

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年6月11日現在

機関番号：21201

研究種目：若手研究（B）

研究期間：2008～2011

課題番号：20700642

研究課題名（和文）ピア・アセスメントにおける他者評価能力の可視化ツールの開発

研究課題名（英文）Development of an evaluation ability visualization system for peer assessments

研究代表者

藤原 康宏（FUJIHARA YASUHIRO）

岩手県立大学・ソフトウェア情報学部・准教授

研究者番号：30305338

研究成果の概要（和文）：実践的な課題やグループワークの場面で、学習者間の相互評価を導入した実践がよく行われている。学習者の他者評価能力を向上させるために、学習者の他者評価能力を可視化するツールを開発した。また、相互評価を導入した実践を行った結果、多くの学習者が相互評価の効果を認識していることが分かった。他者に行った評価が受け入れられているかどうかを主観的にも納得させるためには、トラストの概念が重要であることが明らかになった。

研究成果の概要（英文）：Peer assessment is well known as the effective evaluation way for project based learning, group work. We have implemented an evaluation ability visualization system for peer assessments. We practiced peer assessment for formative evaluation in group exercises. As a result of the practice, we found that the learners felt peer assessment effective. We found that the conception of trust is important to accept the result of peer assessments.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	1,000,000	300,000	1,300,000
2009年度	500,000	150,000	650,000
2010年度	500,000	150,000	650,000
2011年度	800,000	240,000	1,040,000
年度			
総計	2,800,000	840,000	3,640,000

研究分野：総合領域

科研費の分科・細目：科学教育・教育工学、教育工学

キーワード：ピア・アセスメント、教育評価、可視化

## 1. 研究開始当初の背景

近年、学習者が同じ学習コミュニティに属する他の学習者の成果物を評価する相互評価が盛んになっている。その背景の1つとして学力や評価に対する考え方の変化がある。Wiggins(1989)は、標準テストでは断片化した知識を文脈と切り離して評価している本物でない評価になっているのに対して、大人が現実世界で直面する問題解決場面をシミュレートした課題で評価であるべきとするオ

ーセンティック・アセスメント（真正な評価）の考え方を提唱した。オーセンティック・アセスメントの成立に関した実践としては、Coalition of Essential Schoolの教育評価改革(Sizer, 1984)があり、そこでは、真正の評価として、学習発表会における卒業審査と、ポートフォリオが用いられていた。このような背景から、パフォーマンス評価やポートフォリオ評価など、ペーパーテストに替わる評価手法が新しい評価として、実践されている

(Brown & Hudson, 1998). 我が国では、総合的な学習の時間において、成果物に対して自己評価、相互評価を評価に取り入れた様々な実践が報告されている(例えば、国立教育政策研究所, 2003). 以前から、学習者の観察、成果物の評価、相互評価などペーパーテスト以外の方法による評価も行われることはあったが、ペーパーテストを補完する役割として行われることが多かった。近年では、ペーパーテストで測定が困難である場合や、学習の場面と切り離れた評価を行うことが適切とはいえない場合で学習者間評価が取り入れられている。例えば、従来の教科型の学習のように体系的な知識の獲得を目指した学習ではなく、プロジェクト学習のような問題解決志向の学習場面では、狭義の知識獲得を目指したものではないため、その学習場面での状況と切り離して、学習者の個人内の知識のみを評価の対象とするのではなく、文脈の中で総合的に評価されるべきである。

次に、学習者が行った評価について述べる。全ての学習者が他の全ての学習者を評価する場合は、ほぼ公平であるといえる。しかし、評価者を選択しなければならぬ状況では、たまたま甘い評価者にあたったのか、厳しい評価者にあたったのかによって、学習者に不公平感を与え、評価結果を納得して受け入れることが阻害されると考えられる。したがって、公平な評価を行うためには評価者の評価特性を考慮した評価結果を修正したものを学習者に示すことが必要である。植野ら(2006)は、e-ラーニングで得られた大量の多段階評価の相互評価のデータを項目応答理論の1つである Graded Item Response Model(Samejima, 1969) に評価者の特性パラメータを加えるという拡張をして、レポートの評価を行う方法を提案している。それぞれのレポートに対して同一尺度上で合理的な評価を行うことができ、それぞれの評価者の特性として、評価基準の厳しさ、評価の一貫性を求めることができることが特徴としてあげられる。研究代表者らは、潜在能力は推定せず、評価特性パラメータだけを推定することによって、欠測値の多い、少数のデータに適用できるモデルを提案している。このモデルを拡張することで、クラス集団に適用可能な、学習者の評価能力を定義することを目指している。

## 2. 研究の目的

学習者の他者評価能力を向上させるためには、学習者が自己の評価能力について知る必要がある。本研究では、学習者の他者評価能力を可視化するツールや評価についてのコミュニケーションツールを開発することで、学習者の他者評価能力の向上に資することを目指す。

前者は、研究代表者らが開発した、個々の評価者の評価特性の補正方法をもとに、学習者の評価特性をモデル化する。これまでは、評価特性パラメータをそのままフィードバックしてきたが、より学習者に分かりやすい指標、評価の改善のための指針をグラフィカルに分かりやすく表示する方法を提案し、それを実装したシステムの開発を行う。

後者は、評価を受けた学習者が評価者を評価する仕組みを取り入れたシステムの設計・開発を行う。さらに、評価を受けた学習者が、受けた評価が参考になったかどうかや、評価をした側と評価を受けた側との間でのコミュニケーションツールを導入し、他者評価能力だけでなく、相互評価に期待される教育的効果をさらに向上させることを目指す。

## 3. 研究の方法

本研究では、クラス集団などの大規模でない学習者集団を想定している。今回開発するシステムにおける学習者間の関係を図1に示す。図1では、1人の学習者が、2人の評価を行い、2人から評価される状況で、学習者Aを中心としたインタラクションを表している。一方向の矢印は、成果物の評価を、双方向の→は、評価終了後のコミュニケーションを表している。この例では、学習者Aは、評価者として、学習者B、Cの成果物を評価し、システムが行った評価をもとに学習者Aの評価能力について推定し、フィードバックを受ける。また、評価を受けた学習者B、Cは、今度はその評価が適切であったかについて評価し、学習者B、Cの成果物とその評価について、やりとりを行う。逆に、学習者Aは、学習者D、Eから評価され、学習者B、Cの場合とは逆に、受けた評価の妥当性について評価を行う。このように評価をもとにコミュニケーションが生まれることで、学習コミュニティが活性化させることも期待される。

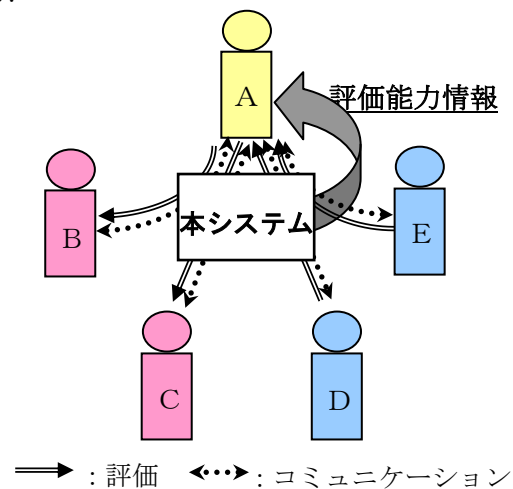


図1 学習者間の関係

次に開発するシステムの構成について述べる。本システムは、図2に示す5つの機能とインタフェースから構成される。学習者間評価の実施機能とし、成果物DBの管理、評価実施の管理、評価結果の管理の各機能があげられる。これらは、以前に開発した相互評価支援ツールの機能を使用する（研究業績欄6）。本研究では、コミュニケーション機能と評価能力推定機能、推定した評価能力を可視化して提示するインタフェースの実装を行う。

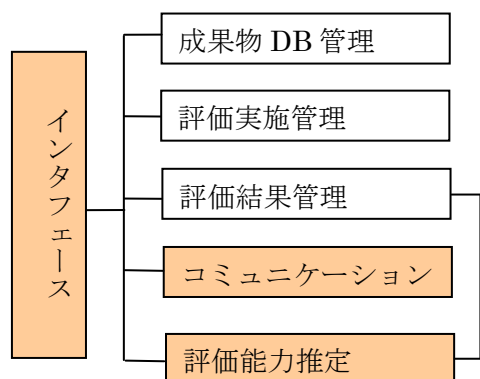


図2 システム構成

本システムでは、学習者間評価を終了した後、評価能力についての推定値をグラフィカルに表示したものと、他者からの評価を表示する。評価能力については、例えば、個人内の評価の一貫性、評価の差をつける（メリハリ）度合い、つけた評価の平均値（甘すぎたり、厳しすぎたりしないか）、他の学習者の行った評価との相関、評価対象の学習者からの評価の妥当性や有用性についての評価などについて、他の学習者が行った評価をもとに算出し、グラフで表示させ、グラフに基づいたアドバイスを表示する。コミュニケーション機能としては、学習者の成果物と学習者の行った他者評価の2つの側面から評価を受ける。成果物の評価では、他の学習者からの評価結果が表示されるだけでなく、評価者とメッセージのやり取りを行い、今後の改善に役立てることができる。また、他の学習者から受けた評価について、評価を行うことができる。他の学習者の成果物に対して行った評価に対しても、同様に他者から評価を評価される。このようなやりとりを通して、客観的に自身の評価能力について理解させるだけでなく、他者に行った評価が受け入れられているかどうかを主観的にも納得させることで、多面的に評価を捉えられ、評価能力の向上に資することができる。

#### 4. 研究成果

2008年度は、学習者の他者評価能力を可視化するツールに関する研究として、学習者

に提示する内容を検討するために、研究代表者らが提案したモデルや綱目応答理論を応用した他のモデルを参考にし、個々の学習者のもつ評価の特性を抽出した。具体的には、安定性、差をつける度合い、評定の厳しさなどのモデル化を行った。評価についてのコミュニケーションツールに関する研究としては、これまでのCSCLのシステムの例も調査し、相互評価の結果を用いて、評価者と評価を受ける人のコミュニケーションのあり方とその実現方法についての設計を行った。

2009年度は、学習者の他者評価能力を可視化するツールに関する研究として、学習者の評価能力を、学習者に分かりやすく表示するためのインタフェースについての設計を行った。コミュニケーションツールに関する研究としては、前年度に引き続き、これまでのCSCLのシステムの例も調査し、相互評価結果を用いて、評価者と評価を受ける人とのコミュニケーションを支援するシステムの設計を行った。

2010年度は、前年度に引き続き相互評価結果を用いて、評価者と評価を受ける人とのコミュニケーションを支援するシステムの実装を行った。評価者と非評価者とのインタラクションを通して、客観的に自身の評価能力について理解させるだけでなく、他者に行った評価が受け入れられているかどうかを主観的にも納得させることを目指すために、コミュニケーションにおけるトラストに関する関連研究調査を行った。

2011年度は、昨年度までに開発したコミュニケーションを支援するシステムについて、予備的評価を実施した。

相互評価を取り入れた授業実践を行い、評価方法による、評価行動の違いについて調べた。1回目の受講生は40名、2回目の受講生は41名であった。2回目では、観点別評価を取り入れた。1回目は114件、2回目は144件の相互評価が実施された。1回目と2回目の差は、一人当たりの評価数と出席者数によるものである。自由記述式のコメント欄には全員が回答し、1回目は181件、2回目は184件の記述が見られた。その内訳を分類すると、1回目は、全体的な印象について意見が89件、体裁や表現についての意見が16件、レポートの条件に対する不足の指摘が27件、内容に関する改善点の提案が49件であった。2回目は、全体の印象が73件、体裁や表現が23件、レポートの条件が23件、改善点が65件であった。2回目は、全体の印象の割合が、1回目よりも減少しているが、別の欄に全体の印象とその理由を記載させたことが影響していると考えられる。また、2回目は、観点別の評価を導入し、不適切を選んだ場合には、その根拠を記入させた。根拠として411件の

記載があった。1つのレポートに対して、1人の評価者が、全11評価項目に対して、2.9項目の指摘を行ったことになる。このことから、観点別評価は、コメントを引き出す効果があったことが示唆された。ワープロのコメント機能を使ってレポート中に記載したコメント数の分布を図3に示す。1回目、2回目ともにコメント数は、評価者やレポートによって差が大きく、1つのレポートに1人が付けた添削箇所の平均値は1回目が4.3箇所、2回目が4.2箇所、中央値は、1回目が3.5箇所、2回目が4.0箇所であった。レポートの記述が不足している場合や、全体にわたって同じ指摘が必要な場合は、添削ではなく、自由記述コメントで対応する傾向が見られた。受講生の行った評価を確認したところ、ほとんどが適切なコメントであり、教員の介入が必要なものはなかった。2回目は、受講生が、観点別評価を記載することで十分であると考へた場合は、添削機能は使用されなかった。そのため、コメント添削がないケースが、1回目と比べて増加した。

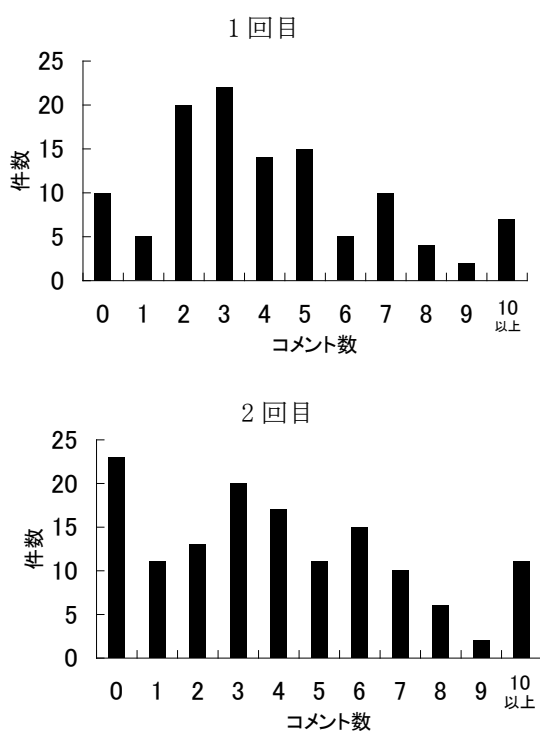


図3 添削コメント数の分布

評価者と非評価者とのインタラクションを通して、客観的に自身の評価能力について理解させるだけでなく、他者に行った評価が受け入れられているかどうかを主観的にも納得させることを目指すために、コミュニケーションにおけるトラストに関する関連研究調査を行った。その結果、学習者を Trustor、評価者を Trustee と対応づけることによって、トラストモデルを適用できることが分かっ

た。評価者をトラストするかどうかは、評価の能力だけでなく、誠実さ、善意が大きな影響を与えており、本研究で取り組んできたコミュニケーションを支援する機能をさらに強化することにより、他者評価がより教育的な効果につながることを示唆された。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

- ① 藤原康宏、大西仁、加藤浩：“継続的な学習者間評価を導入した情報教育の実践”，情報処理学会論文誌，Vol. 49，pp. 3428-3438 (2008)，査読有

[学会発表] (計3件)

- ① 藤原康宏、永岡慶三：“グループワークを取り入れた演習における学習者間レポート相互添削の実践”，電子情報通信学会教育工学研究会 (2010. 11. 26)，東京工業大学
- ② 藤原康宏、永岡慶三：“形成的評価としての学習者間レポート相互添削の実践”，教育システム情報学会第34回全国大会 (2009. 08. 19)，名古屋大学
- ③ 藤原康宏、西久保健太、永岡慶三：“ピアアセスメント支援システムを利用した紙媒体のレポートの相互評価の実践”，電子情報通信学会教育工学研究会 (2008. 07. 19)，福島大学

[その他]

ホームページ

<http://fuji.comm.soft.iwate-pu.ac.jp>

## 6. 研究組織

### (1) 研究代表者

藤原 康宏 (FUJIHARA YASUHIRO)  
岩手県立大学・ソフトウェア情報学部・准教授  
研究者番号：30305338

### (2) 研究分担者

該当なし

### (3) 連携研究者

該当なし