

【ショートレター】

レポート成果物の学生相互評価を成績に加味する妥当性†

—初年次全学必修科目「教養ワークショップ」における取り組み—

太城 康良*・和田 正法*・野田 明*

三重大学教養教育院*

リーディング・ライティング教育を行う「教養ワークショップ」では、新書に対する書評を成果物としている。本研究の目的は、学生による評価を成績に加味することの妥当性を検証することである。相関解析の結果、教員による評価と学生による評価は、ある程度高い正の相関を示した。学生による評価では、単純な評価票を用い、評価法を指示し、匿名性を持たせることで、教員による評価と同様の傾向を持つようになる。条件を整えることで、学生による評価も成績に加味することの妥当性が示された。

キーワード：ライティング教育，成績評価，学生間評価，ピアレビュー，相関解析

1. 背景・目的

レポートの評価は多くの場合、教員が1名で評価する。マーク式などの客観試験と異なり、論述答案やレポートの採点は内容の多様性やそれに対する主観も排除しにくい。評価の信頼性を高めるためには、複数名で評価を行うことが望ましい。しかし現実的には、レポート課題の内容に精通した教員が複数名で評価にあたることは職務負担面からも困難なことがある。学生自身が評価に参加すれば、複数名での評価が実現し、かつ、評価基準の理解を向上させることで自分が書く文章に対する内省への契機にもなる。

本稿では、三重大学教養教育の「教養ワークショップ」におけるレポート成果物（書評）の評価をモデルに、書評の教員による評価と学生による評価との相関を解析し、学生ピア評価を最終成績に加味することの妥当性の検証を目的とする。

2. 方法

2.1. 対象授業の概要

三重大学教養教育は、自律的・能動的学修力の育成を理念の一つに掲げている。これを実現するためには「読む」「書く」の能力向上が必須と捉え、アクティブ・ラーニング科目の中に、リーディング・ライティング教育を行う「教養ワークショップ」を開講している。教養ワークショップは全学部で1年次必修である。5つの学部（人文・教育・医・工・生物資源）の学生約1,350人が、

教員21人で担当する44クラス（2017年度）に分けられ、各クラスは2学部以上の学生で混成されている（和田・野田 2019）。

授業内容は、スライドと指導手引書によって標準化されている。授業では班毎に論説系の新書を読み、本の要約に自らの論評を加えた書評を作成する。文章の読解、要約、作成に関する解説に加え、5名程度でのグループ演習も行う。学生は毎回メモ（読書シート）を作成しつつ本を読み、本についての意見交換をする。書評の原案を作成した後は、班内外で書評のピアレビューを行い、学生が互いにブラッシュアップして最終稿を完成させる（和田・野田 2019）。

2.2. 成果物（書評）の評価

ピアレビューでは評価票を用いたコメントや採点（ピア評価）を実施する。評価票は大きく書評の構成や字数など7つの単純なチェック項目と10段階の総合評価の二つで構成される（図1）。10段階の総合評価は教員が成績判定に用いるものと同一の尺度で、絶対評価である。全クラス共通で使用する授業スライドでは「ピアレビューの意義」や「普通に真面目な書評の評価は7にする」など、安易に9や10の高評価を出さないよう判断基準の指導をしている。また、ピアレビューの中間段階でもこの評価票を用いてコメントすることを学生に体験させ、互いの書評の改善につながるような観点で評価させる指導もしている。

書評のチェックリスト、兼、評価シート

教員は、査読者のチェック・採点内容にも目を通します。公正・厳格にチェック・採点してください。

受講曜日・時間 いずれかに○印
火5-6 火7-8 火9-10 木1-2 木3-4 木5-6 木7-8

担当教員

被査読者(書評を書いた人)
学籍番号
氏名 班

査読者(書評を読んだ人、このシートの記入者)
学籍番号
氏名 班

評価①(問題がある)、評価②(若干改善の余地がある)、評価③(問題ない)で評価してください。項目10の総合評価は、10段階で評価してください。

1. 指示された構成に従っている。	000
2. 指示された書式に沿っている。	000
3. 指定された文字数で書かれている。	000
4. 新書の内容が分かりやすく紹介されている。	000
5. 新書について十分な評価がなされている。	000
6. 新書の内容と評者の意見の区別が明確である	000
7. 新書を読みたくするような書評である。	000
教員による追加項目	
8	000
9	000
10. 総合評価 (10段階) 評価点⑤以下は評定D相当、⑥はC、 ⑦はB、⑧はA、⑨はAA、#はAA完璧	0000000000

10段階評価

図1 書評の評価票

ピアレビューの最終段階で、学生は自分の班以外の3名以上の書評を評価する。評価者となる学生は被評価者となる学生には匿名である。本研究ではこれを「ピア評価 (班外)」と呼ぶことにする。また、一部のクラスでは、学生は自分の班の自分以外の班員の書評3~5名を、同様に評価した。本研究ではこれを「ピア評価 (班内)」と呼ぶ。さらに、教員が書評について10段階の採点を行ったものを「教員評価」と呼ぶ。

2.3. 解析方法

解析は2017年度の42クラス(履修登録者約1,350人)のうち、解析可能なデータが得られた17人の教員が担当する38クラス(1,237人)を対象とした。出席不足、未提出など、最終成績と書評得点に対応しない不合格者のデータ74件を除外し1,163件を解析した。ピア評価(班内)の解析においては2018年度の1クラスも対象とし、2017年度の1クラスと併せて、教員1名が担当する2クラス計59人の学生データを解析した。

評価票の10段階の総合点について、学生2~4名の評価を平均したものをピア評価の評点とし扱い、教員による評点と対応させ、相関係数の算出と分散図の作成を、統計ソフト「R」(Ihaka・Gentleman 1996)を用いて行った。一般に相関係数Rの解釈は1が「完全な正の相関」、0が「無相関」とされるが、ここでは自然科学的な連続量でなく質問紙調査を用いた離散的な段階数値であるため、「意味のある正の相関 ($R \geq 0.5$)」「ある程度

意味のある相関 ($0.3 \leq R < 0.5$)」の指標で結果を解釈した(水本 2011)。

3. 結果

図2より、教員評価とピア評価(班外)の得点の間には、相関係数0.73の「意味のある正の相関」がみられた(→結果①)。また、教員評価とピア評価(班内)の得点の間には0.25、教員評価とピア評価(班外)の間には0.73の相関係数がみられ、ピア評価(班外)の方が、より高く教員評価と正の相関を示した(→結果②)。

図3より、解析対象を2クラス(図2)から38クラスに拡大しても、教員評価とピア評価(班外)の間には、相関係数0.36の「ある程度意味のある正の相関」がみられた(→結果③)。最終成績と教員評価の得点の間には相関係数0.63の「意味のある正の相関」を示した(→結果④)。最終成績とピア評価(班外)の得点の間には、0.31の「ある程度意味のある正の相関」を示した(→結果⑤)。

4. 考察・結論

結果①について、学生も評価者として教員と同様の基準を有するとみなせる。ピアレビューの最終段階だけでなく中間段階でも評価票を用いたコメントを体験し、10段階の総合評価に至る具体的で単純な7つの項目を学生と教員で共有していること、書評の改善につながるよう安易に高評価を出さないなどの評価の観点を授業で指示していること(和田・野田 2019)が相関を高めるこ

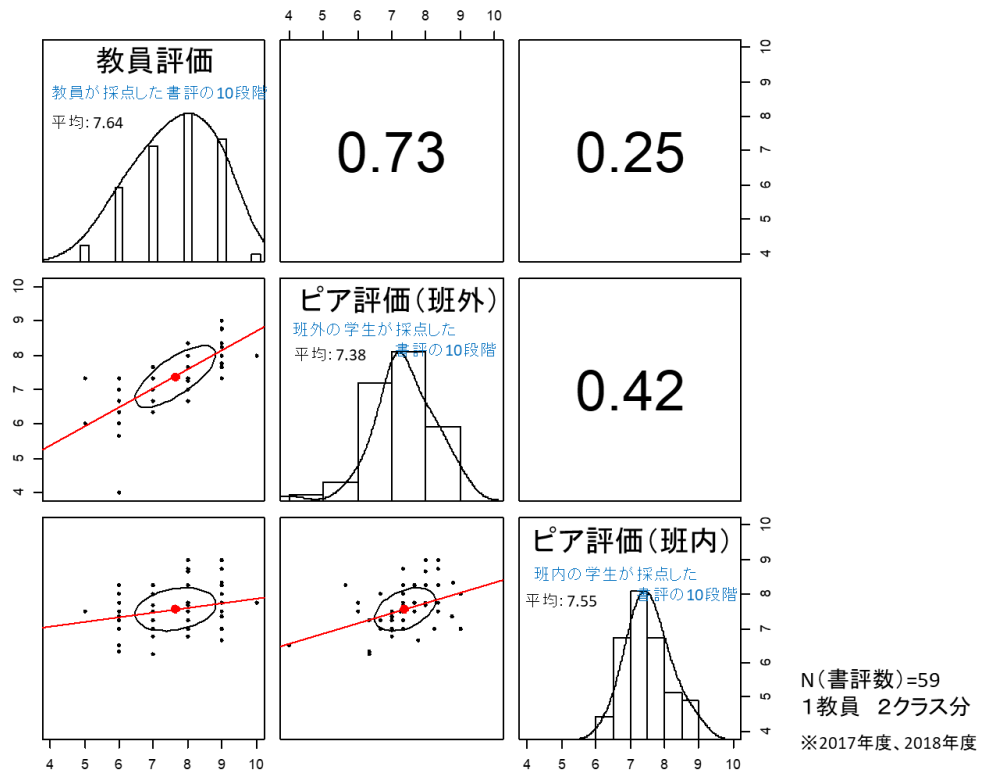


図2 教員評価とピア評価(班外、班内)との相関解析

とに寄与していると考えられる。

評価票の評価項目が多く記述量が多いと評価者は内容を吟味せずに評価を行い、その評価が形骸化することが、グループ活動の評価で指摘されている(太城 2014)。教養ワークショップの評価票は、必ずしも教育評価に精通していない教員や学生にも広く評価内容が理解でき、簡便に精度よく実施できるように工夫したことで、相関が高まったと考えられる。

結果②について、匿名性が高い方が学生は教員と同様に評価する。原因の一つは、班内も匿名ではあるが3~5の少人数であり、班外は自分の班を除くクラス内25名程度の中の誰かという状況で匿名性が上がり低評価も含めより公正に評価しやすくなった可能性が挙げられる。もう一つの原因として、教員と班外の学生は新書の内容に初めて触れる場合が大半で評価者としての観点が近い可能性が挙げられる。評価者として観点の獲得は自らの成果物の内省にも生かされ教育的にも効果が期待できる。

結果③について、ばらつきはあるものの、結果①②の傾向は、解析したすべてのクラスで成立する。解析方法で記載したように、出席不足、未提

出など、最終成績と書評得点に対応しない不合格者74件のデータが除外されている。出席不足、未提出者の書評の完成度は低いと予測されるため、除外された群は、本来すべての分散図の左下方に位置すると推察され、①②の傾向を損なうものではない。

PROGスコアでは、本学の学生のジェネリックスキルに学部ごとの有意差はないことが確認されており(太城 2018)、対象とするクラス・学生の評価能力よりも、指導力など教員に起因する要素がばらつきに影響したと推測される。

結果④、結果⑤について、成果物(書評)の良し悪しが、最終成績にも反映されている。ただし、結果④については総合評価の中には書評の教員評価が含まれており高い相関係数を示したと解釈できる。しかし、成績評価における提出物の内容の割合は最大で40%であることをシラバスに明記しており、その影響は限定的であると考えられる。その状況下でも「ある程度意味のある相関」を示したことは、良い書評を書ける学生は、出席・提出・議論への貢献など、他の項目も良好な評価を受けている可能性が高い。

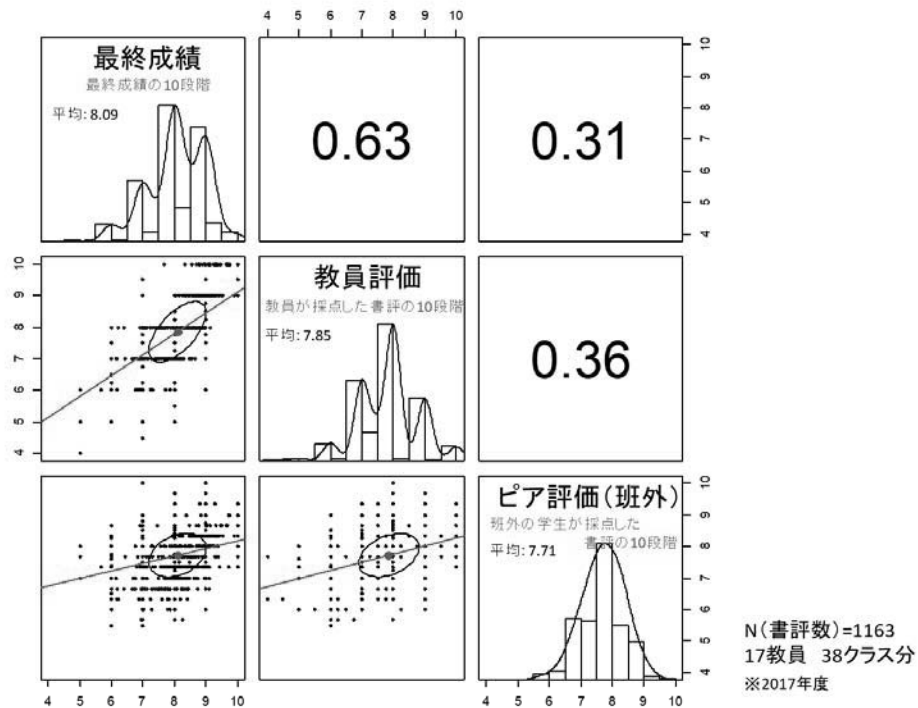


図3 総合評価、教員評価、ピア評価(班外)の相関解析

以上の結果と考察により、単純な評価票で評価の観点を授業で十分指導し、事前に評価票を用いる体験をするなどの条件を整えることで、レポート成果物のピア評価は、成績評価の根拠として十分に妥当であり、教員評価と同様に最終成績に加味することが可能であると結論づけられる。当初、学生による評価は談合や甘い採点により機能しない懸念もあったが、評価法の指示や評価票の工夫により、その可能性は払拭できた。本研究は、複数名の評価によるレポート評価の信頼性の向上、学生による評価で教員のレポート採点を補完、あるいは代替による職務負担の軽減策などの論拠となりうる。

附記

本稿は、太城康良，和田正法，野田明（2019）「レポート成果物の教員評価と学生ピア評価の相関解析—初年次全学必修科目『教養ワークショップ』における取り組み—」第25回大学教育研究フォーラム発表論文集142，で報告した内容にさらなる考察を加えたものである。

参考文献

Ihaka R and R Gentleman. (1996) R: a language for data analysis and graphics. J. Comp.

Graph. Stat. 5:299-314.

水本篤 (2011)「質問紙調査における相関係数の解釈について」外国語教育メディア学会 (LET) 関西支部 メソドロロジー研究部会 2011 年度報告論集, 63-73.

太城康良ほか (2014)「ピア評価の導入による PBL 行動評価の改善の試み」医学教育 45(suppl.) 91.

太城康良ほか (2018)「三重大学アクティブラーニング型授業における成績評価法の検証」第24回大学教育研究フォーラム発表論文集180.

和田正法，野田明 (2019)「書評を通したライティング教育の課題と成果」第25回大学教育研究フォーラム発表論文集54.

† Yasura Tashiro*, Masanori Wada* and Akira Noda*: Validity of student-peer evaluation on report of book reviews: A case study on the first-year mandatory reading-writing course “Liberal Arts Seminar.”

* College of liberal arts and sciences, Mie University 1577 Kurimamachiyacho Tsushi, Mie 514-8507 JAPAN.