

間主観的評価における自己評価結果整合性に基づく助言法の予備的検討

Preliminary Examination of the Advice Method Based on the Consistency of the Self-Assessment Result for Inter-Subjective Assessments

眞坂 美江子^{*1}, 渡辺 博芳^{*2}, 宮崎 誠^{*3}

Mieko MASAKA^{*1}, Hiroyoshi WATANABE^{*2}, Makoto MIYAZAKI^{*3}

^{*1} 帝京大学 情報電子工学科

^{*1} Graduate School of Science & Engineering, Teikyo University

^{*2} 電気通信大学

^{*2} University of Electro-Communications

^{*3} 帝京大学ラーニングテクノロジー開発室

^{*3} Learning Technology Laboratory, Teikyo University

Email: masaka@ics.teikyo-u.ac.jp

あらまし：ルーブリックやチェックリストを用いた間主観的評価は、学習者が自分の成長を把握し、今後の改善点を見つけることで、持続的な成長を促す効果が期待されている。しかし、学習が指標になれていない状況で間主観的評価をすることは難しく、評価指標作成者の意図と大きく異なる評価をする場合がある。本研究は、学習者の間主観的評価において、評価結果間の不整合を検知し再評価の助言を行うことができないか、その可能性について統計的検証を行った。

キーワード：間主観的評価、ルーブリック、自己評価支援

1. はじめに

ルーブリックやチェックリストを用いた間主観的評価において、継続的な自己評価は、学習者に自身の成長を把握し、今後の改善点を見つけることで、持続的な成長を促す効果が期待されているが、評価すべき指標の増加に伴う、評価疲れが問題となる。我々は、自己評価者の回答パターンに応じて最適な評価指標を提示することにより、より少ない評価指標の数で、すべての評価指標を用いた場合と同等の自己評価結果を得る「適応型自己評価」の開発を進めている[1]。検討を進めていく中で、学習者が評価指標になれていない状況で間主観的評価による自己評価をすることは難しく、評価指標作成者の意図と大きく異なる評価をする学習者が存在していることが明らかとなった。本研究は、この問題に着目し、学習者の間主観的評価において、自己評価結果の整合性を検証することで、不整合が生じている自己評価結果を検知して、再評価を促す助言をできるのではないかと仮説を立て、予備的検討を行った。

2. 帝京大学情報電子工学科における汎用的能力評価

本研究は、帝京大学情報電子工学科で実施しているルーブリックを用いた汎用的能力評価の結果を用いて、再評価を促す助言の実現可能性を検討した。帝京大学情報電子工学科では、9つの汎用的能力を定め、各能力を自己評価するためのルーブリックを作成している。また、各能力の具体的な評価指標として、英語コミュニケーション力を除く8つの能力について8～35項目から成るチェックリストを作成

している。ルーブリックとチェックリストは階層構造となっており、学習者は、対象とする能力について、まずチェックリストを用いて個々の観点で評価を行った後、ルーブリックにより総合的に評価する。

3. 評価支援手法の提案

本研究では、以下の2つの助言法について実現の可能性を検討する。両者とも、学習者が一度行った自己評価結果について学習者群全体の評価結果との比較により整合性をチェックし、必要に応じて再評価を促すことで、適切な自己評価に近づくよう支援しようとするものである。

方法1：チェックリストによる評価結果間の不整合に関する助言

方法2：チェックリストによる評価結果とルーブリックによる評価間の不整合に関する助言

4. 実証分析

帝京大学情報電子工学科では、在学期間4年間にわたって、授業や実習において統一した評価指標を用いて行った自己評価結果が、評価システム内に保存されている。本研究は、汎用的能力の1つである情報リテラシーについて、2020年から2023年に実施した1431回の自己評価結果を用いて分析を行った。情報リテラシーは、13のチェックリストとルーブリックの1つの観点で構成されている。チェックリストは、0:あまりできていない、1:ある程度できている、2:できている、の3段階評価、ルーブリックは、レベル0からレベル4の5段階評価となっている。

4.1 チェックリスト評価結果間の不整合検知

情報リテラシーで定められた 13 のチェックリストについて各チェックリスト間の相関係数を求めたところ、0.23~0.63 と全体的に弱い相関がみられた。チェックリストは、大きく「情報機器の活用(チェックリスト(以下、chk と略す)1~5)」「情報の収集(chk6~9)」「情報の分析(chk10~12)」の3つに分類されており、同一の分類間では他に比べて相関係数が大きい傾向が見られたため、ここでは分類ごとに分け、順序ロジットモデルによるチェックリスト評価結果間の不整合検知を試みた。各分類の最後のチェックリスト (Chk5, Chk9, Chk13) を目的変数、その他のチェックリストの評価結果を説明変数として順序ロジスティック回帰分析を行った結果、すべての分析において説明変数および定数項の P 値は、1%有意であった。また、尤度比検定の結果からも回帰モデルは有意であると判定された。

表1は、順序ロジットモデルによる推定したチェックリストの結果と実際の自己評価の結果における的中率をまとめたものである。chk5 について考察すると、ch5 が生成された順序ロジットモデルにより「2:できている」と推定された場合、99%が、実際の自己評価で「2:できている(73%)」もしくは「1:ある程度できている(26%)」と回答していることが分かる。つまり、モデルによる推定結果が「2:できている」であるにも関わらず、自己評価を「0:できていない」と回答している学習者は、学習者群全体と不整合が生じているとして検知できる。chk9,chk13 についても同様の傾向が見られており、上記の分析結果を用いて、例えば、関連するチェックリストにすべて回答した段階で、相互にチェックリスト評価結果の推定を行い、的中率が一定値を下回るチェックリストに対し、再評価を促す助言ができると考えられる。ただし、評価には偏りがあり今回使用した 13 のチェックリストにおいて、「0:できていない」と自己評価した学生は少数であった。助言を行う際には、各推定結果の信頼度にも考慮し、どのチェックリストに対し助言を行うべきかを決定する必要があることも明らかとなった。

4.2 チェックリストとルーブリック間の不整合検知

次に、学習者がすべてのチェックリストに回答した後、ルーブリック評価を行う際の助言の可能性について検討する。

図1は、情報リテラシーに関する各チェックリストの自己評価結果を0~2の数値に換算し、加算した結果とルーブリック評価の結果をバブルチャートで示したものである。バブルの大きさは人数を示している。ルーブリックの評価が高くなるにつれて、チェックリスト評価の合計値も大きくなる傾向がみられる。しかし、分布から大きく外れる評価をするケースも散見されるため、分布から大きく外れる学習者を抽出し、再評価を促す方法を検討する。

表1 推定結果

推定結果 \ 自己評価結果	chk5			chk9			chk12		
	0	1	2	0	1	2	0	1	2
0	0.50	0.06	0.01	1.00	0.02	0.00	0.76	0.08	0.00
1	0.50	0.71	0.26	0.00	0.74	0.17	0.21	0.75	0.27
2	0.00	0.23	0.73	0.00	0.24	0.83	0.03	0.17	0.73

0:あまりできていない 1:ある程度できている 2:できている

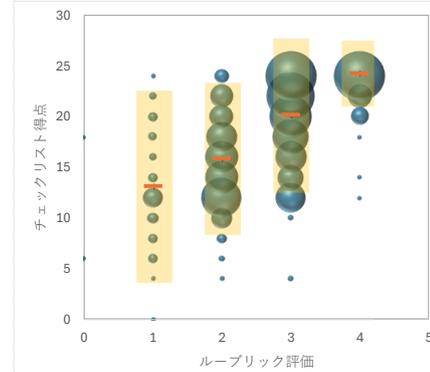


図1 ルーブリック評価とチェックリスト得点

図1に、チェックリストの合計に対するルーブリックレベルごとの平均値を信頼区間95%として区間推定した結果を黄色枠で重ねた。ルーブリック評価の結果が高くなるに従い、区間推定の下側信頼限界、上側信頼限界も高くなる傾向が見られ多くの実データが推定範囲内に含まれていることが確認できた。

以上の結果から、平均値の区間推定を用いて、チェックリストによる評価結果とルーブリック評価の結果が学習者群全体と一致していない学習者を検知し、再評価を促す助言ができると考えられる。

5. まとめ

本研究は、学習者が指標になれていない状況における間主観的評価において、評価指標作成者の意図と大きく異なる評価をする学習者が存在するという課題に対し、自己評価結果の整合性を学習者群全体の結果と比較することで、不整合が生じている自己評価結果を検知して、再評価を促す助言法を提案しその可能性について検証を行った。分析の結果、学習者が行ったチェックリスト評価結果間の不整合、チェックリスト評価結果とルーブリック評価間の不整合の2点において、再評価の助言ができる可能性があることが明らかとなった。

謝辞

本研究は JSPS 科研費 24K06217 及び帝京大学先端総研インキュベーション助成金 23-98 の助成を受けた。

参考文献

(1) 眞坂美江子, 渡辺博芳, 宮崎誠: ルーブリックとチェックリストを用いた間主観的評価における適応型自己評価のための問題定義, 第86回情報処理学会全国大会講演論文集(4), pp.4.265-266, 2024