分析 日本教育工学会論文誌 Vol.27増刊号pp.229-232. (2003)
- 2) 下村勉、白井靖敏、鷲尾敦、他9名「総合的な学習の時間」の成果と課題についての実践的検証 三重大学創造開発研究センター研究報告第12号 pp. 1-10. (2004)
- 3) 下村勉、白井靖敏、鷲尾敦、他19名「総合的な学習の時間」の成果と課題についての検証と分析 三重県高等教育機関連絡協議会研究報告書 (2004)
- 4) 下村勉、白井靖敏、鷲尾敦、他17名「総合的な学習の時間」の実践検証と充実・発展への提案 三重県高等教育機関連絡協議会研究報告書 (2005)
- 5) 天野昌和、下村勉、須曾野仁志 Webベース相互評価システムの開発と効果 日本教育工学会第17回全国大会講演論文集pp.595-596. (2001)
- 6) 吉長裕司、金川明弘、川畑洋昭 打鍵技術の習熟過程における学習者の自己評価と客観評価について 日本教育工学会論文誌 Vol.27-1 pp.71-81. (2003)
- 7) 下村勉、織田守矢 学習者の自己評価と客観評価との統合評価法 電気通信学会論文誌、J63-A, 8, pp.483-490. (1980)
- 8) T.Shimomura, M.Oda, K.Senda Study on Experimental Application of Unified Evaluation Method in Educational Technology. The Transactions of IECE of Japan, Vol.65-E, No.11.(1982)

SUMMARY

By reviewing the difference between the self-appraisal and the evaluation by others in the analysis, the reliability of these types of evaluation was examined. As a result, I considered it possible to convert the evaluations made by other students into an objective index. In this analysis, I found that more than 70% of the students are self-disciplined and lenient to others.

The evaluation standards and assessment capabilities of students could be estimated by viewing the way they evaluate their peers and themselves, that is, whether they tend to be severe or lenient on themselves.

In order to increase the reliability of the evaluation, the students need to be trained to make a good evaluation, and the evaluation they make should be subjected to numerical adjustment.