

オンライン授業におけるループリックを活用した 相互評価の信頼性の検討

Investigation of Reliability on Peer Assessment with Rubric
in an Online Class of the Graduate School

鈴木 伸子* 石川 奈保子* 向後 千春**
Nobuko SUZUKI* Naoko ISHIKAWA* Chiharu KOGO**

早稲田大学大学院人間科学研究科* 早稲田大学人間科学学術院**
Graduate School of Human Sciences, Waseda University*
Faculty of Human Sciences, Waseda University**

＜あらまし＞本研究では、大学院でのオンライン授業においてループリックを活用した相互評価の信頼性を検討した。相互評価における信頼性の確保に重要とされている一貫性について、評価者3人の採点がどれほど一致しているかという点から検証を行った。その結果、ループリックの各観点、および評点の合計点ランクにおいて高い一致率がみられた。

＜キーワード＞ 相互評価 ループリック オンライン授業 eラーニング ピア・アセスメント

1. はじめに

近年、大学等の授業において相互評価が盛んになっている。従来の教員評価に代わり相互評価が行われる背景には、大規模公開オンライン講座 (Massive Open Online Courses, 以下 MOOC) 等の普及がある。相互評価は MOOC の成績評価でも使われている仕組みであり、教員側の採点に伴う負荷を軽減することができる。このような効率面に加え、教員評価だけの場合に比べ、フィードバックが多様になる、学習者の自律性や内省が促進されるなど学習効果面でのメリットもあげられる (植野ほか 2008)。

高等教育における相互評価に関する実践や研究は、近年多くの事例が報告されている (桑原ほか 2014, 金子・登り口 2008 など)。今後、高等

教育の実践に相互評価を導入していくには評価の信頼性・妥当性の検討も必要である。そこで本研究では受講者同士の相互評価を取り入れた大学院でのオンライン授業を対象として、妥当性の前提となる相互評価の信頼性を検討した。

藤原ほか (2007) は、相互評価における信頼性の確保には一貫性、評価への専念、公平性が重要であると述べている。本研究では一貫性の点から信頼性の検証を行った。

2. 方法

本調査は、X 大学大学院にて毎年開講している専門科目で行った。授業は大学が提供する LMS (Learning Management System) 上で配信される e ラーニングで行われた。本年度 2016 年の履修者 27 名のうち 20 名が全期間を受講し、2 回

表 1 レポート評価のループリック

観点	評価				
	5	4	3	2	1
1 自分の領域への応用	「課題図書の考え方で面白いと思ったこと」「それをどう自分の領域に応用するか」「どちらもわかりやすく書かれています」	「課題図書の考え方で面白いと思ったこと」「それをどう自分の領域に応用するか」「どちらもわかりやすく書かれています」	「課題図書の考え方で面白いと思ったこと」「それをどう自分の領域に応用するか」「どちらも書かれています」	「課題図書の考え方で面白いと思ったこと」「それをどう自分の領域に応用するか」「どちらも書かれています」	「課題図書の考え方で面白いと思ったことが書かれていません。あるいは2つ以上書かれています」
2 実現可能性、効果、障害	「自分の領域へ応用することの実現可能性、実現した場合の効果、実現に向けての障害の3つが、根拠もふくめてわかりやすく書かれています」	「自分の領域へ応用することの実現可能性、実現した場合の効果、実現に対する障害の3つがすべて書かれています」	「自分の領域へ応用することの実現可能性、実現した場合の効果、実現に対する障害のうち、1つあるいは2つしか書かれていません」	「自分の領域へ応用することの実現可能性、実現した場合の効果、実現に対する障害が3つとも書かれていません」	「自分の領域へ応用することの実現可能性、実現した場合の効果、実現に対する障害が3つとも書かれていません」
3 内容への興味	「このレポートは大変興味深く、他人に読むことを勧めたいと思った」	「このレポートは興味深いと思った」	「このレポートは興味深いと思った」	「このレポートにはあまり興味が持てなかった」	「このレポートには興味が持てなかった」
4 レポートの読みやすさ	「誤字脱字が多く、かつ段落分けが適切で、全体がスムーズに読める」	「誤字脱字がなく、かつ段落分けが適切である」	「誤字脱字が多少あるが全体の段落分けが適切である」	「誤字脱字はないが、段落分けが適切ではない箇所が一部ある」	「誤字脱字が多くあり、かつ段落分けが適切ではなく、全体がスムーズに読めない」
5 レポート全体の出来	「とても優れている」	「優れている」	「合格水準だと思う」	「不十分な所がある」	「不十分な所が多い」

のレポート作成とそれに対する相互評価を行つた。評価は1つのレポートについて3人の受講者が評価者となり採点を行つた。評価者は教員がランダムに割りあて、受講者には自身が評価するレポートの作成者および自身のレポートを評価した評価者が特定できないように配慮した。

前年の授業では5点満点の絶対評価を行い、その結果から、学生の相互評価に一定の信頼性があることが示された(鈴木・向後2016)。本年度の授業ではループリックを採用し、より詳細な観点からの評価を行つた。レポート評価に使用したループリックは5つの観点からなり、観点ごとに5段階で評価した(表1)。各観点の評価を評点とし、評点の合計点(評点合計)についてA~Eの5段階でランク付けをおこなつた(表2)。ランクは22点以上25点以下をA、18点以上22点未満をB、14点以上18点未満をC、10点以上14点未満をD、5点以上10点未満をEとした。

3. 結果

一貫性は、評価者3人の評価がどれだけ一致しているかという一致度から検証した。一致度は1つのレポートに対して「3人の評価が一致」をS3、「2人の評価が一致」をS2、「3人とも評価が異なる」をS0、として判定した(表3)。判定結果から、2人以上の評価(S3+S2)が一致した割合を一致率とした。一致率は、ループリックの各評点および評点合計のランクの2つで判定した。その結果、観点1~5の各評点の一致率はいずれも70%を超えており、どの観点にも目立った偏りは見られなかつた(表4)。また評点合計のランクでは一致率が87.5%となつた(表5)。

表2 学生aに対する相互評価の例

評価者	観点1	観点2	観点3	観点4	観点5	評点合計	合計点ランク
評価者1	4	4	3	4	4	19	B
評価者2	4	2	4	5	3	18	B
評価者3	5	5	2	4	4	20	B

表3 相互評価の一致度判定の例(課題1)

被評価者	評点			ランク			一致度
	評価者1	評価者2	評価者3	評価者1	評価者2	評価者3	
a	19	18	20	B	B	B	S3
b	18	15	17	B	C	C	S2
c	13	14	13	D	C	D	S2
d	18	20	23	B	B	A	S2
e	8	19	13	E	B	D	S0
f	20	18	19	B	B	B	S3
g	11	11	9	D	D	E	S2

表4 ループリック観点別一致度(全課題)

一致度	観点1	観点2	観点3	観点4	観点5	合計	%
S3	5	8	3	2	8	26	13.0%
S2	23	22	26	28	22	121	60.5%
S0	12	10	11	10	10	53	26.5%
合計	40	40	40	40	40	200	100.0%
一致率 (S3+S2)	70.0%	75.0%	72.5%	75.0%	75.0%	73.5%	

表5 合計点ランク一致度(全課題)

一致度	課題1	課題2	合計	%
S3	4	2	6	15.0%
S2	13	16	29	72.5%
S0	3	2	5	12.5%
合計	20	20	40	100.0%
一致率 (S3+S2)	85.0%	90.0%	87.5%	

4. 考察

これらから相互評価の信頼性について考察する。前年度の授業における同様の調査では一致率が70%以上を示しており信頼性は高いと判断した(鈴木・向後2016)。本年度は、より詳細な観点からの検証を目的としてループリックを採用した。その結果、一致率はループリックの全観点で70%を超え、評点合計のランクでも90%近い一致率であった。このことから相互評価の一貫性は高く、信頼性が確保されたと判断できる。

ループリックを相互評価に取り入れることは、評価の視点が明確になり信頼性が向上するだけでなく、受講者が評価基準を意識することで授業における成果物の品質向上も期待できる。今後、大規模オンライン講座や、高等教育機関の初年次教育の事例とも比較し、相互評価におけるループリック採用の効果なども検証したい。

参考文献

- 藤原康宏、大西仁、加藤浩(2007) 学習者間の相互評価に関する研究の動向と課題(展望論文). メディア教育研究, 4(1): 77-85
 金子大輔、登里口泰久(2008) 相互評価やグループ学習を支援するシステムの開発と基礎的情報教育での利用. 日本教育工学会論文誌, 31(Suppl.): 33-36
 桑原千幸、喜多敏博、合田美子、根本淳子、鈴木克明(2014) 初年次キャリア教育科目における相互評価学習の実践と進路選択自己効力の向上. 日本教育工学会論文誌, 38(2): 79-89
 鈴木伸子、向後千春(2016) 大学院でのオンライン授業における相互評価の信頼性と妥当性の検討. 日本教育工学会研究報告集, JSET16-2, pp.5-12
 植野真臣、ソムマーン ポクポン、岡本敏雄、永岡慶三(2008) ピアアセスメントにおける評価者特性を考慮した項目反応理論. 電子情報通信学会論文誌, D, 情報・システム J91-D(2): 377-388