

令和5年度中学校英語ライティングパフォーマンステストに関する調査結果報告書

青森公立大学 教授 丹藤 永也

1 回答数

回収率 26.5% (151校中40校) 回答者数：60人

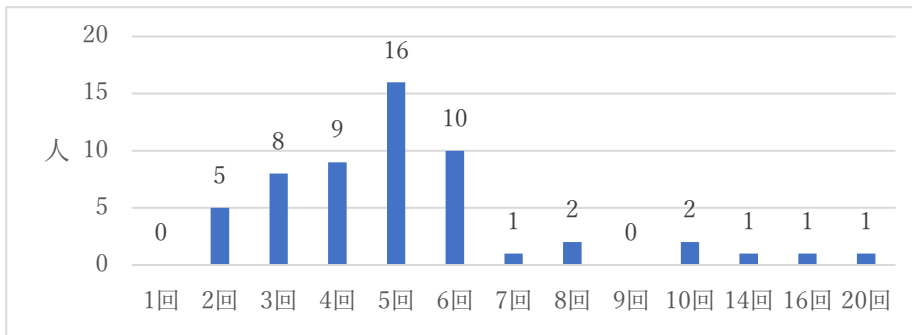
内訳 1学年：16人 (26.7%) 2学年：13人 (21.7%) 3学年：31人 (51.7%)

2 実施状況

実施した：56人 (93.3%) 内訳 1学年：16人 2学年：12人 3学年：28人

実施しなかった：4人 (6.7%) 内訳 2学年：1人 3学年：3人

3 回数



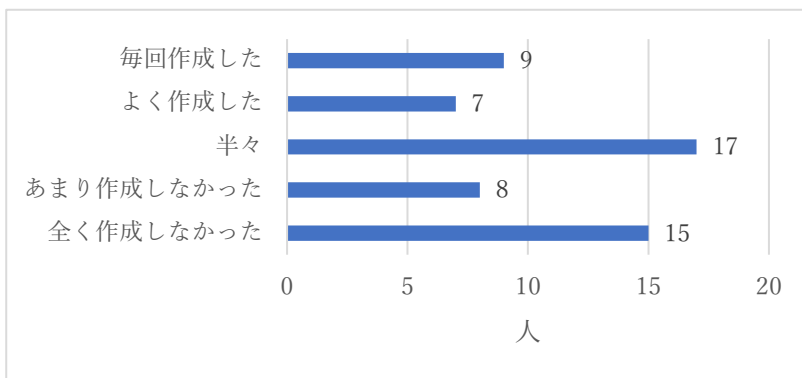
全体の回数：303回 1人あたり平均回数：5.4回

最多が5回で16人 (28.6%)、次いで6回 (17.9%)、4回 (16.1%)

定期テストを活用した人数：45人 (80.3%)

全体における定期テストの割合：57.8% (175回)

4 ルーブリックの作成



毎回・よく作成した：16人 (28.6%)

半々：17人 (30.4%)

あまり・全く作成しなかった：23人 (41.1%)

5 パフォーマンステストの評価者（複数回答 総回答数 65）

教科担任：34人（60.7%） 手分け：16人（28.6%） その都度：12人（21.4%）
ALT：3人（5.4%）

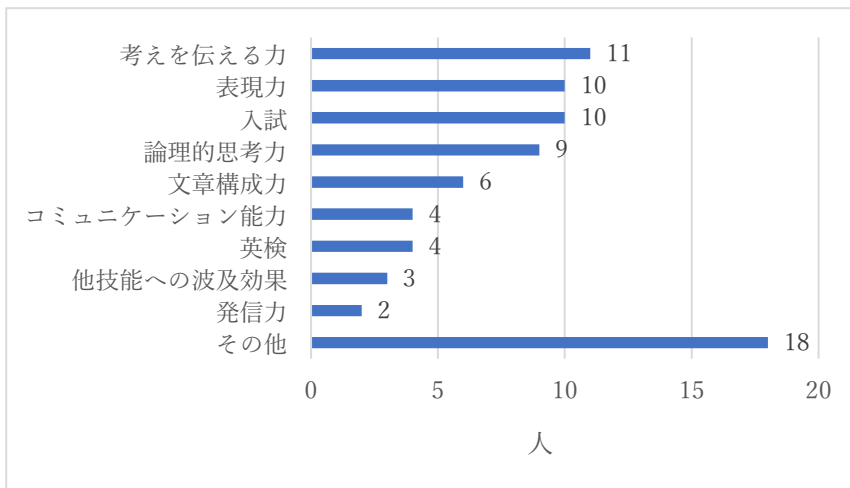
6 技能統合型（英語を聞いて／読んで／話して→書く）ライティングパフォーマンステストの回数

実施率：51.8%（29人）
全体における技能統合型の割合：36.0%（109回）

7 まとまりのある英文の割合と重要度

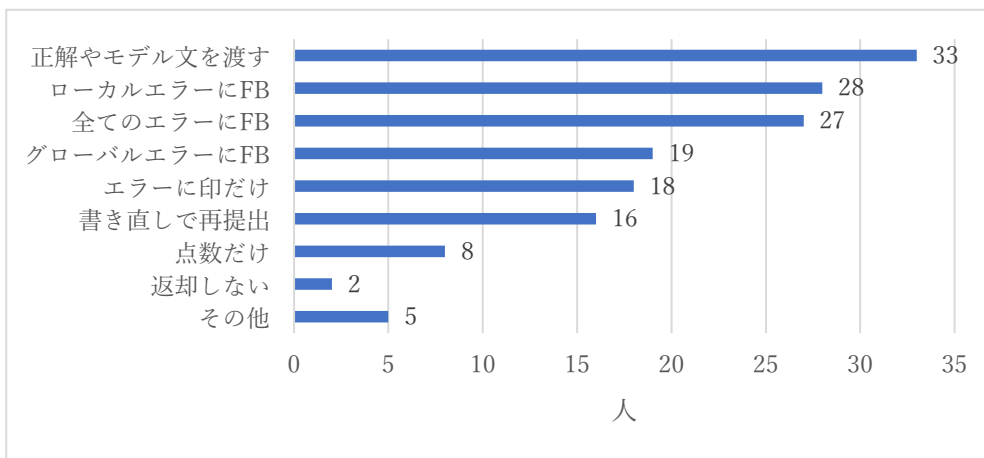
全体における割合：72.9%（221回）
とても重要：41人（73.2%） 重要：11人（19.6%） 半々：4人（7.1%）

8 7の回答の理由（複数回答 総回答数 77）



思考力・構成力：15人（19.5%） 伝達力・コミュニケーション能力：15人（19.5%）
試験等：14人（18.2%） 表現力・発信力：12人（15.6%）

9 英作文に対するフィードバック（FB）の方法（複数回答 総回答数 156）



直接FB：74人（47.4%） 間接FB：51人（32.7%） FBなし：10人（6.4%）

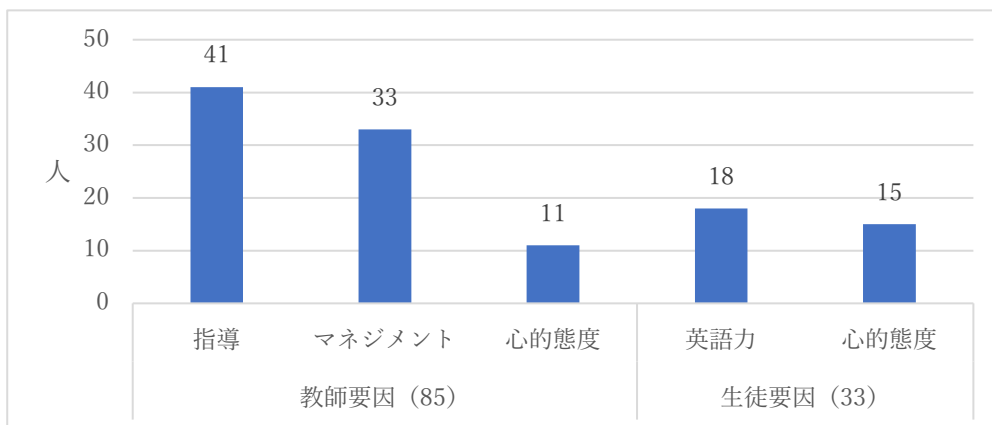
10 採点・添削・フィードバックへの自信

自信・ある少しある：16人（28.6%）

半々：27人（48.2%）

自信ない・あまりない：13人（23.2%）

11 ライティングパフォーマンステストの困難点（複数回答 総回答数 118）



教師要因：72.0%（85人）

指導：34.7%（41人） マネジメント：28.0%（33人） 心的態度：9.3%（11人）

生徒要因：28.0%（33人）

英語力：15.3%（18人） 心的態度：12.7%（15人）

教師側に要因がある困難点の詳細（人）

指導 (41)	マネジメント (33)	心的態度 (11)
評価の信頼性のぶれ (10)	採点・添削・評価の時間 (22)	採点に自信がない (9)
採点・添削 (10)	活動・テストの時間設定 (7)	英語力に自信がない (2)
指導 (7)	ALTの都合 (3)	
課題設定 (7)	返却までの時間 (1)	
個人差対応 (4)		
評価規準・ループリック (3)		

生徒側に要因がある困難点の詳細

心的態度 (15)	英語力 (12)
英作文が嫌い (6)	誤りが多い (5)
自信がない (6)	個人差が大きい (4)
必要性を感じていない (3)	添削が理解できない (3)

12 まとまりの英文の採点で設定する観点と配点

100点満点でまとまりのある英文を採点する際の観点と配点を決めてもらい、このデータを被験者内要因での1元配置分散分析を行った結果、 $F(4.78, 263.04) = 23.89, p < .001, \eta^2 G2 = .302$ で有意であり、 $\eta^2 G2$ 効果量も大きかった。そこで、Holmの方法を用いて多重比較を行ったところ、有意差が現れたのは以下の通りである。

・グローバル

内容 > 構成 > 一貫性 = 論理性 = 談話 = 独創性

・ローカル

文法 > 文構造 > 語彙 = 綴り > 語数 = 全体の正確さ

・グローバル vs. ローカル

内容 > 文法/文構造/語彙/綴り/語数/全体の正確さ

構成 > 語数/全体の正確さ

文法 > 一貫性/論理性/談話

語彙 > 談話/独創性

綴り > 談話/独創性

グローバルな観点では内容を最も重視し、次いで構成、さらに一貫性、論理性、談話、独創性（この4つに差はない）を重視していることが明らかになった。ローカルな観点では文法を最も重視し、次いで文構造、その次に語彙と綴り、さらに語数と全体の正確さの順で重視していることが明らかになった。そして、内容は、すべての他の観点との間に有意差が見られた ($p < .05$)。つまり、まとまりの英文の採点において、教員は内容を最も重視しているということが示唆された。

	観点の人数	合計	合計得点の割合	観点の平均	標準偏差
グローバル					
1. 内容	49/56	1660	29.6	33.88	17.99
2. 構成	35/56	700	12.5	20.00	11.79
3. 一貫性	16/56	300	5.4	18.75	9.24
4. 論理性	14/56	300	5.4	21.43	10.57
5. 談話	2/56	35	0.6	17.50	3.31
6. 独創性	1/56	20	0.4	20.00	2.67
ローカル					
7. 文構造	25/56	335	6.0	13.40	7.74
8. 文法	43/56	795	14.2	18.49	10.76
9. 語彙	32/56	437.5	7.8	13.67	9.22
10. 綴り	31/56	377.5	6.7	12.18	8.84
11. 語数	3/56	70	1.3	23.33	5.74
12. 全体の正確さ	9/56	250	4.5	27.78	11.07
13. その他	4/56	320	5.7	80.00	22.79

13 考察

・ライティングパフォーマンステストの実施率 93.3%は、文部科学省令和4年度「英語教育実施状況調査」の全国の割合 93.3%と同じであった。この全国の割合は、スピーキングとライティング両方とライティング単独を合わせた数値であり、本県では十分実施されていると判断できる。ただし、実施状況には個人差が大きく、教師間、学校間での学力等の差が懸念される。

・ライティングパフォーマンステストの回数は1人あたり5.4回であった。全国は1校あたりの平均が4.7回であった。単位は違うが、回数も十分実施されていると判断できる。ただし、定期テストで実施

している割合が 57.8%と高く、適切なフィードバックが行われているかが懸念される。

・ルーブリックの作成率が、「毎回・よく作成した」が 28.6%と低い。ルーブリックは評価の妥当性、信頼性を確保するのに必要不可欠なものであるため、簡略化したものでいいので、作成する必要がある。そうでなければ教師の主観的で偏った評価が行われてしまう恐れがある。

・評価者は教科担任単独が 60.7%で最も多かった。実用性の面から理解できるが、評価の信頼性の問題がある。そもそも 1 人だと、評価基準（観点や割合）が正しいかどうかを検討されていないことや、評価の信頼性（最初の生徒と最後の生徒の評価に差がある）などの懸念が生じる。自分の取組について検証する機会を設ける必要がある。

・技能統合型パフォーマンステストの実施率は 51.8%で、パフォーマンステスト全体からの割合は 36.0%であった。技能統合型の場合、最初の技能で理解できない場合、肝心のライティングができないことも想定されるため、難しい側面もあるが、令和 5 年度の「全国学力・学習状況調査」でも出題されていることから、この型への対応は今後の重要課題の 1 つであると言える。

・まとまりのある英文の実施は、パフォーマンステスト全体の 72.9%であった。また、まとまりの英文を、「とても重要・重要」とした回答は 92.9%（52 人）と高かった。さらに、まとまりの英文がなぜ重要かという理由では、思考力・判断力と伝達力・コミュニケーション能力が同じく 19.5%（15 人）、表現力・発信力が 15.6%（12 人）であった。これらのことから、まとまりの英文を書くパフォーマンステストは十分認識され、適切な目的のもと実施されていると判断できる。また、採点する際の観点では、1 元配置一分散分析の結果、内容や文法が特に重視されていることが明らかになった。まとまりの英文でなければ確認できない一貫性や論理性などに重きを置く必要があると考える。

・フィードバック（FB）に関しては、エラーを直接修正する直接 FB が 47.4%、印や記号で間接的にエラー箇所示す間接 FB が 32.7%を占めた。FB なしも 6.4%あった。FB に関しては生徒の実態に合わせたものを組み合わせて行う必要があり、生徒が誤りの原因に気づき、修正点を理解し、主体的に自己調整を図るようにすることが重要であるため、簡単でもいいので FB は行うべきであると考え。

・自分の採点・添削・フィードバックに「自信がある・少しある」教員は 28.6%（16 人）と低く、「自信がない・あまりない」教員も 23.2%（13 人）と相当数あった。テストの困難点にも挙げられているが、具体的には、自分の作成した評価基準が適切なかわからない、生徒が書いた表現が正しいか判断できない、英語力に自信がない等が挙げられ、改善が急務であると考え。このような課題には演習形式の研修が有効であると言われているため、行政側がそのような機会を設定することが求められる。

・困難点では、教員側に多くの問題があると感じている教員が多かった。最も多かったのが指導についてで、これらを解決する研修の設定が重要であると同時に、教員が自己調整を図りながら取り組んでいく姿勢も必要であると考え。

14 さいごに

全国の中学校 3 年生を対象とした令和 5 年度「全国学力・学習状況調査」の結果を見ると、パフォーマンステストに関連する「書くこと」は 5 問で 24.1%、「話すこと」[やり取り] は 4 問で 14.5%、[発表] は 1 問で 4.2%と、驚くほど低い正答率であった。「書くこと」で実際に英文を書く問題では、大問 8 の (2) が「社会的な話題に関して読んだことについて、考えとその理由を書く」問題で正答率が 20.1%（無解答 28.9%）、大問 10 の「日常的な話題について、事実や自分の考えなどを整理し、まとまりのある文章を書く」問題で正答率が 7.7%（無解答 20.9%）であった。「話すこと」は大問 1 が [やり取り]

で、即興で会話を続ける問題が4問あり、正答率は順に19.0%（無解答22.7%）、9.4%（無解答18.1%）、13.4%（無解答19.4%）、16.1%（無解答17.8%）で、大問2が「社会的な話題に関して聞いたことについて、考えとその理由を話す」問題で正答率が4.2%（無解答18.8%）であった。

問題の難易度や課題設定の仕方、採点基準など、様々な要因が絡んでいるが、ここでは議論しない。ただ、令和4年度「英語教育実施状況調査」では、言語活動やパフォーマンステストは十分実施されているという結果が出ているので、この矛盾をどう処理すべきなのか、大きな課題として残る。

今後はこの課題解決のために、新たな施策が講じられることになるのだろうが、個人的には、もう教育現場に負担をかけることはやめてもらいたい。デジタル化による効率化や教科書、指導内容のスリム化に行政が率先して取り組んでもらいたい。令和5年度「全国学力・学習状況調査」の結果を、現場の教員はどのように捉えているのだろうか。授業実践の結果が文部科学省の設定する基準まで届いていない現状は、教師には辛すぎる現実である。