



Beyond the Final Exam: The Impact of Formative Assessment on Performance, Motivation and Anxiety in University L2 Writing

Emilia D'Alessio

University of Salerno, Italy

Abstract

Assessment practices can significantly influence both achievement and students' engagement with second-language writing. This study examines the impact of a combined assessment approach, incorporating continuous formative assessment alongside final summative assessment, compared with a summative-only model in an Italian higher-education context. Quantitative achievement data were analysed across three academic years, with equalised samples of 47 students per assessment condition for each year. Students assessed through summative-only procedures achieved a mean GPA of 19.16/30, compared with 23.74/30 for students experiencing combined assessment. Pass rates were also consistently higher in the combined-assessment groups. Survey data from 34 students in the combined-assessment condition explored learners' perceptions of feedback, motivation and assessment anxiety. A strong positive relationship was found between feedback and motivation ($r = .938$), while feedback was strongly positively associated with the anxiety-related composite score ($r = .970$). However, the reported indirect effect through the anxiety-related composite was not statistically established. These findings support the value of formative assessment and effective feedback in promoting learning (Black & Wiliam, 1998; Hattie & Timperley, 2007) and suggest that integrating regular formative opportunities with summative assessment may improve writing performance while supporting motivation and potentially helping students manage perceived assessment anxiety; the latter possibility requires further research. The study highlights assessment design as an integral component of university-level L2 writing pedagogy.

Keywords: *formative assessment; L2 writing; feedback; motivation; examination anxiety; higher education*

1. Introduction

In university second-language (L2) writing, a single end-of-course examination offers a summative measure of performance but little opportunity to use feedback before grades are awarded. Formative assessment, by contrast, creates opportunities to practise, interpret criteria and revise approaches while learning is still in progress. The distinction between assessment of learning and assessment for learning is central to the case for combining these purposes [1, 2]. Within Italian higher education, where final examinations remain influential, such a combined approach may be especially relevant.

The present study asks whether students experiencing formative assessments with feedback alongside a final writing examination achieve different outcomes from those assessed only through a final examination. It also examines how perceived feedback relates to motivation and examination anxiety and whether anxiety mediates an association between feedback and writing grades. The paper is based on an earlier dissertation conducted at the University of Salerno and reports its findings with explicit attention to methodological limitations raised during conference review.

2. Background and Theoretical Framework

Formative assessment is effective when evidence of performance is used to support subsequent learning rather than merely to record achievement [1, 3]. Timely, interpretable feedback can clarify standards and enable students to regulate their work, although learners must understand and act on that information for it to influence future performance [2, 4]. These principles are particularly pertinent to writing, where practice and response to comments can occur across successive tasks. Combining formative tasks with a summative examination preserves final certification while giving learners earlier information about the quality of their writing.

Motivation and examination anxiety are related but distinct constructs. Perceptions of competence and the value attached to achievement may shape academic emotions [5], while feedback may reinforce confidence and self-regulation [2, 6]. Nonetheless, a relationship between perceived feedback and



motivation or anxiety does not itself demonstrate that feedback caused psychological change. This distinction is essential in the present non-experimental classroom setting.

3. Methodology

3.1 Setting, Assessment Conditions and Sampling

The dissertation describes undergraduate English-writing assessment in the Faculty of Modern Languages and Literature at the University of Salerno. The combined-assessment condition involved two formative writing assessments, one in each semester, using the summative marking rubric. These assessments contributed no marks to the final grade, but students received formative feedback before the end-of-year written examination. In the comparison condition, assessment relied on the final examination without the same continuous feedback arrangements. Accordingly, the analysed achievement measure was the final writing-examination grade, not an average of formative and summative marks.

The source reports a population of 434 students distributed across three course years (290 in combined assessment and 144 in traditional assessment). For grade comparisons, the study equalised the groups at 47 students per assessment condition in each year, resulting in an analytical sample of 282 students (141 per assessment type). The dissertation describes random sampling within available groups, but does not document random assignment of students to the two assessment conditions. Sampling students from existing conditions should not be confused with randomly allocating the intervention; baseline equivalence and possible selection or instructional differences cannot therefore be assumed.

3.2 Measures and Analysis

Final writing grades were extracted from university records and analysed descriptively by condition and course year. Italian examinations were evaluated on a thirty-point scale; 18/30 was the pass threshold, although failing marks below 18 were retained in the reported grade distributions. Pass/fail rates and chi-square analyses were reported. A voluntary electronic questionnaire was completed by 34 students from the combined-assessment condition, administered approximately five months after their examinations. Items assessed perceived feedback, motivation and examination anxiety on five-point scales; anxiety questions were adapted from the Westside Test Anxiety Scale. Six teachers contributed to a pilot review of questionnaire clarity.

The original grade-level SPSS dataset was reanalysed using a two-way ANOVA with assessment type and course year as factors, together with descriptive GPA summaries and pass/fail counts. The questionnaire analyses (one-sample t-test, Pearson correlations, simple regressions and PROCESS Model 4 mediation) are reported from the original dissertation and recovered outputs. The exact composite scoring and reverse-coding procedure could not be reconstructed from the archived files. Accordingly, *Anxi_Sc* is treated as an anxiety-related questionnaire score rather than a validated measure of objectively observed change in anxiety severity.

4. Results

4.1 Examination Performance

Analysis of the original 282 GPA records confirmed a mean of 19.16 (SD = 5.59; $n = 141$) for summative-only assessment and 23.74 (SD = 3.55; $n = 141$) for combined assessment, a difference of 4.57 grade points using unrounded means. Combined-assessment means were higher in every year (Table 1). A two-way ANOVA of individual grades yielded a significant association with assessment type, $F(1, 276) = 68.93$, $p < .001$, and academic year, $F(2, 276) = 3.75$, $p = .025$, but no significant assessment-type-by-year interaction, $F(2, 276) = 1.88$, $p = .154$. The grade distributions include scores below the passing threshold of 18. The group differences are observational, not causal estimates.

Table 1. Mean GPA by academic year and assessment condition ($n = 47$ per cell).

Course year	Summative mean GPA	Combined mean GPA
Year 1	19.70	24.30
Year 2	17.45	23.32
Year 3	20.34	23.60



Table 2. Two-way ANOVA of GPA (N = 282).

Effect	df	F	p
Assessment type	1, 276	68.93	< .001
Academic year	2, 276	3.75	.025
Assessment × Year	2, 276	1.88	.154

Table 3. Pass rates by academic year and assessment condition (n = 47 per cell).

Course year	Summative only	Combined assessment
Year 1	78.7%	100.0%
Year 2	38.3%	95.7%
Year 3	68.1%	93.6%

Original SPSS cross-tabulations (47 students per year and condition, N = 282) recorded summative-only passes of 37, 18 and 32 (78.7%, 38.3%, 68.1%) for Years 1–3; combined-assessment passes were 47, 45 and 44 (100.0%, 95.7%, 93.6%). Within-condition Pearson chi-square tests of Year × Pass/Fail were significant for summative-only assessment, $\chi^2(2, N = 141) = 17.467$, $p < .001$, but not combined assessment, $\chi^2(2, N = 141) = 2.903$, $p = .234$. The latter test had three of six expected cells (50%) below five (minimum 1.67), limiting the reliability of its asymptotic p-value. Across both conditions pooled, $\chi^2(2, N = 282) = 14.446$, $p < .001$. These tests assess differences across years; they are not direct tests of the between-condition pass-rate difference.

4.2 Motivation and Feedback

Among 34 valid questionnaire respondents, the motivation composite score had a mean of 18.65 (SD = 3.51, observed range 9–25). Against the theoretical scale midpoint of 15, the original one-sample test was $t(33) = 6.07$, $p < .001$, with Cohen's $d = 1.04$. Feedback and motivation scores were strongly positively correlated ($r = .938$, $p < .001$). Their simple regression was statistically significant, $F(1, 32) = 232.404$, $p < .001$, $R^2 = .879$, with slope $B = 0.591$. The adjusted R^2 of .875 describes a model adjusted for sample size and number of predictors, not independent proof of prediction stability. This association is observational, and its magnitude warrants examination of how overlapping questionnaire items entered the composite scores.

4.3 Examination Anxiety and Mediation

The original SPSS correlation matrix reported a positive association between feedback and the anxiety-related score ($r = +.970$, $p < .001$), consistent with the positive regression coefficient ($B = +0.991$), but contrary to the negative correlation sign described in the dissertation narrative. The model was significant, $F(1, 32) = 506.534$, $p < .001$, $R^2 = .941$ (adjusted $R^2 = .939$). The score had $M = 21.94$ (SD = 5.69, range 9–35; $n = 34$). Questionnaire content concerns students' perceptions of anxiety or stress reduction following feedback, but the archived data do not document the exact composite scoring/reverse-coding rule. The positive relationship can therefore be reported for the recorded score; it cannot establish that objectively measured anxiety decreased.

The original dissertation reports a PROCESS Model 4 mediation analysis of 34 respondents. The reported indirect effect of feedback on GPA through Anxi_Sc was $B = 0.923$, with a bootstrap 95% CI $[-0.260, 1.861]$. Because this interval crosses zero, the indirect effect was not statistically established. The reported path coefficients were feedback → Anxi_Sc (a): $B = 0.991$, $p < .001$; Anxi_Sc → GPA controlling for feedback (b): $B = 0.931$, $p = .081$; total effect (c): $B = -0.134$, $p = .320$; and direct effect (c'): $B = -1.057$, $p = .054$. The reported indirect estimate is consistent with $a \times b$ ($0.991 \times 0.931 \approx 0.923$). These values were verified against the dissertation's Table 4.11, but the archived PROCESS procedure was not independently rerun.

5. Discussion

The recovered student-level GPA dataset resolves the earlier discrepancy between pooled and year-specific means: both yield a combined-assessment mean of 23.74 when Year 2 is correctly entered as 23.32. Combined assessment was associated with significantly higher GPA in the two-way ANOVA, with no detectable assessment-by-year interaction. The separate pass/fail tabulations also show a higher descriptive pass percentage under combined assessment, although the within-condition chi-square tests investigate changes by year rather than directly comparing the conditions. Because allocation to



assessment conditions was not random, these associations cannot establish that formative feedback itself caused the score differences.

The strong feedback–motivation association is compatible with accounts of feedback as a resource for self-regulated learning [2, 6]. The strong positive association with Anxi_Sc may reflect greater perceived relief from anxiety among students rating feedback more positively, if higher scores represent greater perceived anxiety reduction. Without the verified score-construction rule, however, this interpretation remains conditional; neither causal effects nor actual before-and-after anxiety change can be inferred. Retrospective self-report and shared measurement method could also inflate associations.

For practice, the findings support further evaluation of a low-cost approach in which educators incorporate two rubric-aligned, feedback-rich writing tasks before the final examination without necessarily changing the weighting of the summative grade. Clear criteria and opportunities to use feedback may be as important as providing comments themselves [2–4]. This is a context-specific recommendation for testing and refinement, not evidence that the model will produce comparable gains across all institutions or learners.

6. Limitations and Future Research

The Scientific Committee highlighted two central limitations: the 34-response questionnaire sample and the absence of random allocation to assessment conditions. The small voluntary survey, drawn from one university department and completed about five months after the examinations, limits generalisability and may introduce non-response and recall bias. Students were sampled from existing assessment conditions rather than randomly allocated, so prior proficiency, instructor practices and other differences could confound the observed grade contrasts. The exceptionally high questionnaire correlations, possible item overlap and unverified scoring direction of the anxiety-related composite also constrain interpretation; high multicollinearity was reported in the exploratory multiple-predictor GPA regression. The reported mediation analysis should be treated as exploratory rather than causal.

Future studies should use prospectively specified measures, documented allocation procedures, baseline language proficiency, independently validated anxiety and motivation instruments, and privacy-preserving linkage between survey measures and subsequent grades. Multi-institution replication, controlled or carefully matched comparisons, and qualitative accounts of how students actually use formative feedback would extend the present work.

7. Conclusion

In this single-institution observational comparison, combined assessment was associated with higher writing grades and pass rates. The reconstructed student-level analysis showed a significant GPA difference by assessment condition but no significant assessment-by-year interaction. Feedback ratings were strongly associated with motivation and an anxiety-related composite among 34 respondents; the reported indirect effect on GPA was not statistically established. Non-random allocation, survey size and measurement limitations warrant cautious interpretation and independent replication.

REFERENCES

- [1] Black P. and William D., “Assessment and Classroom Learning”, *Assessment in Education: Principles, Policy & Practice*, 5(1), 1998, pp. 7–74.
- [2] Nicol D. J. and Macfarlane-Dick D., “Formative assessment and self-regulated learning: A model and seven principles of good feedback practice”, *Studies in Higher Education*, 31(2), 2006, pp. 199–218.
- [3] Sadler D. R., “Formative assessment and the design of instructional systems”, *Instructional Science*, 18, 1989, pp. 119–144.
- [4] Hattie J. and Timperley H., “The Power of Feedback”, *Review of Educational Research*, 77(1), 2007, pp. 81–112.
- [5] Pekrun R., “The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice”, *Educational Psychology Review*, 18, 2006, pp. 315–341.
- [6] Bandura A., “Perceived self-efficacy in cognitive development and functioning”, *Educational Psychologist*, 28(2), 1993, pp. 117–148.
- [7] Gibbs G. and Simpson C., “Conditions under which assessment supports students’ learning”, *Learning and Teaching in Higher Education*, 1, 2004–2005, pp. 3–31.