

演習における相互評価による学習効果と問題点

坪倉篤志^{*1}

^{*1} 日本文理大学・九州工業大学

Learning effect and problem about Peer review during exercise

Nippon Bunri Univ. and Kyushu Institute of Technology ^{*1}

Until now, we have been working on Peer review using various environments. In both cases, a dedicated system has been constructed and implemented, and devices have been introduced that do not cause problems that are predicted by using general systems. However, we have not confirmed problems with general systems. Therefore, this time, we tried to evaluate mutual using moodle's forum and investigated the problem by using general system. This time we will report on the results.

キーワード: 相互評価, コントロールレス, 問題点

1. はじめに

これまで我々は大学教育において、マルチメディアコンテンツ制作に向けた教育として、演習を中心に取り組んできた。授業で取り扱ってきた分野としては、CG(DTP, Illustration, Photo レタッチ)や3D-CGにおけるModeling, Web ページ(HTML, CSS, JavaScript), アニメーション(Flash)などである。授業では、資料配布や課題管理としてLMS/CMSであるmoodleを用い、授業運営してきた。課題は、毎授業で複数の課題を課している。提出期限は、基本的に翌授業日の前日とし、授業日の朝に教授者が、提出物を評価している。課題提出者は自身の評価結果を閲覧できる。授業では全体向けにレビューと評点が下がっている場合は、基本的には再提出を全体に向けて促していた。しかし再取り組みと再提出へ取り組む学習者は限られていた。

相互評価による効果は様々な先行研究により報告されている¹⁻³⁾。そこで、学習者の課題への取り組み意欲の向上と、自身の提出物が持つ要改善事項に対するアウェアネス向上による、再度課題への取り組みや、再提出の改善に向けた取り組みとして、相互評価や振り返り学習に取り組んできた。

2. 相互評価への取り組み

2.1 概要

相互評価時の学習者が行う活動は、1,他学習者提出物の閲覧, 2,他学習者提出物の評価, 3,他学習者による自己提出物に対する評価結果の閲覧, の3種類である。これら活動から、学習者は、同課題へ他学習者が取り組んだ成果物の閲覧や、評価への取り組みからの学習、自己提出物に対する第三者による評価結果に触れ、様々な学習をする。これらから、学習者は自身の課題への取り組みを振り返り、自身の取り組みの改善や提出物の改善に取り組む(図1)。

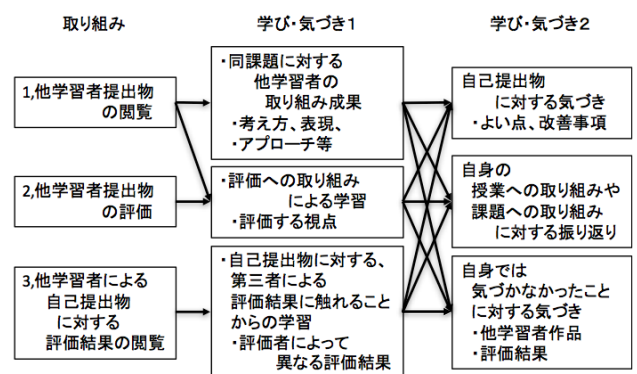


図1, 相互評価での取り組みと学びや気づき

これまで我々は学習者の課題への取り組み意欲向上と、自身の提出物が持つ要改善事項に対するアウェアネス向上に向け、3種類の相互評価へ取り組んできた。

2.2 イベント型⁴⁾

簡易な相互評価でありイベント性として、Web 展覧会を導入した。取り組む課題は、テーマを設定し、企画立案から制作まで各学習者が独自に取り組む演習として実施した。これは、通常の課題提出は、教員のみが閲覧し、評価を行う。Web 展覧会は、最終的には全学習者に公開される。そのため、通常の課題程支出と比較すると、学習者の意欲が高まり、課題への取り組みが改善することを期待して実施した。実名は非公開とし各学習者が独自に設定したペンネームで公開。評価は投票形式で、評価者が、よいと判断した作品に対し投票件数には上限を設けず投票させた。投票結果は全体に公開した。実施結果より、一定の効果が確認できた。しかし、日常的な課題評点の低い学習者は、展覧会向けの作品制作が間に合わない場合もあり、効果が見られない場合もあった。



図2，イベント型相互評価：Web 展覧会

また、学習者によって、展覧会における相互評価への取り組み行動には、偏りが見られ、期待していた学習効果にばらつきがみられた。また Web 展覧会では、一過性のイベントとなってしまう、要改善事項に気づいたとしても、これを生かす機会が、次の授業期になってしまう。

2.3 繰り返し型^{5~6)}

そのため、日常的な課題を対象に、繰り返し相互評価に取り組み、気づきからの改善に向けた取り組みを促進できるよう、継続・繰り返し型を試行した。評価項目はルーブリックで定義し、公開と解説。課題はコンテンツが中心となるため、評価項目に対し評価者依存特性を設定し開示した。相互評価の問題として、評価者の評価結果における妥当性も問題視されていた。そこで、評価に取り組んだ結果は、教授者の評価結果とも比較できるようにした。これにより評価者は、自身が取り組んだ評価への取り組みが、評価者依存特性が「ない」評価項目は、教員と評価者の評価結果の比

較から評価の妥当性を学べるようにした。また評価者依存特性の「ある」評価項目は、評価者によって、評価が異なることを学べるようにした。これにより、評価者も、被評価者も評価結果に対しても、評価結果の閲覧から学べるようにした。実名は非公開とした。こちらも一定の効果が確認でき、特に上位層の学習者の改善が高い傾向があった。しかし、下位層への効果や取り組みは限定的であった。

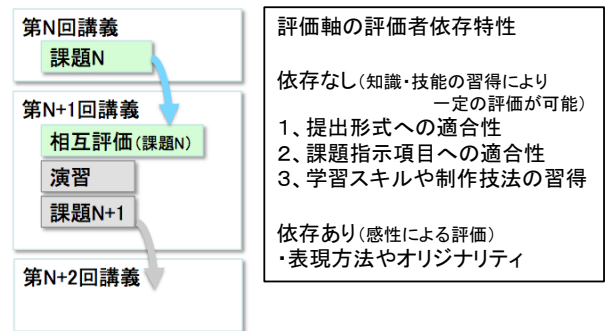


図3，継続・繰り返し型相互評価と評価者依存特性

2.4 組み込み型⁷⁾

相互評価では、様々な活動から、様々な学習に取り組む。相互評価における特定の段階の取り組みと、課題への取り組みとで、関連する活動を連動させると、相互評価や課題への取り組みが、より効果的になると考えた。そこで組み込み型として、課題への取り組みに相互評価を組み込んだ。

情報伝達に向けた企画立案時(図4)、企画立案者は、伝達情報の分析や構成について考える。このプロセスにて企画立案者は、視聴者像を想定しながら、伝達情報の構成や内容調整を行う。この伝達情報の構成は、様々な情報の調査・分析を行いながら、アイデアを広げる。おおよその方向性が決まってきたら、アイデアの絞り込みを行い、企画を構成し、ブラッシュアップに取り組む。

個人が持てる視点や情報収集能力には限りがある。これらプロセスにおいて、第三者による評価結果や第三者の成果を閲覧することは、学習者にとって非常に参考になると考えた。そこで、企画立案に取り組む課題の中で、各個人が一旦企画立案した後、企画のブラッシュアッププロセスに相互評価を組み込んだ。相互評価への取り組みは様々な活動(図1)がある。そのため、学習者にセルフコントロールさせると、活動に偏りが出る傾向が高くなってしまふ。そのため、ブラッシュ

アッププロセスと相互評価を段階的に取り組める段階的相互評価(図5)で実施できるようにした。

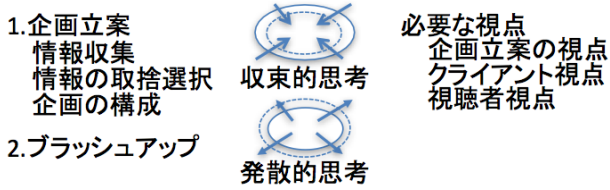


図4, 企画立案とブラッシュアップ, 必要な視点

Step	相互評価	作業
0		作品の提出: 提出物
1	1段階目	他学習者提出物の閲覧と評価
2		提出物の再検討と見直し①
3	2段階目	被評価結果の閲覧
4		提出物の再検討と見直し②

図5, 組み込み型相互評価: 段階的相互評価

以上のように, これまで3種類の取り組みを試行してきた。いずれの取り組みも一定の効果が見られ, 特に成績上位層に与える影響が高かった。しかし, 成績下位層に対する効果が限定的であった。そのため, これらに対する対策が課題となっている。

3. 調査

3.1 調査目的

以上のように, 様々な相互評価に取り組んできた。これらは, 各種相互評価への取り組みから, 発生する問題点などを考慮した取り組みとなっていた。特に, コントロールレスな状況下での学習者行動の偏りについては, データで確認ができてない。そこで, 本研究では, 一度基本に立ち戻り, 一般的な相互評価環境下で, 学習者行動の偏りについて調査を行う。調査結果から閲覧数, 相互評価数, 閲覧時間などの相関関係から, 学習者行動の偏りの分析を行う。

3.2 調査方法

今回, 調査で用いる演習は, 「コンテンツを用いた情報伝達の上で, 伝達情報と伝達対象に対する分析結果を用いた相互評価」にて調査をおこなった。

調査環境は, moodle(1.9.19+)のフォーラムモジュールを用いる(図6)。相互評価で用いる課題は, テーマに沿った, テーマに対する「もの・こと」分析結果と, 「ターゲット層」分析結果である。テーマは事前に用意していた5種類の中から, 事前調査を行い, 受講生の希望が均等分布するテーマにて実施した(事前に用

意したテーマ: 犬派・猫派, ご飯派・パン派, iOS派・Android派, Mail派・Line派, 一人暮らし派・実家派)。今回は「実家派・一人暮らし派」とし, 「実家派」は, 実家派である考えを「一人暮らし派」へ, 「一人暮らし派」は, 一人暮らし派である考えを「実家派」へ伝えるコンテンツ制作を目的に, 分析に取り組みさせた。相互評価では, まず「もの・こと」分析結果を, 相互評価に取り組みさせた。次に時間的に余裕のある学習者に対しては, 「ターゲット層」分析結果を, 相互評価に取り組みさせた。分析では, コンセプト分析用紙(図6)を紙で用意し, 手書きで取り組みさせた。また相互評価には, 各学習者が持つ携帯電話のカメラ機能で撮影させた画像を用いた。この画像ファイルを, 各学習者が作成したディスカッショントピックスに掲載させた。タイトルは, 自分が「実家派」か「一人暮らし派」を明記させ, 学籍番号を付与させた。相互評価では, 各学習者が閲覧した提出物に, コメントを記入させた。

名前や学籍番号は, オープンとした。学習者に課した取り組みは, お互いの提出物を閲覧。評価はフォーラムのコメント機能にて, コメントバック。コメント内容は, 自分と同じスタンスの場合, 参考になったこと, 指摘できること。自分と異なるスタンスの場合, ターゲット層の視点で, 参考になったこと, 指摘できることの視点で, 建設的な指摘や, 感想に限定するよう説明。内容によっては教育的指導を行うと事前に説明してある。

閲覧数やコメント数は, 目安として全提出の半分程度は閲覧すること。時間的に余裕があれば, さらに多くの提出物の閲覧に取り組むこととした。評価コメントは5件程度を目安とした。取り組み手順や時間は, 特に制約を与えず, セルフコントロールとした。

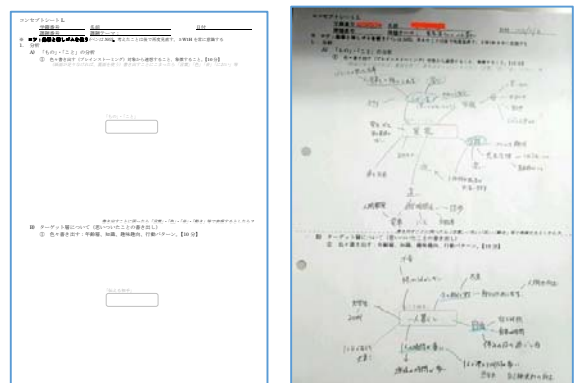
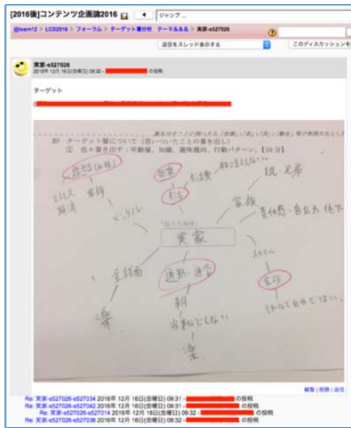


図6, コンセプト分析用紙

(左: フォーム, 右: 分析結果例)



Re: 実家-4527026-4527034 2016年 12月 16日(金曜日) 09:31 - [redacted] の投稿
 Re: 実家-4527026-4527042 2016年 12月 16日(金曜日) 09:31 - [redacted] の投稿
 Re: 実家-4527026-4527014 2016年 12月 16日(金曜日) 09:32 - [redacted] の投稿
 Re: 実家-4527026-4527036 2016年 12月 16日(金曜日) 09:32 - [redacted] の投稿

図7, フォーラムモジュールを用いた相互評価
 (上:一例 下:コメントバック)

3.3 実施結果

2016年11月2日に実施した, 大学2年生向けの授業, コンテンツ企画論(90分授業)の第5回授業にて, 事前調査から分析, 相互評価に取り組んだ. 分析手法は, 第4回授業に解説・演習済みである. 出席者数は21名であった.(図8)



図8, 授業実施教室の様子

3.4 分析結果

分析では moodle のログデータを用いた. moodle のログデータは, フォーラムの閲覧(forum view form: 記事一覧の閲覧), フォーラム内ディスカッションの閲覧(forum view discussion)・ディスカッションの作成(forum add discussion)・修正(forum update), ディスカッションに対するコメント送信(forum add post)が記録される. 今回の分析では, これらを用いた.

まずは, 取り組み時間についてである. 表1に各相互評価回と全体の平均値, 最大値, 最小値の計算結果を示した. さらに, 最大値・最小値間の倍率を表2に示した. これらより, 全体では, 最小取り組み時間者から最大取り組み時間者の間に約3倍~約17倍の差

があることがわかる. 個別の相互評価内では, 約2.6倍から約251倍である.

表1, 取り組み時間(値分布)

		総取り組み時間	フォーラム閲覧(記事一覧)	ディスカッション		
				閲覧	追加	コメント送信
全体	average	1:16:17	0:28:13	0:47:43	0:00:04	0:00:15
	max	1:43:08	0:56:40	1:17:38	0:00:07	0:00:33
	min	0:33:16	0:03:13	0:20:18	0:00:02	0:00:02
相互評価1	average	0:37:08	0:12:23	0:24:36	0:00:02	0:00:07
	max	0:53:53	0:35:52	0:38:52	0:00:04	0:00:15
	min	0:20:35	0:00:30	0:08:21	0:00:01	0:00:00
相互評価2	average	0:39:08	0:15:51	0:23:08	0:00:02	0:00:08
	max	0:54:45	0:33:28	0:47:16	0:00:04	0:00:26
	min	0:10:08	0:00:08	0:09:43	0:00:00	0:00:00

表2, 取り組み時間(Max/Min)

max/min (time)	総取り組み時間	フォーラム閲覧(記事一覧)	ディスカッション		
			閲覧	追加	コメント送信
全体	3.1	17.6	3.8	3.5	16.5
相互評価1	2.6	71.7	4.7	4.0	-
相互評価2	5.4	251.0	4.9	-	-

図9に全学習者の取り組み時間の分布を示した.

これを, 各学習者別に各取り組みに対する時間割合を計算した結果が図10である. いずれの結果からも, 学習者間で取り組み分布にバラツキを確認でき, 学習者によって, 記事一覧の閲覧が長い場合と, ディスカッションの閲覧が長い場合に, 大きな差があることがわかる.

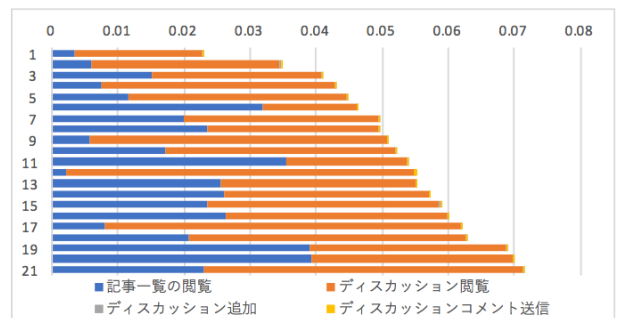


図9 各取り組み時間
 (総取り組み時間でソート)

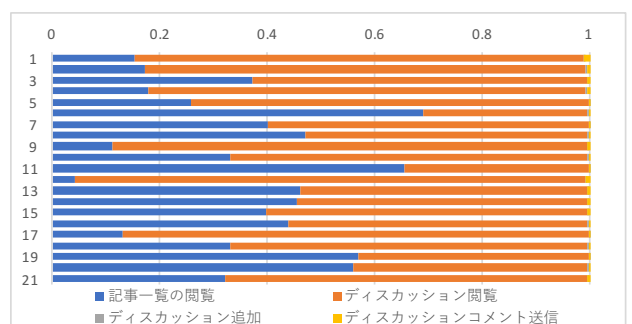


図10 各取り組み時間割合
 (総取り組み時間でソート)

次に、行動数である。ログから各行動をピックアップし回数を分析した。分析結果を表 3、4 に示した。行動数においても、全体では約 3 倍～約 10 倍の差が見られ、相互評価別になると、最大約 26 倍の差が見られる。

閲覧作品数は、全体の平均値が約 25.67 と目標値を上回るものの、最低値が 10 と目安の半分程度であった。対して最大値は 38 と、目安の約 1.8 倍であった。コメント数は、全体の平均値が約 8.52 と目安を下回り、最低値は 2 と目安の 2 割と大きく下回った。対して最大値は 14 と目安の約 1.4 倍であった。

表 3 行動数 (値分布)

	アクション数	フォーラム閲覧 (記事一覧)	ディスカッション			閲覧記事 (Uniq)	
			閲覧	追加	コメント送信		
全体	average	82.05	23.95	47.19	2.10	8.52	25.67
	max	147.00	52.00	110.00	3.00	14.00	38.00
	min	33.00	8.00	11.00	1.00	2.00	10.00
相互評価 1	average	43.00	13.05	24.48	1.00	4.29	11.05
	max	90.00	41.00	43.00	1.00	7.00	18.00
	min	18.00	3.00	8.00	1.00	0.00	6.00
相互評価 2	average	39.05	10.90	22.71	1.10	4.24	14.62
	max	97.00	20.00	79.00	2.00	7.00	24.00
	min	11.00	1.00	3.00	0.00	0.00	4.00

表 4 行動数 (Max/Min)

max/min (count)	アクション数	フォーラム閲覧 (記事一覧)	ディスカッション			閲覧記事 (Uniq)
			閲覧	追加	コメント送信	
全体	4.5	6.5	10.0	3.0	7.0	3.8
相互評価 1	5.0	13.7	5.4	1.0	—	3.0
相互評価 2	8.8	20.0	26.3	—	—	6.0

閲覧作品割合 (ユニーク閲覧数/出席者数) の計算結果を図 11 である。1 番目の学習者は、コメント目安値に近い閲覧数であった。目安の半分以上を閲覧した学習者は、9 番目以降である。以上より出席者の約 57% は目標を達成していることがわかる。

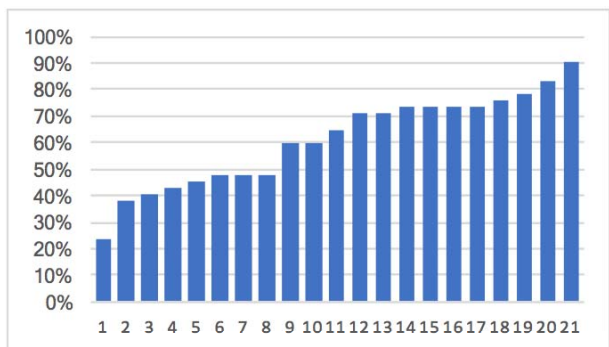


図 11 閲覧作品割合

次に閲覧時間や閲覧回数とコメント数との関係性について相関で分析した。総閲覧時間と閲覧数 (ユニーク) の計算結果を図 12 に示した。相関係数は、-0.15 と無相関であった。

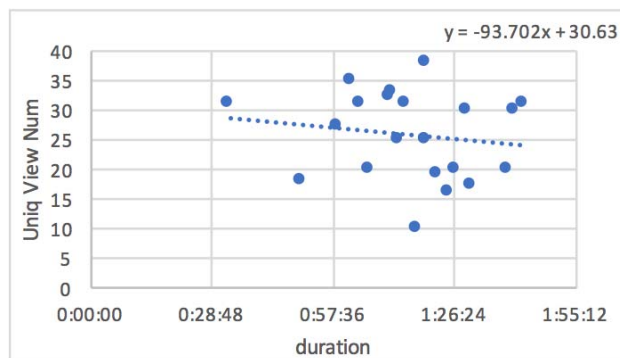


図 12 総閲覧時間と閲覧数(ユニーク)
(相関係数 -0.15 : 無相関)

次に総閲覧時間とコメント数の計算結果を図 13 に示した。こちらの相関係数は 0.17 と、無相関であった。

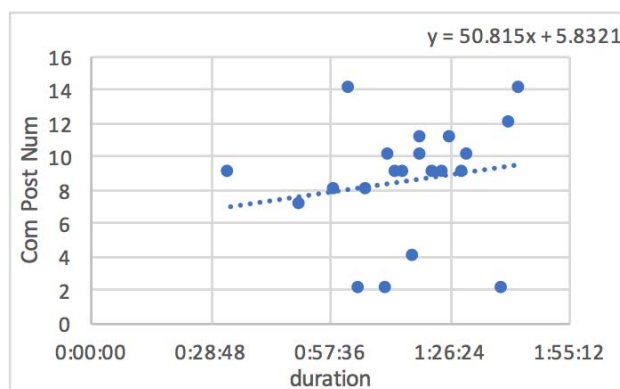


図 13 総閲覧時間とコメント数
(相関係数 0.17 : 無相関)

さらに、コメント数と閲覧数 (ユニーク) の計算結果を図 14 に示した。こちらの相関係数は 0.30 と弱い相関があることがわかった。

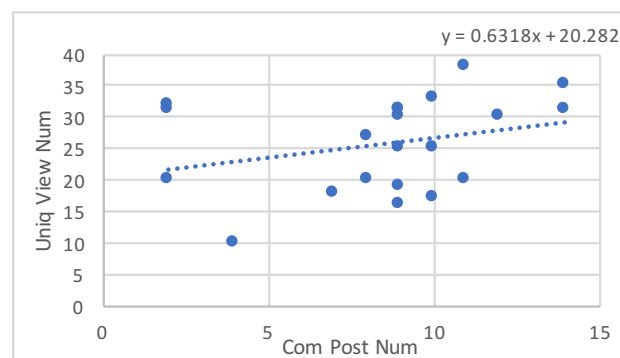


図 14 コメント数と閲覧数 (ユニーク)
(相関係数 0.30 : 弱い相関)

いずれの相関も、無相関か弱い相関であり、閲覧数 (ユニーク)、コメント数、閲覧時間には、あまり関係性が見られないことがわかった。

これらより、閲覧数の多さや、閲覧時間の長いことが、評価数の増につながる訳ではない。閲覧数の多さ

参 考 文 献

と閲覧時間の長さについても、関係性は無いことがわかった。以上より、学習者別の行動量には、バラツキを確認した。また、他学習者提出物の閲覧や、評価行動は、学習者によるセルフマネジメントでは、偏りが発生することが明らかになった。

4. まとめ・展望

今回、一般的な相互評価環境として、簡易な相互評価が手軽に可能な moodle のフォーラムモジュールを用い、相互評価に取り組んだ。今回取り組んだ相互評価は、特別な支援システムや方法を実施せず、取り組み内容と、取り組み目標を目安として示し、学習者の自主性に任せた。この結果、学習者間の行動量や行動時間、取り組んだ行動種に、学習者間のバラツキが大きく見られた。また閲覧数や閲覧時間と評価への取り組みに相関は見られなかった。これらより、閲覧や閲覧時間、コメントに、学習者によって偏りのあることを示している。

相互評価への取り組みから期待される学習効果は、相互評価への取り組みである、1,他学習者提出物の閲覧, 2,他学習者提出物の評価, 3,他学習者による自己提出物に対する評価結果の閲覧の3種類への取り組みが重要である。そのため、相互評価を用いた学習から、期待される学習効果を得るためには、学習者行動をある程度のコントロールし、取り組みが必要な活動を取り組ませる必要がある。

しかし、学習者能力には、差異がある。相互評価では、多くの他学習者作品の閲覧や深い解釈、そして多くの学習者からの評価結果（被評価結果）の閲覧や深い解釈から、学習につなげることが、高い学習効果につながると考える。

以上より、相互評価への取り組みにおいて、学習者行動を、ある程度のコントロールしながら、能力の高い学習者に対しては、ある程度の自由度を許すことにより、より高い学習効果を望める相互評価への取り組みが可能になると考えている。

なお、今回は、moodle のフォーラムモジュールを用いたため、コメントに対する閲覧記録がログに残らず、分析できなかった。今後は、専用のシステムを用いた詳細な分析が必要だと考えている。

- (1) 藤原康宏,大西仁,加藤浩, 学習者間の相互評価に関する研究の動向と課題, メディア教育研究 4(1),77-85,2007
- (2) 布施泉,岡部成玄,多段階相互評価法による学習の実践と効果 日本教育工学会論文誌 33(3), 287-298,(2010)
- (3) 中山晃,高木正則,勅使河原可海,全員参加型の協調学習のための性格を考慮したグループ編成方法の研究,情報処理学会研究報告. CSEC, 2011-CSEC-52(33),1-6,(2011)
- (4) 坪倉,松原, マルチメディアに対応した発表会・展覧会を行う Web システムの構築: システム構築とアンケート評価,電子情報通信学会技術研究報告, 教育工学 107(391), 31-36,(2007)
- (5) 坪倉篤志,松原伸人,足立元,西野和典, マルチメディア課題の評価における評価者依存の少ない評価手法の構築の試み~課題要件に対応した評価基準の明確化と試行結果について~,教育システム情報学会 第35回 全国大会 論文集, pp249-250, ISSN 1342-9477, (2010)
- (6) 坪倉篤志, 松原伸人, 林敏浩, 足立元, 西野和典, 制作課題における評価者視点の学習のための相互評価システムの研究, JSiSE research report 29(7), 21-28, 2015-03
- (7) 坪倉,竹内, 相互評価における課題提出者と評価・閲覧者の組み合わせが与える影響,教育システム情報学会研究報告 30(1),57-64,2015-05