

課題研究 評価法開発

報告者 研究開発部創意実践課
課長 瀧上 祐太

本校の既 SSH で開発した課題研究の評価基準（巻末の参考資料）を用いて評価基準改善アルゴリズムを開発した。生徒の自己評価と教員評価の認識差を手がかりとして、課題研究における評価基準をどの観点から優先的に改善すべきかを定量的に抽出し、指導改善とルーブリック・評価基準の精緻化に資する独自の評価法を開発した成果を示す。

一 一致度分析に基づく評価基準改善アルゴリズムの開発

目的

ルーブリック評価は、評価観点を明示し、学習目標と評価の対応を可視化できる利点をもつ一方で、同一の評価基準を用いても評価者によって達成判定がずれるという課題を抱えている。とくに探究活動の評価では、観察可能な行動だけでなく、根拠の十分性、研究計画の妥当性、結論の限定条件など、複合的な判断を要する観点が含まれるため、評価者間の認識差が拡大しやすい。この問題は、評価の信頼性を低下させるだけでなく、指導すべき観点と評価基準そのものの改善点を見えにくくする。

本研究で用いた 33 設問は、本校が既 SSH で開発した評価基準であり、探究活動における評価観点を、研究の着想、仮説形成、方法設計、結果整理、考察、発表という進行に沿って配列した段階的な評価設問群である。こうした設問群は、探究活動のどの局面を評価するかを明示できる一方で、各設問について達成要件の共有が十分でなければ、自己評価と他者評価の間に認識差が生じる可能性がある。

そこで本研究の目的は、生徒の自己評価と教員評価の認識差を定量的に分析し、その差が生じやすい設問を抽出することで、課題研究における評価基準のどこを優先的に改善すべきかを体系的に判断する“評価基準改善アルゴリズム（スクリーニング手法）”を構築することである。

探究の評価では、根拠の妥当性、仮説形成、範囲設定、結論の限定条件など質的判断を含む観点において認識差が拡大しやすく、これが評価の信頼性や指導改善の焦点を不明確にする要因となっている。

本研究では、観測一致率、Cohen の κ 、設問別達成率差 Δ_i 、FP 率を指標として認識差の大きさと頻度を捉え、改善優先度を定量的に決定するためのスクリーニング・アルゴリズムを開発することを目的とする。

データと方法

対象は、ポスター番号で生徒自己評価と教員評価を対応付けられるポスター（N=44）である。各設問は 2 件法（達成=1、未達=0）として扱い、原データ上で達成が複数水準で記録されている場合は「0 を未達、0 より大を達成」として 2 値化した。グループ研究のため、同一ポスターに複数の生徒回答または複数の教員回答が存在する。そこで、ポスター単位で設問ごとの達成割合を算出し、達成割合が 0.5 以上を達成（1）、0.5 未満を未達（0）として平均値を代表値と定めた。

比較単位は「1 ポスター×1 設問」であり、設問は 33 項目であるため、比較総数は $44 \times 33 = 1,452$ 比較である。

指標の定義

① 観測一致率 p_0 : 生徒と教員の判定が一致した割合であり、 $p_0 = \frac{\text{一致数 (両者達成+両者未達)}}{\text{比較総数}}$ と定義した。

② 混同行列：生徒判定を基準として以下を定義した。

区分	判定条件	内容の意味（解釈）
TP (True Positive)	生徒＝達成 (1) 教員＝達成 (1)	生徒・教員ともに「達成」と判断した一致。肯定的評価が一致している領域。
TN (True Negative)	生徒＝未達 (0) 教員＝未達 (0)	生徒・教員ともに「未達」と判断した一致。 改善の必要性について認識がそろっている領域。
FP (False Positive)	生徒＝達成 (1) 教員＝未達 (0)	生徒は「できている」と判断したが、教員は「未達」と判定した不一致。 認識差が最も生じやすい領域で、基準理解のズレを示す。
FN (False Negative)	生徒＝未達 (0) 教員＝達成 (1)	生徒は「できていない」と判断したが、教員は「達成」と判定した不一致。 生徒が自分の達成を過小評価したケース。

さらに、達成率の偏りによって偶然に一致する分を補正するため、Cohen の κ を用いた。

$$\kappa = \frac{p_0 - p_e}{1 - p_e}$$

ここで偶然一致率 p_e は、

$$p_e = p_S p_T + (1 - p_S)(1 - p_T)$$

とした。

p_S は全比較 1,452 件において生徒が達成 (1) と判定した比率、 p_T は全比較 1,452 件において教員が達成 (1) と判定した比率である。すなわち、

$$p_S = \frac{\text{生徒が達成と判定した総数}}{1452}, \quad p_T = \frac{\text{教員が達成と判定した総数}}{1452}$$

である。

③ 設問別指標

設問別の認識差を直接に捉えるため、達成率差 Δ_i と FP 率を導入した。設問 i における生徒達成率 $p_{S,i}$ は、44 ポスター中で生徒側代表値が達成 (1) となった割合、教員達成率 $p_{T,i}$ は、44 ポスター中で教員側代表値が達成 (1) となった割合である。すなわち、

$$p_{S,i} = \frac{\text{設問 } i \text{ において生徒側が達成と判定したポスター数}}{44}, \quad p_{T,i} = \frac{\text{設問 } i \text{ において教員側が達成と判定したポスター数}}{44}$$

と定義した。

これに基づき、達成率差 Δ_i を

$$\Delta_i = p_{S,i} - p_{T,i}$$

と定義した。 Δ_i が正で大きいほど、その設問では生徒の方が教員よりも達成と判定しやすいことを示す。

また、設問 i における FP 率は、

$$\text{FP 率}_i = \frac{FP_i}{44}$$

と定義した。ここで FP_i は、設問 i において「生徒 = 達成、教員 = 未達」となったポスター数である。したがって、FP 率は、認識差が実際の評価場面でどの程度頻発しているかを示す。

③ 介入優先度の段階化

本分析では、 Δ_i を L 段階、FP 率を F 段階に区分し、 $L \times F$ のマトリクスによって指導設計への介入優先度を A ~ E に分類した。閾値は、本データの観測分解能 $1/44 \approx 0.023$ と実測分布に基づき、0.10、0.25、0.40 を境界として設定した。これらは、おおむね 5 件、11 件、18 件規模の差または不一致に対応するため、散発的な差、一定程度みられる差、頻発する差、高頻度の差を区別する実務的区分として解釈できる。

本データに基づく、指導設計における介入優先度表(N=44)

$\Delta_i \setminus \text{FP}$	F1 (FP < 0.10)	F2 (0.10 ≤ FP < 0.25)	F3 (0.25 ≤ FP < 0.40)	F4 (0.40 ≤ FP)
L4 (0.40 ≤ Δ_i)	C (注意)	C (注意)	B (優先)	A (最優先)
L3 (0.25 ≤ Δ_i < 0.40)	C (注意)	C (注意)	B (優先)	A (最優先)
L2 (0.10 ≤ Δ_i < 0.25)	D (安定)	D (安定)	C (注意)	B (優先)
L1 (0 ≤ Δ_i < 0.10)	D (安定)	D (安定)	C (注意)	C (注意)
L0 (Δ_i < 0)	E (逆方向)	E (逆方向)	E (逆方向)	E (逆方向)

結果① 全設問をプールした一貫性 (プール 33 項目)

全設問をプールした観測一致率は $p_0 = 0.613$ であった。混同行列における不一致は 562 件 (=FP+FN) であり、その内訳は FP=438 件 (77.9%)、FN=124 件 (22.1%) であった。したがって、不一致は主として「生徒は達成と判断したが教員は未達と判断した (FP)」方向に偏っていた。

全比較を通じた達成率 (全ポスター・全項目平均) は、生徒側 $p_S = 0.838$ 、教員側 $p_T = 0.622$ であり、生徒の達成判定は教員より 0.216 高かった。Cohen の κ は 0.073 であり、周辺比率の偏りを考慮すると、偶然一致を差し引いた正味の一致は小さい水準にとどまった。

表1 全体一致度 (N=44、33項目)

指標	値
比較総数	1,452
観測一致率 p_0	0.613
Cohen の κ	0.073
生徒達成率 (平均) p_s	0.838
教員達成率 (平均) p_T	0.622
TP (生徒=達成、教員=達成)	779
TN (生徒=未達、教員=未達)	111
FP (生徒=達成、教員=未達)	438
FN (生徒=未達、教員=達成)	124
偶然一致率 p_e	0.582443
不一致数 (FP+FN)	562
FP 比率 (FP/(FP+FN))	0.779359

設問別結果② (達成率差 Δ_i と FP 率による原因候補のスクリーニング)

設問別に達成率差 Δ_i と FP 率を算出したところ、 Δ_i が大きく、かつ FP 率が高い設問は、先行研究比較、仮説の根拠、範囲設定、結果予測、対照設定、結論の意味づけ、適用限界、今後の課題の絞り込み、データ計画など、研究の妥当化や根拠要件に関わる観点に集中した。

たとえば、設問 11「先行研究と比較して仮説に独創性があることを確認している」は $\Delta_i = 0.43$ 、FP 率 = 0.52、設問 9「仮説を検証するための適当な範囲または対象を決定している」は $\Delta_i = 0.41$ 、FP 率 = 0.48、設問 24「今後の課題が現在の結論をもとに十分に絞り込まれている」は $\Delta_i = 0.41$ 、FP 率 = 0.43 であり、最優先介入群 A に分類された。

一方、設問 12「実験や調査の手順が簡潔にまとめられている」は $\Delta_i = 0.02$ 、FP 率 = 0.09、設問 32「予定した内容を制限時間内に発表できた」は $\Delta_i = 0$ 、FP 率 = 0.02 であり、認識差が小さい設問であった。また、設問 30「原稿から目を離しアイコンタクトしていた」では $\Delta_i = -0.20$ 、FP 率 = 0.09 であり、生徒の自己評価が教員評価より低い逆方向のずれが確認された。

以上より、認識差は全設問に一樣に生じるのではなく、設問が要求する判断の性質に応じて偏在することが示された。

評価の一致性

設問	内容	p_0	p_s	p_T	Δ_i	FP 率	介入優先度
1	好奇心を喚起する研究内容である	0.82	1	0.82	0.18	0.18	D
2	テーマを深く追究する意欲があふれる研究内容である	0.59	0.93	0.61	0.32	0.36	B
3	研究への熱意が他者へ伝わる発表内容である	0.57	0.86	0.61	0.25	0.34	B
4	研究に対する真摯な態度が他者へ伝わる発表内容である	0.68	0.96	0.73	0.23	0.27	C
5	研究動機が明確である	0.89	0.98	0.91	0.07	0.09	D
6	仮説が明確である	0.68	0.96	0.73	0.23	0.27	C
7	仮説を立てるために十分な情報が提示されている	0.41	0.71	0.34	0.36	0.48	A
8	仮説を検証するための正確なデータを収集するように計画されている	0.43	0.77	0.52	0.25	0.41	A
9	仮説を検証するための適当な範囲または対象を決定している	0.46	0.86	0.46	0.41	0.48	A
10	仮説の検証結果について適切な予測を立てている	0.55	0.84	0.48	0.36	0.41	A
11	先行研究と比較して仮説に独創性があることを確認している	0.39	0.77	0.34	0.43	0.52	A
12	実験や調査の手順が簡潔にまとめられている	0.84	0.89	0.86	0.02	0.09	D

13	制御変数と操作変数を明確に定義している	0.64	0.84	0.57	0.27	0.32	B
14	計画通りにデータを集めている	0.71	0.89	0.71	0.18	0.25	B
15	適切な場面で対照実験または比較調査を行っている	0.52	0.82	0.48	0.34	0.41	A
16	不適切なデータ取得時に繰り返し実験または再調査を行っている	0.59	0.73	0.48	0.25	0.32	B
17	適切な範囲で信頼できる証拠を集めている	0.61	0.86	0.59	0.27	0.32	B
18	実験や調査の結果が提示されている	0.8	0.89	0.73	0.16	0.18	D
19	結果の傾向またはパターンを特定している	0.73	0.86	0.68	0.18	0.23	D
20	結果が表・グラフで客観的に整理されている	0.77	0.89	0.71	0.18	0.21	D
21	結果から適切な結論を導き出している	0.71	0.86	0.66	0.21	0.25	B
22	研究全体を理解した上で結論の意味を考察している	0.41	0.82	0.5	0.32	0.46	A
23	科学的判断に基づき結論の適用限界を考察している	0.41	0.66	0.39	0.27	0.43	A
24	今後の課題が現在の結論をもとに十分に絞り込まれている	0.55	0.86	0.46	0.41	0.43	A
25	ストーリーの展開が理解しやすいよう工夫されている	0.75	0.91	0.8	0.11	0.18	D
26	説明方法が理解を促進するよう工夫されている	0.71	0.89	0.73	0.16	0.23	D
27	要点を明確に伝えることができた	0.8	0.93	0.86	0.07	0.14	D
28	研究の社会的価値を適切に伝えることができた	0.61	0.84	0.61	0.23	0.27	C
29	聴衆のニーズを把握しそれに基づく結果提示ができた	0.5	0.68	0.59	0.09	0.3	C
30	原稿から目を離しアイコンタクトしていた	0.66	0.68	0.84	-0.2	0.09	E
31	スライド/ポスターのデザインが理解を促進するよう工夫されている	0.66	0.89	0.64	0.25	0.3	B
32	予定した内容を制限時間内に発表できた	0.91	0.93	0.98	-0	0.02	E
33	質問に分かりやすく適切に対応できた	0.8	0.91	0.84	0.07	0.14	D

考察 (Δ_i ×FP 率にもとづく介入優先度モデルの含意と妥当性)

本分析では、生徒の自己評価と教員評価の認識差を、設問別の達成率差 Δ_i と FP 率によって定量化し、指導改善および評価基準精緻化の対象を抽出した。ここで重要なのは、こうした課題が個別の設問文の問題にとどまらず、ルーブリック評価一般が抱える評価者間不一致の問題と連続している点である。ルーブリック評価は、観点や到達目標を可視化できる一方、各観点の達成要件の解釈が評価者間で必ずしも一致しない場合、同一の基準を用いても判定がずれる。そのため、評価の明示性それ自体は、判断の一致を自動的に保証しない。

本研究で用いた設問群も、探究活動の各局面を段階的に配列しているが、設問が並んでいること自体と、その達成要件が評価者間で共有されていることは別問題である。実際、本分析では全体の不一致が FP 方向、すなわち「生徒は達成と判断するが、教員は未達と判断する」方向に強く偏っていた。これは、単なる採点誤差ではなく、達成とみなすための条件や判断基準が、生徒と教員の間で十分に一致していない可能性を示す。

この点を把握するために導入した Δ_i は、生徒達成率と教員達成率の差として定義され、生徒の方が教員より達成と判定しやすい方向の系統差の大きさを表す。一方、FP 率は、その差が実際の評価場面でどの程度の頻度で不一致として出現しているかを示す。したがって、 Δ_i が差の大きさを、FP 率が差の頻度を表すため、両者を併用することにより、認識差を単なる一致率以上に具体的に把握できる。すなわち、本手法は、評価の不一致を「存在する」と述べるだけでなく、どの設問で、どの方向に、どの程度、どれほど頻繁に生じているかを分解して示す点に意義がある。

また、本研究では 0.10、0.25、0.40 を閾値として段階化したがる、これらは恣意的に設定した値ではない。N=44 における観測分解能は $1/44 \approx 0.023$ であり、0.10、0.25、0.40 はそれぞれ、おおむね 5 件、11 件、18 件規模の差または不一致に相当する。したがって、これらの閾値は、微小な差と実務上無視しにくい差とを区別する境界として解釈できる。さらに、実測範囲は Δ_i が $-0.159 \sim 0.432$ 、FP 率が $0.023 \sim 0.523$ であり、設定した

各区分に実際の設定問が配分されうる分布となっていた。このことから、本区分は、設定問を介入優先度別に整理する規則として機能していると判断できる。

このような前提のもとで、 L (Δ_i の大きさ) \times F (FP 率の高さ) のマトリクスによる介入優先度 (A~E) は、「ズレが大きく、かつ頻発する設定問」を最優先 (A) として抽出し、「ズレはあるが頻度が低い設定問」や「頻度はあるが差が小さい設定問」を次順位 (B/C) として整理できる。したがって、このマトリクスは、認識差を単に記述するための表ではなく、どの設定問を優先的に調整すべきかを示す介入優先度表として機能する。

実際に、 Δ_i と FP 率がともに大きかった設定問は、「仮説の根拠」、「対象範囲の妥当性」、「対照設定」、「結論の意味づけ」、「適用限界」、「次課題の絞り込み」など、研究の妥当化に関わる観点に集中していた。これらの設定問では、生徒は「記述した」「実施した」という事実にもとづいて達成と判断しやすい一方、教員は「十分な根拠があるか」「論理が整合しているか」「限定条件まで示されているか」といった質的要件を含めて判定している可能性が高い。そのため、不一致は単なる採点のばらつきではなく、達成要件の共有不足に由来すると考えられる。

一方、「手順整理」、「結果提示」、「時間内発表」など、観察可能な行動や成果物に直接対応する設定問では、 Δ_i と FP 率は小さかった。これらは評価者が参照する情報が比較的一致しやすく、判断基準も共有されやすいためである。したがって、認識差は評価全体に一樣に存在するのではなく、設定問が要求する判断の性質、とくに質的要件の複合性に依りて偏在することが示された。

以上より、本手法は、ルーブリック評価一般が抱える評価者間不一致の課題に対して、その不一致を設定問単位で可視化し、指導改善および評価基準精緻化の優先順位へ接続する方法として有効である。すなわち、評価基準の明文化、必須要件の列挙、到達例・非到達例の提示、指導時の説明重点の再配置といった改善を、どの設定問から優先的に行うべきかを定量的に示す点に、本手法の実践的意義がある。

ただし、本分析は 2 値化したデータとポスター単位の代表値に基づいており、達成程度の濃淡やグループ内の評価分散までは十分に反映していない。また、閾値は本データの標本数と分布に適合するように設定したものであり、そのまま他年度・他校・他教科に一般化できるとは限らない。したがって、今後は、複数年度・複数集団への適用、相互評価や外部評価との比較、多段階尺度での再分析を通して、閾値の安定性、分類規則の再現性、手法の汎用化可能性を検証する必要がある。

まとめ

本研究では、生徒の自己評価と教員評価の認識差を、達成率差 Δ_i と FP 率によって定量化し、その差が生じやすい評価観点を抽出した。これにより、認識差は全体に一樣に生じるのではなく、仮説の根拠、対照設定、適用限界、次課題の絞り込みなど、探究の妥当化に関わる質的観点到に偏在することが明らかとなった。一方、手順整理や結果提示など観察可能な観点では認識差が小さく、到達状況が共有されやすいことも示された。

これらの指標を用いて Δ_i (差の大きさ) \times FP 率 (差の頻度) のマトリクスを構築し、A~E の改善優先度分類として整理することで、評価基準のどの観点を優先的に改善すべきかを体系的に決定できる“評価基準改善アルゴリズム (スクリーニング手法)”を開発した。

このアルゴリズムは、ルーブリック評価でしばしば問題となる評価者間不一致を、単なる指摘にとどめず“改善すべき観点的優先順位”として可視化することを可能にした点に意義がある。今後は、抽出された A 群観点を中心に、達成要件の明確化、到達例・非到達例の提示、指導の重点化といった改善を進めることで、評価の信頼性向上と指導改善の双方に資する枠組みとして発展させることが期待される。

参考資料

(1) 英国 GCSE における「コースワーク」による課題研究の評価

課題研究の評価は、下表に示すコースワークの評価基準①～④と、⑤校内活動日数、⑥校外連携活動回数、⑦外部発表回数、⑧コンテスト受賞数、の全8項目で評価した。評価は、複数の教員によって行った。論文評価は、各自がコースワークの指標に従って採点をした後、持ち寄って、能力領域の得点を決定した。判断に迷う際は、担当教員が、生徒の研究を説明、他の教員が質問することによって得点を決定した。

①能力領域 P (Planning) 「計画すること」

点	評価	評価基準
8点	P.8a	ある適切な手法を計画し伝えるために詳細な科学的知識と理解を用いており、そこに、正確で信頼できる証拠を生成する必要性と、予測をした場合に予測を正当化する必要性を考慮している。
	P.8b	計画を述べるために、適切な場合に先行研究から関連する情報を用いている。
6点	P.6a	ある手続きを計画し伝えるために科学的知識と理解を用い、重要な諸要因を特定したり、変化させたり、抑制したり、考慮に入れたり、また、適切な場合に予測を行っている。
	P.6b	証拠を収集するために、ある適当な範囲と大きさを決定している。
4点	P.4a	根拠を確実にするような証拠を収集するように計画している。
	P.4b	証拠に関してふさわしい準備物や情報源の利用を計画している。
2点	P.2a	単純な手順を説明している。

②能力領域 O (Obtaining Evidence) 「証拠を得ること」

点	評価	評価基準
8点	O.8a	ある適切な範囲で信頼できる証拠を得たり記録したりするための精確な手順と技能を用いている。
6点	O.6a	十分に体系的で正確な証拠を集め、また、適切な場面で繰り返しや確認をしている。
	O.6b	収集した証拠を明確にかつ正確に記録している。
4点	O.4a	活動に十分で適切な証拠を収集している。
	O.4b	証拠を記録している。
2点	O.2a	ある単純で安全な手順を用いていくらか証拠を収集している。

③能力領域 A (Analyzing and Considering Evidence) 「証拠を分析し考察すること」

点	評価	評価基準
8点	A.8a	詳細な科学的知識と理解を用いて、証拠を処理して導かれた根拠の確かな結論を説明している。
	A.8b	予測がなされている場合に、どの程度まで結論がその予測を支持するかを説明している。
6点	A.6a	ある結論に向けて証拠を処理するために、ふさわしい図解や図表、グラフ（適切な場合に最適にあてはまる線が引かれている）、あるいは数字で表す方法を作ったり用いたりしている。
	A.6b	証拠に合致するある結論を導き、それを科学的知識と理解を用いて説明している。
4点	A.4a	証拠を説明するための基礎として単純な図解や図表やグラフを用いている。
	A.4b	証拠中の傾向とパターンを特定している。
2点	A.2a	証拠によって何が示されるかを簡潔に述べている。

④能力領域 E (Evaluating) 「評価すること」

点	評価	評価基準
6点	E.6a	証拠の信頼性と、それが結論を支持するに十分かどうかについて、変則を説明しながら、批評的に考察している。
	E.6b	付加的に関連する証拠を与える更なる研究について、詳細に記述している。
4点	E.4a	何らかの変則を特定しながら、証拠の質について批評している。
	E.4b	手順の適切性について批評し、また不適切な場合に、それを改善するための変更点を示唆している。
2点	E.2a	用いた手順や得られた証拠に関連した批評を行っている。

(2) 本校で独自に開発した評価基準による課題研究の評価

平成 22~30 年度のコア SSH および重点枠等で国内外の連携校と協力して開発した課題研究の評価基準である。各基準について「良 1 と悪 0」の 2 段階による自己評価と高校教員による他己評価の 5 領域「態度、計画、証拠の収集法、分析と考察、発表技術」の結果を得点とした。

本校 SSH が連携校と開発した評価基準

領域	No	評価基準	良・悪	
態度	1	「知る」ことに対する好奇心があふれる研究内容である。	1・0	
	2	テーマを深く追究する意欲があふれる研究内容である。	1・0	
	3	研究への熱意が他者へ伝わる発表内容である。	1・0	
	4	研究に対する真摯な態度が他者へ伝わる発表内容である。	1・0	
研究プロセス	計画	5	研究動機が明確である。	1・0
		6	仮説が明確である。	1・0
		7	仮説を立てるために十分な情報が提示されている。	1・0
		8	仮説を検証するための正確なデータを収集するように計画されている。	1・0
		9	仮説を検証するための適切な範囲または対象を決定している。	1・0
		10	仮説の検証結果について適切な予測を立てている。	1・0
		11	先行研究と比較して仮説に独創性があることを確認している。	1・0
	証拠の収集法	12	実験や調査の手順が簡潔にまとめられている。	1・0
		13	変えない条件（制御変数）と変える条件（操作変数）を明確に定義している。	1・0
		14	計画通りにデータを集めている。	1・0
		15	適切な場面で対照実験または比較調査を行っている。	1・0
		16	不適切なデータを取得した際に繰り返し実験または再調査を行っている。	1・0
		17	適切な範囲で信頼できる証拠を集めている。	1・0
	分析と考察	18	実験や調査の結果が提示されている。	1・0
		19	実験結果または調査結果の傾向またはパターンを特定している。	1・0
		20	実験や調査の結果がグラフや表を用いて客観的に整理されている。	1・0
		21	実験や調査結果から適切な結論を導き出している。	1・0
22		研究全体を正確に理解した上で結論の持つ意味を考察している。	1・0	
23		科学的な判断に基づいて結論の適用の限界について考察している。	1・0	
24		今後の課題が現在の結論をもとに十分に絞り込まれている。	1・0	
発表技術	25	ストーリーの展開は聴衆が容易に理解できるように工夫されている。	1・0	
	26	説明方法は聴衆の理解を促進するように工夫されている。	1・0	
	27	要点を明確に伝えることができた。	1・0	
	28	研究の社会的な価値を適切に伝えることができた。	1・0	
	29	聴衆のニーズを的確に把握し、それに基づいた研究結果を提示できた。	1・0	
	30	原稿から目を離して、聴衆とアイコンタクトしていた。	1・0	
	31	スライドのデザインは聴衆の理解を促進するように工夫されている。	1・0	
	32	予定していた内容を制限時間内（終了のベルまで）に発表できた。	1・0	
	33	聴衆からの質問に分かりやすくかつ適切に対応できた。	1・0	