Integrated reading-into-writing: Developing a task blueprint

Brooke Donnelly

UNSW College

Integrated reading-into-writing assessment has become increasingly common in second language English for Academic Purposes assessment. It is viewed as a means of potentially increasing the authenticity and therefore validity of assessment that aims to measure a candidate's readiness for the demands of academic writing. While there is a body of research seeking to identify the distinct sub-skills involved in an integrated reading to write task, there is much more limited work addressing the design of rubrics for integrated tasks for specific contexts. This paper describes the development of an integrated reading-into-writing task and rubric for students at B2+/C1 CEFR level in a PreMasters Direct Entry Program (DEP) at a university language centre in Australia. The project aimed to develop a blueprint for an integrated reading-into-writing task and rubric that captures the core constructs of synthesis writing and can be used as a basis for assessment writers to adapt to suit the parameters of other specific integrated reading into writing tasks in other programs or contexts.

Key words: integrated assessment, reading-into-writing, rubric development, synthesis writing, English for academic purposes, university entry

Introduction

Academic writing is a complex task that requires the integration of multiple skills, and multiple sources of information often in different modalities. University students are required at both undergraduate and postgraduate levels to produce 'content responsible' texts (Leki & Carson, 1997) that are based on and appropriately integrate information and ideas from academic sources. Much research has shown that in academic settings there are commonly assigned writing tasks that are reading based, require the integration of reading and writing and are crucial for academic success (Horowitz, 1986; Hale et al. 1996; Leki & Carson, 1997; Rosenfeld et al, 2001; Huang, 2010). Indeed, there is increasing evidence that the ability to successfully integrate source information in written texts is an important threshold ability that determines the readiness of an individual for the demands of academic study (Huang, 2010;

Cumming 2013, Sawaki et al, 2013; Cumming 2014; Cumming et al, 2016).

Integrated reading-into-writing assessment has become increasingly common in second language English for academic purposes assessment (Göktürk Sağlam & Tsagari, 2022; O'Grady & Taşkesen, 2022). It is viewed as a means of potentially increasing the authenticity and therefore validity of assessment that aims to measure a candidate's readiness for the demands of academic writing (Cumming, 2013; Gebril & Plakans 2013). Candidates are required to select, connect, organise, transform, and integrate content from sources in an integrated writing task and such tasks are therefore said to reflect the demands of tertiary academic writing because candidates are required to demonstrate both receptive reading skills and productive writing skills in the final product (Cumming, 2013; Yu, 2013). However, reading and writing processes change when completing an integrated writing task and take on new dimensions, making the reading-into-writing construct unique (Wolfesberger, 2013) and therefore presenting a challenge to assessment writers in developing appropriate scoring instruments that capture the unique construct. While there is a body of research seeking to identify the distinct sub-skills involved in an integrated reading-into-writing task (Sawaki & Lee, 2013; Gebril & Plakans, 2013; Gebril & Plakans, 2014; Shin & Ewert, 2015; Plakans et al 2016), there is much more limited work addressing the design of rubrics for integrated tasks for specific contexts (Chan et al., 2015; Janssen et al., 2015; Uludagh & McDonough, 2022b).

This paper describes the development of an integrated reading-into-writing task and rubric for students at B2+/C1 CEFR level in a PreMasters Direct Entry Program (DEP) at a university language centre in Australia. Previous work on quality assurance of outcomes in DEPs across the Australian university sector recognises the value of DEPs in developing fundamental academic language skills beyond a discrete language skills focus, but has identified the need to develop scales for integrated skills assessment for use in cross-institutional settings (Roche & Booth, 2021). This project aimed to develop a blueprint for an integrated reading-into-writing task and rubric that captures the core constructs of synthesis writing and that can be used by assessment writers as a basis from which to adapt in order to suit the parameters of specific integrated reading into writing tasks in other programs or contexts.

Theoretical definition of the assessment construct

Integrated reading-into-writing assessment moves away from the traditional focus in language assessment on discrete skills testing that seeks to measure reading and writing as independent traits (Cumming, 2013; Gebril & Plakans, 2013). Reading and writing are inextricably linked in an integrated task – a task dependency that means that a candidate cannot write effectively if they cannot understand the

input (Cumming, 2014). This requires a redefinition of the writing construct for an integrated task to recognise the necessary interdependence of the two skills in the written product.

Grabe and Zhang (2013) provide a summary of common reading/writing tasks in academic settings that are seen as important by both faculty and students across disciplines, as shown below:

1. Taking notes from text
2. Summarising information from a text
3. Paraphrasing information from text
4. Combining (synthesising) information from multiple sources
5. Producing a critical synthesis by comparing multiple points of view from texts
6. Answering written exam questions
7. Writing a literature review or extended research paper or essay
8. Summarising and critiquing assigned texts

They argue that the process of synthesis, or the combination and interpretation of information from two or more texts to produce a new unique text, is central to the tasks identified above. Indeed, writing and learning from multiple sources is a complex and demanding task that is typical of and commonly assigned in academic settings across disciplines (Hirvela, 2016; Castells et al. 2023).

Synthesis writing, included as mediation in the Common European Framework of Reference for Languages Companion Volume (CEFR CV, see Council of Europe, 2018), involves both reading and writing processes that occur in a non-linear iterative manner. Discourse synthesis research suggests that writers employ a number of processes in synthesis writing — understanding the organisation of the input and selecting a structure for their own text; selecting relevant information from the input to include in the writing; and connecting or integrating information from the input in the writing, as well as integrating information from the input with the writer's own ideas (Spivey & King, 1989; Plakans, 2009; Nelson & King, 2023).

The discourse synthesis framework developed by Spivey and King (1989) has informed research into the processes employed in integrated writing tasks to identify a potential construct. Ascension-Delaney (2008) found that different reading-into-writing tasks elicited different processes that reflected discourse synthesis and were not simply a combination of reading and writing ability. Similarly, Plakans (2009)

found that writers of higher quality syntheses more frequently used the processes of selecting, organising, and connecting than those who produced lower quality syntheses who tended to focus more on language difficulties. Further evidence is provided by Yang and Plakans (2012) who, in an investigation of strategy use in the TOEFL reading-listening-writing task, found that higher performing candidates employed a complex set of inter-related strategies and concluded that the use of discourse synthesis strategies may support comprehension and thus retention for writing. Knoch and Sitajalabhorn (2013) summarise the sub-processes revealed in such a task as:

1. Selecting ideas from the input text(s) for use in writing
2. Connecting (synthesising) ideas from the input texts
3. Transforming or paraphrasing the language used in the source text(s) to present the selected ideas
4. Selecting an organisational structure for the written product (this is often not the same as the input)
5. Connecting the ideas selected from the input texts in the writing; and connecting ideas from the input texts with the writer's own ideas

If it is ultimately desirable to use an integrated reading-into-writing score to make confident predictions about university readiness in both reading and writing, it is necessary to clearly establish how each skill is captured in and interacts in the construct. While reading ability is clearly important in completing an integrated writing task successfully, it is less clear how inferences about reading proficiency may be drawn from an integrated writing task score (Plakans, 2023). The role of reading appears to be related to the quality of source use in the written product. Studies that have investigated source use in integrated tasks have found that comprehension of source texts may influence how they are used. Cumming et al. (2005, 2006) found that writers with higher proficiency summarised more and showed better integration of sources in their writing than mid-level proficiency writers who relied more on paraphrasing and copying from sources, while low proficiency writers used sources the least. Similarly, Shin and Ewert (2015) found that overreliance on sources in an integrated task, or conversely underuse, correlated with lower reading ability. Gebril and Plakans (2013) also found that low proficiency writers relied heavily on direct copying of words and phrases from the source texts. They suggest that scoring should focus on different features at low and mid to high proficiency levels with a greater focus on the quality of source integration at higher levels.

The ability to appropriately select, organise, and integrate source information appears

to be a consistent predictor of high and low performance on integrated reading-into-writing tasks, and to be related to reading ability. In a study of the TOEFL integrated reading-listening-writing, Sawaki et al. (2013) found that writing was structurally related to reading and listening input and identified a higher order comprehension factor that distinguished between high and low performers. They concluded that candidates performing below the university entry threshold were experiencing difficulties comprehending the input texts and selecting and organising information from the texts. Castells et al. (2023) also found that reading comprehension scores are significantly correlated with text organisation in the written product in an integrated task, as well as the accuracy of representation of source ideas in the written text and the relevance of the ideas included from the sources. In summary, the research described here suggests that lower proficiency readers may struggle with the reading demands in integrated tasks and that this inhibits how they select, organise, and transform information from sources in their writing. It would appear that the quality of integration and transformation of source information, as well as the accuracy and relevance of the source information in the written product are important to the construct of integrated reading-into-writing, and therefore it is a recommended task type for higher proficiency learners (e.g., CEFR B2 and above).

Research has also revealed expected differences between lower and higher proficiency writers in the written output of an integrated task. Fluency, or length, has been a consistent predictor of performance across much research (Cumming et al. 2005, 2006; Gebril & Plakans, 2009; Plakans et al, 2016) with higher proficiency writers consistently producing longer texts. Gebril and Plakans (2013) found that 35% of score variance on an integrated reading to writing task could be attributed to fluency and accuracy, although grammatical accuracy only distinguished low from mid to high proficiency levels but did not distinguish between mid and high proficiency writers. Plakans et al. (2016) similarly showed that fluency, accuracy, and complexity explained 46% of score variance in a study of the TOEFL integrated reading-listening-writing test, though complexity had the least predictive value, while Sawaki et al. (2013) also found that sentence conventions (accuracy) and productive vocabulary significantly distinguished high from low performers on the TOEFL integrated task. It appears that while accuracy may be a feature that distinguishes lower proficiency writers, at higher levels where language-related issues appear less frequently and to distinguish between mid to high proficiency writers, text organisation and development may play a greater role. In a study investigating rater processes when scoring an integrated reading into writing task, Gebril and Plakans (2014) found that raters paid more attention to language-related features at lower levels, to text organisation at mid proficiency, and to development and effective integration of

sources with high proficiency writers. This finding supports the research reviewed above which suggests that the quality of organisation and integration of source material is a feature that distinguishes high proficiency performance on an integrated reading-into-writing task.

In summary, and on the basis of the literature reviewed above, the following features will form the basis of rubric development:

1. Relevance of chosen source information
2. Length (under length for low proficiency)
3. Under/over reliance on source information in the written text (low proficiency)
4. Integration of source information in the written text (logic of connections made between source information and between source information and the writer's 'voice')
5. Transformation (high proficiency: summarisation of source material; mid to high proficiency: paraphrasing; low proficiency: direct copying)
6. Organisation of the written text (coherence and cohesion)
7. Grammatical/lexical accuracy
8. Grammatical/lexical complexity (subsumes range)

The construct defined here forms the basis for rubric development for an integrated reading-into-writing task developed for an English language DEP at an Australian university language centre.

Research context

The university centre is at the time of writing in the development phase of a Pre-Masters pathway program. The Pre-Masters allows students whose undergraduate results are marginally below the academic entry requirements for their desired Masters program guaranteed entry into a Masters by coursework program on successful completion of the pathway component. The Pre-Masters is to be offered in three discipline streams – Engineering; Business; Art, Design and Architecture – with students completing a mix of discipline-specific courses for credit that contribute to their Master's degree. Students who do not meet the university's English language entry requirements with an overall score gap of no more than 0.5 band score IELTS equivalent, are also required to complete an additional 120-hour Academic English for Higher Studies unit. The English language entry requirement for the English supported option of the program is 6.0 IELTS equivalent overall for courses requiring a 6.5 for entry, or 6.5 IELTS equivalent overall for courses requiring a 7.0 for entry.

Curriculum and assessment rationale

Short answer synthesis task

The integrated reading-into-writing assessment described in this paper is the first of two linked integrated writing assessments in the Academic English for Higher Studies curriculum. The curriculum takes students through a guided research project based on a discipline-specific problem-solution scenario. The course first focuses on explanation and analysis of the given problem and the language and discourse features of causality and depth of explanation. In the latter part of the course, the problem analysis is extended to evaluation and argumentation with a focus on extending the problem analysis to the proposal and evaluation of one or more strategies or solutions for managing the identified issue.

The short answer synthesis task (SAST) is the integrated writing task administered at mid-course and the focus of this project. It is a timed, supervised task requiring students to produce a short answer of approximately 300 words, synthesising information from three texts. The focus is on explanation — causes and/or consequences of the issue identified as the research topic — and the task is informative in that an adequate performance simply requires the integration and summary of information from the given source texts in a neutral manner without evaluation or argumentation. The paragraph level focus of the task is intended to act as a scaffold for the final summative writing task, the critical analysis essay (CAE). This is a 600 – 700-word, timed, supervised essay which requires the more complex process of synthesis-based argumentation, requiring the student to propose and evaluate one or more solutions to the problem analysed in the SAST.

This design decision has been made on the basis of research which shows that producing a synthesis-based argument requires higher-level critical thinking processes in constructing the task representation than summary tasks (Ascención Delaney, 2008), and that producing an argument synthesis appears to require more complex and more frequent patterns of referring to sources than producing an informative synthesis (Vandermeulen at al., 2020). The SAST and CAE aim to measure a student's ability to successfully perform synthesis writing in an academic context as a measure of preparedness for university study. They do this in different genres and at an increasing level of complexity. The rubric developed for the SAST and described here can form the basis of the rubric for the CAE with adaptations for the additional argumentation and different genre of the final product.

Reading guides

To scaffold the integrated reading into writing tasks in the Pre-Masters curriculum, students receive a Source Bank consisting of two academic articles related to the

designated research topic. This Source Bank is the starting point for a series of integrated written and spoken tasks that draw on these and additional texts. Multiple studies have underscored the need for students to be supported with explicit attention to reading comprehension with the texts that they will use in reading to writing tasks (Grabe & Zhang, 2013; Grabe & Zhang, 2016; Hirvela, 2016) and to this end, students undertake guided reading of an extended extract from each of the full Source Bank texts supported by a structured reading guide and the teacher. The Reading Guide consists of a series of activities as suggested by Grabe and Zhang (2013) to support students with preparation for reading and writing tasks including activities to support understanding of text structure, key points, main arguments and key vocabulary, as well as identifying connections or contradictions between the information or arguments made in the texts.

Input text specifications

Input texts in integrated reading-into-writing tasks should contain a significant proportion of language, otherwise the input task is likely be merely a source of ideas or content that will not trigger meaningful integration of skills or transformation of input language (Knoch & Sitajalabhorn, 2013). It was thus decided that in order to maintain the authenticity of an academic reading-to writing task with language-rich input, each input text for the SAST should be longer than the required written answer of 300 words but cognitively manageable within the constraints of a timed writing task.

Three input texts with a combined total of 1700 – 1800 words are provided in the SAST:

- Sources 1 & 2: a 600 – 650-word extract from each of the Source Bank texts (from the section selected for Reading Guide)
- Source 3: a 500 – 600-word extract from a third unseen text
- source content should address causes of, reasons for, and impacts or consequences of the designated research topic or issue in line with the problem explanation focus of the SAST

The Source Bank texts may be drawn from authentic academic journal articles, textbooks, or conference proceedings. The extract selected from each text for detailed reading with the Reading Guide should not exceed 2000 words. This length is manageable within the constraints of the course and lesson time available but provides an extended reading experience for the students to develop some understanding of the topic, while also providing sufficient length for test writers to select more than one 600 – 650-word extract for variation in test versions. These 600 – 650-word extracts chosen for use in the SAST are ungraded on the basis that

students have previously been exposed to these texts with support from the Reading Guide and their teacher. Source 3, in contrast, is unseen prior to the assessment task and should be accessible to the students under the time constraints of the assessment. It should also reflect the desired B2+/C1 proficiency level required for university admission (Sawaki et al, 2013; Gebril & Plakans, 2013). The 500 – 600-word Source 3 extract should be drawn from an additional authentic academic text and should be modified to C1+ level with a readability score of between 40-50 on the Flesch Reading Ease Index (equivalent to easier end of university level range). The Text Inspector (Weblingua, n.d.) is used in our context, though other tools are available (see Appendix A for full text and prompt specifications).

Task pilot

A pilot task based on the specifications outlined in Appendix A was administered to a class of 10 students at the centre completing a 13-week DEP with an English language requirement. The students had achieved IELTS 6.5 for entry and required IELTS equivalent 7.0 on exit. They were highly motivated to participate in additional writing development activities.

The students were provided with two texts on the topic of intensive animal farming, each of them around 2000 words in length, and each accompanied by a structured reading guide. Students were given one week to read the first text and complete the activities in the Reading Guide before participating in a teacher facilitated workshop to review the texts and participate in structured discussion of the topic. This process was repeated with the second text with the second workshop considering explicit connections between the information provided in the two texts. One week after the second workshop, the students completed the pilot task with an extract of 600 words from each of the texts and an additional unseen graded extract of 500 words. The time allowed for the pilot task was 80 minutes based on timing allowed in common proficiency tests for writing and reading tasks of similar length. However, on observation that 40% of the responses were marginally under length, this has been increased to 90 minutes for the next pilot.

Rubric development

Rationale for rubric development

The construct identified in the literature review for this project is multifaceted, consisting of a number of discrete sub-skills and it is important for the rubric to represent these underlying constructs as separate, measurable skills (Knoch & Sitajalabhorn, 2013). An analytic rubric was developed for the task as it provides the ability to capture the identified constructs separately and it also provides a useful means for students to receive detailed formative feedback on the identified aspects of

their performance (Uludag & McDonough, 2022a). This is an important consideration given that students will need to undertake the same synthesis processes in the more complex written argument of the critical analysis essay at the end of the course. It is also a rubric format that is commonly used at the centre and with which teachers are familiar.

The pilot was administered prior to the development of the scoring instrument on the basis of literature which suggests that the use of actual student writing samples to understand student performance levels in the different scoring domains is central to an evidence-based approach to the design of scoring instruments (Ewert & Shin, 2015). Similarly, the elicitation of rater perceptions of scoring criteria in relation to student work also plays an important role in the standardisation of scoring instruments — this kind of qualitative feedback and consideration of teacher challenges and priorities when applying evaluation criteria can help to more explicitly link assessment and classroom practice in contexts where teachers are also assessors (Uludag & McDonough, 2022b). Rubric development therefore utilised an iterative process based on evaluation of the pilot samples and rater feedback and input (North & Docherty, 2016; O'Grady & Taşkesen, 2022).

Process for rubric development

In the initial phase, a group of four Pre-Masters curriculum writers were consulted to give feedback on the pilot samples in relation to the identified task construct. All have 12 – 18 years' experience in teaching academic English in tertiary contexts and were selected for their knowledge of the course content and the alignment this would provide with assessment outcomes. The group considered the input and then ranked the writing samples in order from those they felt achieved a superior task outcome to those that performed more poorly. They were then provided with the list of identified constructs and asked to comment on each of the samples in relation to those criteria.

Based on this feedback, six rating criteria were developed from the eight construct areas as shown in Table 1. Relevance and length were combined as under length samples tended to omit key points included in the input. Under or over reliance on the source information was felt to reflect a poorer ability to integrate the source information in the written product and therefore was combined with integration of the source information in the written text. This result reflected the findings in the literature reviewed that under or overreliance on source information is a product of poorer reading proficiency. Finally, grammar and vocabulary were separated to give a more accurate representation of the performance in each domain.

Table 1
Development of Rubric Criteria from Identified Constructs

	Construct	Criteria
1.	Relevance of chosen source information	Content and relevance
2.	Length (under length for low proficiency)	
3.	Under/over reliance on source information in the written text (low proficiency)	Synthesis
4.	Integration of source information in the written text (logic of connections made between source information and between source information and the writer's 'voice')	
5.	Transformation (high proficiency: summarisation of source material; mid to high proficiency: paraphrasing; low proficiency: direct copying)	Transformation
6.	Organisation of the written text (coherence & cohesion)	Organisation
7.	Grammatical/lexical accuracy	Grammar
8.	Grammatical/lexical complexity	Vocabulary

The second phase of rubric development involved mapping the criteria to existing public performance criteria in order to develop descriptors for performance levels on the existing institutional scale for reporting results. The Content and Relevance, Synthesis, and Transformation criteria explicitly involve reading ability and the selection and use of information from the input sources and thus were mapped to CEFR mediation strategies (Council of Europe, 2018) as shown in Table 2. The CEFR mediation strategies reflect a move away from discrete description of performance in the four skills and a recognition that language use often involves reception and production simultaneously (Piccardo et al., 2019).

Table 2
Mapping Rubric Criteria to Public Performance Criteria

Criterion	Mapping
Content and relevance	CEFR Relaying specific information
Synthesis	CEFR Processing text in writing
Transformation	CEFR Adapting language
Organisation	IELTS Coherence and cohesion
Grammar	IELTS Grammatical range and accuracy
Vocabulary	IELTS Lexical resources

CEFR 'Relaying specific information' is seen as a purely informational task and it is suggested that by B2+, a user should be able to "reliably relay . . . information from . . . particular sections of long, complex texts (Council of Europe, 2018;

p. 93). Higher end rubric performance descriptors for Content and Relevance were thus designed to reflect inclusion of all key points from the input material, and lower-level descriptors to reflect omission of key points or inclusion of irrelevant material.

For Synthesis, the literature reviewed above suggests that higher level performance on reading-to-write integrated tasks tends to include more summarisation of source material and higher quality integration of source material in the written product. Similarly, CEFR 'Processing text in writing' recognises summarisation and interpretation of relevant source material at C1+; comparison, contrast, and synthesis of information from source material at B2+; and simpler summarisation of main points from source material at B2. Descriptors for the Synthesis criterion were therefore developed to reflect these aspects of performance, incorporating under or over reliance on the source material at lower proficiency as outlined previously.

The CEFR 'Adapting Language' descriptors used as the basis for development of the Transformation criterion also reflect the literature reviewed with the ability to paraphrase complex information in simpler language reflected at B2 and the ability to paraphrase, adapt, and interpret complex information reflected at C1. These aspects of performance were therefore incorporated into descriptors for the Transformation criterion with less flexible paraphrase and direct copying incorporated at lower proficiency in line with the literature reviewed.

Finally, the Organisation, Grammar, and Vocabulary criteria were felt by the original group of consulting teachers to more directly reflect aspects of the students' writing proficiency and therefore were mapped to key aspects of the IELTS writing descriptors of Coherence & Cohesion, Grammatical Range & Accuracy, and Lexical Resources (IELTS, 2023a) as these are public descriptors which provide a common reference point with which teachers and other stakeholders are more familiar. The selected aspects of these public IELTS descriptors were aligned with the CEFR levels using the public IELTS and CEFR concordance published by the IELTS organisation (IELTS, 2023b).

However, as the rubric incorporates more granular performance levels than the IELTS descriptors and as the CEFR mediation strategies include only broad descriptions of performance at B2, B2+, and C1 levels, it was necessary to develop more granular performance descriptions at the performance levels required of the institutional rating scale. In order to develop these more granular descriptions, the final phase of rubric development employed an iterative multi-cycle process based on the pilot samples and expanding circles of rater feedback and input (Harsch & Martin 2012; O'Grady & Taşkesen, 2022). A group of five senior teachers first rated the pilot samples using the draft rubric. Feedback from discussion of this process and the challenges experienced was used to adapt the draft rubric descriptors. The process was repeated

twice more with different groups of classroom teachers to ensure input was received at all levels from course developers to classroom instructors.

Discussion & Conclusion

The development of the scoring instrument for this integrated assessment task has been limited by the availability of only a small number of student samples for rubric development and its initial pilot. Though these samples encompassed a range of performance levels, the description of performance levels will obviously be more robust with further review and development based on an expanded set of samples and rater feedback. Further validation work also needs to be undertaken to provide more robust evidence that the scoring domains represent the underlying constructs (Knoch & Sitajalabhorn, 2013) and it is intended that this work and further revision of the rubric will be undertaken with a second pilot group of around 160 students which is currently underway.

However, the approach employed in this project for task and rubric development based on actual samples of student writing and feedback (Ewert & Shin, 2015), mapping to public proficiency descriptors, and iterative feedback and input (North & Docherty, 2016; O'Grady & Taşkesen, 2022) at all levels from course developers to classroom instructors has produced a set of evaluation criteria that we are confident not only reflect performance aligned with institutionally recognised tests of proficiency, but also reflect important instructional aspects at the classroom level. The process has enabled appropriate external proficiency standards for integrated reading-into-writing, via alignment with the CEFR, to be incorporated into the centre's assessment design and delivery (Roche & Booth, 2021). The evaluation criteria are also valuable from a pedagogical perspective, having been shaped by the teacher assessors and their understanding of the construct and aspects of performance. This provides an important bridge between assessment and classroom practice (Uludag & McDonough, 2022) and enables teachers to more confidently incorporate essential aspects of successful performance in the integrated reading-into-writing task in their teaching and learning activities (c.f. Harsch & Martin 2012; O'Grady & Taşkesen, 2022).

The short answer paragraph focus of the task, intended to capture the essential elements of synthesis writing without a specific genre focus, can also provide a basis from which rubrics can be developed for other more complex tasks in different contexts. The designed adaptability of the base scoring instrument may assist with the development of an integrated reading-into-writing scale for use in cross-institutional settings (Roche & Booth, 2021). Indeed, it is hoped that, in line with the remit of the integrated assessment project, assessment writers in other contexts

with student cohorts at the B2+/C1 level may be able to use the rubric as the basis for development of rubrics for tasks in their own contexts.

References

Ascención Delaney, Y. (2008). Investigating the reading-to-write construct. *Journal of English for Academic Purposes, 7*(3), 140–150. https://doi.org/10.1016/j.jeap.2008.04.001

Castells, N., Minguela, M. & Nadal, E. (2023). Writing a synthesis versus reading: Strategies involved and impact on comprehension. *Reading and Writing, 36,* 849–880. https://doi.org/10.1007/s11145-022-10341-y

Chan, S., Inoue, C., & Taylor, L. (2015). Developing rubrics to assess the reading-into-writing skills: A case study. *Assessing Writing, 26,* 20-37. https://doi.org/10.1016/j.asw.2015.07.004

Council of Europe (2018). *Common European framework of reference for languages: Learning, teaching, assessment companion volume with new descriptors.* Council of Europe.

Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly, 10*(1), 1-8. https://doi.org/10.1080/15434303.2011.622016

Cumming, A. (2014). Assessing Integrated Skills. In *The Companion to Language Assessment,* A.J. Kunnan (Ed.). John Wiley & Sons. https://doi.org/10.1002/9781118411360.wbcla131

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing, 10*(1), 5-43. https://doi.org/10.1016/j.asw.2005.02.001

Cumming, A., Kantor, R., Baba, K., Eouanzoui, K., Erdosy, U., & James, M. (2006). Analysis of discourse features and verification of scoring levels for independent and integrated tasks for the new TOEFL. *TOEFL Monograph No. MS-30 Rm 05-13.* ETS.

Cumming, A., Lai, C. & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes, 23,* 47-58. https://doi.org/10.1016/j.jeap.2016.06.002

Ewert, D., & Shin, S. Y. (2015). Examining instructors' conceptualizations and challenges in designing a data-driven rating scale for a reading-to-write task. *Assessing Writing, 26*, 38-50. https://doi.org/10.1016/j.asw.2015.06.001

Gebril, A., & Plakans, L. (2009). Investigating source use, discourse features, and process in integrated writing tests. *Spaan Fellow Working Papers in Second or Foreign Language Assessment, 7*(1), 47-84. Ann Arbor.

Gebril, A., & Plakans, L. (2013). Toward a transparent construct of reading-to-write tasks: The interface between discourse features and proficiency. *Language Assessment Quarterly, 10*(1), 9-27. https://doi.org/10.1080/15434303.2011.642040

Gebril, A. & Plakans, L. (2014). Assembling validity evidence for assessing academic writing: Rater reactions to integrated tasks. *Assessing Writing, 21,* 56-73. https://doi.org/10.1016/j.asw.2014.03.002

Göktürk Saglam, A. L., & Tsagari, D. (2022). Evaluating perceptions towards the consequential validity of integrated language proficiency assessment. *Languages, 7*(1), 65. https://doi.org/10.3390/languages7010065

Grabe, W. & Zhang, C. (2013). Reading and writing together: A critical component of English for academic purposes teaching and learning. *TESOL Journal, 4*(1), 9-24. https://doi.org/10.1002/tesj.65

Grabe, W., & Zhang, C. (2016). Reading-writing relationships in first and second language academic literacy development. *Language Teaching, 49*(3), 339-355. https://doi.org/10.1017/S0261444816000082

Harsch, C., & Martin, G. (2012). Adapting CEF-descriptors for rating purposes: Validation by a combined rater training and scale revision approach. *Assessing Writing, 17*(4), 228-250 https://doi.org/10.1016/j.asw.2012.06.003

Hirvela, A. (2016). *Connecting reading & writing in second language writing instruction.* University of Michigan Press.

Hale, G., Taylor, C., Bridgeman, B., Carson, J., Kroll, B., & Kantor, R. (1996). A study of writing tasks assigned in academic degree programs. *TOEFL Research Report 54*. ETS.

Horowitz, D. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly, 20,* 445–462. https://doi.org/10.2307/3586294

Huang, L. (2010). Seeing eye to eye? The academic writing needs of graduate and undergraduate students from students' and instructors' perspectives. *Language Teaching Research, 14*(4), 517-539. https://doi.org/10.1177/1362168810375

IELTS (2023a). *IELTS writing band descriptors.* https://s3.eu-west-2.amazonaws.com/ielts-web-static/production/Guides/ielts-writing-band-descriptors.pdf

IELTS (2023b). *IELTS and the CEFR.* https://ielts.org/organisations/ielts-for-organisations/compare-ielts/ielts-and-the-cefr

Janssen, G., Meier, V., & Trace, J. (2015). Building a better rubric: Mixed methods rubric revision. *Assessing Writing, 26,* 51-66. https://doi.org/10.1016/j.asw.2015.07.002

Knoch, K. & Sitajalabhorn, W. (2013). A closer look at integrated writing tasks: Towards a more focussed definition for assessment purposes. *Assessing Writing, 18*(4), 300-308. https://doi.org/10.1016/j.asw.2013.09.003

Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly, 31,* 39-69. https://doi.org/10.2307/3587974

Nelson, N., & King, J.R. (2023). Discourse synthesis: Textual transformations in writing from sources. *Reading and Writing, 36,* 769–808. https://doi.org/10.1007/s11145-021-10243-5

North, B., & Docherty, C. (2016). Validating a set of CEFR illustrative descriptors for mediation. *Cambridge English: Research Notes, (63),* 24-33. https://www.cambridgeenglish.org/images/301971-research-notes-63.pdf

O'Grady, S., & Taşkesen, Ö. (2022). Developing a rating scale for integrated assessment of reading-into-writing skills. *Language Learning in Higher Education, 12*(1), 159-183. https://doi.org/10.1515/cercles-2022-2049

Piccardo, E., North, B., & Goodier, T. (2019). Broadening the scope of language education: Mediation, plurilingualism, and collaborative learning — The CEFR companion volume. *Journal of e-Learning and Knowledge Society, 15*(1). https://doi.org/10.20368/1971-8829/1612

Plakans, L. (2009). Discourse synthesis in integrated second language assessment. *Language Testing, 26*(4), 561–587. https://doi.org/10.1177/0265532209340192

Plakans, L. (2023). Assessment of Integrated Skills. In The Encyclopedia of Applied Linguistics, C.A. Chapelle (Ed.). https://doi.org/10.1002/9781405198431.wbeal0046.pub2

Plakans, L., Gebril, A., & Bilki, Z. (2016). Shaping a score: Complexity, accuracy and fluency in integrated writing performances. *Language Testing, 36*(2), 161 – 179. https://doi.org/10.1177/0265532216669537

Roche, T., & Booth, S. (2021). A collaborative approach to assuring standards: Using the CEFR to benchmark university pathway programs' English language outcomes. *Language Teaching Research Quarterly, 26,* 18-38. https://doi.org/10.32038/ltrq.2021.26.02

Rosenfeld, M., Leung, S., & Oltman, P. (2001). The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels. *TOEFL Monograph Series MS-21.* ETS. https://www.ets.org/research/policy_research_reports/publications/report/2001/ibxd.html

Sawaki, Y., Quinlan, T., & Lee, Y. (2013). Understanding learner strengths and weaknesses: Assessing performance on an integrated writing task. *Language Assessment Quarterly, 10*(1), 73–95. https://doi.org/10.1080/15434303.2011.633305

Shin, S. Y., & Ewert, D. (2015). What accounts for integrated reading-to-write task scores? *Language Testing, 32*(2), 259-281. https://doi.org/10.1177/0265532214560257

Spivey, N., & King, J.R. (1989). Readers as writers composing from sources. *Reading Research Quarterly, 24*(1), 7-26.

Uludag, P., & McDonough, K. (2022a). Exploring EAP instructors' evaluation of classroom-based integrated essays. *TESOL Journal, 13*, e670. https://doi.org/10.1002/tesj.670

Uludag, P., & McDonough, K. (2022b). Validating a rubric for assessing integrated writing in an EAP context. *Assessing Writing, 52*. https://doi.org/10.1016/j.asw.2022.100609

Vandermeulen, N., van den Broek, B., Van Steendam, E., & Rijlaarsdam, G. (2020). In search of an effective source use pattern for writing argumentative and informative synthesis texts. *Reading and Writing, 33,* 239-266. https://doi.org/10.1007/s11145-019-09958-3

Wolfersberger, M. (2013). Refining the construct of classroom-based writing-from-readings assessment: The role of task representation. *Language Assessment Quarterly, 10*(1), 49-72. https://doi.org/10.1080/15434303.2012.750661

Yang. H. & Plakans, L. (2012). Second language writers' strategy use and performance on an integrated reading-listening-writing task. *TESOL Quarterly, 46*(1), 80-103. https://doi.org/10.1002/tesq.6

Yu, G. (2013). The use of summarization tasks: Some lexical and conceptual analyses. *Language Assessment Quarterly, 10*(1), 96-109. https://doi.org/10.1080/15434303.2012.750659

DOI: https://doi.org/10.61504/ZKYH1877

Brooke Donnelly is an Education Manager at UNSW College. Brooke has managed English pathway programs for tertiary education for over 10 years and developed a range of blended EAP curricula and integrated assessments for these contexts. She is currently also a co-convenor of the English Australia Assessment SIG.

b.donnelly@unswcollege.edu.au

Many thanks to Eve Courtney, Bianka Malecka, and Simon Schubert at UNSW College for their work on the curriculum and assessment design.

Many thanks also to Phil Godber, Curriculum and Assessment Lead at UTS College, for his feedback on the project.

Appendix A

Task Blueprint

Input and output specifications for the integrated short answer task are outlined below. Specific aspects of the input text specifications, prompt specifications, or administration guidelines may be adapted to suit the constraints or parameters of tasks in different contexts.

1. Input materials

1.1 Source Bank reading texts

- Authentic academic texts that may be drawn from research/journal articles, textbook extracts, conference proceedings – no limit on length
- Section(s) selected for analysis with reading guide: maximum 3000 words (to avoid complex data or methodology descriptions, for research/journal articles the selected sections should be drawn from the introduction, background or literature review, or discussion sections)

1.2 SAST assessment input texts

Total input length: 1700 – 1800 words

- Source 1: extract of 600 – 650 words from Source Bank Text 1 covered in the Reading Guide
- Source 2: extract of 600 – 650 words from Source Bank Text 1 covered in the Reading Guide
- Source 3: adapted extract of 500 – 600 words from an additional authentic academic article C1+ level with a readability level of between 40-50 on the Flesch Reading Ease Index (equivalent to easier end of university level range)

1.3 Input text content

- Two of the sources should present positions, arguments, conclusions or information that is complementary and support each other
- One of the sources should present a position, argument, conclusion or information that at least in part contradicts or offers a different interpretation of the position, argument, conclusion or information given in the other two texts.

1.4 Source 1 and Source 2 adaptation:

Both these extracts should not be adapted for level or readability as students have already received comprehension support in class and with the Reading Guides associated with these texts. Extracts may need to have the first or last sentence, some reference words or connectors adapted in order that the extract is logical as a stand-alone text.

1.5 Source 3 adaptation:

Source 3 should meet the following specifications:

a. C1+ level

b. Flesch Reading Ease score between 40-50 (easier end of university level range)

Instructions for adapting Source 3:

1. Run the selected extract through Text Inspector using the institutional account.
2. In the *Analysis Options* under the text entry box for analysis, choose '*Reading'* from the drop-down list, and tick *'Exclude all digits'*. This will provide an accurate score.
3. The analysis should be further refined by choosing 'Use custom known words list' option in *Analysis Options.* Input all the words at C2 and above which the students are expected to know. This may include key vocabulary from the Source Bank texts covered in the Reading Guides. To find which words are C2 and above, go to: http://www.englishprofile.org/wordlists/text-inspector
4. Click 'Analyse' and check the Flesch Reading Ease
5. Then click 'Scorecard' to see the full analysis - texts should be in the C1 level at 50%+.
6. The text metrics should be included on the master text document below the text after the number of words (eg. 670 words 67% C1+ Flesch 44.3).

2. Prompt specifications

2.1 Time allowed

Total time allowed is 80 minutes divided per section as below:

a. 40 minutes – reading, note-taking, and planning

b. 40 minutes – writing

The 40-minute reading and planning time is to allow candidates time to read the unseen Source 3 and to re-read Sources 1 and 2 in light of the question prompt. While some candidates may choose to start writing their response in this time, to encourage note-taking and planning to support synthesis, candidates will not be able to enter a written response until the initial reading and planning time is finished.

2.2 Task instructions

Prompt and task instructions should be set out as below. Insert task specific information where italicised.

(insert assessment number) Short answer synthesis task

Write a **300-word short answer** in response to the question below using relevant information from the sources provided to support your answer.

(insert question prompt)

Instructions

This is a short answer task:

- your answer should be 2 paragraphs
- you do not need to write an introduction and conclusion

You should:

a. identify the information in the sources provided that is most relevant to the question

b. identify connections (similarities and/or differences) between the information in the different sources

c. write your answer and synthesise the information from the sources to support your answer

d. you must refer to all three sources in your answer

e. you must paraphrase/summarise the information you include from the sources in your answer and include in-text references as appropriate

Time allowed (80 minutes)

a. 40-minutes: Reading, note-taking, & planning – you will be provided with paper for note-taking

b. 40 minutes: Writing – type your answer directly into the text box in the online quiz

You will not be able to start typing your answer until the 40 minutes of reading, note-taking, and planning time is finished.

2.3 Standard prompts

A prompt for each test version should be chosen from the standard prompts below to suit the characteristics and content of the extracts chosen for each version of the assessment.

1. What are two reasons for . . . given in the texts provided?
2. What are two consequences of . . . given in the texts provided?
3. What is one cause and one consequence of . . . given in the texts provided?
4. What are the different arguments made about/views given on the problem of . . . in the texts provided?
5. What are the different conclusions drawn about the problem of . . . in the texts provided?

2.4 Written response specifications

Length	• 300 words (+/- 10%)
Format	• short answer 2-3 paragraphs • no introduction or conclusion required
Inclusion of input	• all sources must be included in the written synthesis • answers that do not include reference to all three sources receive a capped grade of 60

Four design measures are in place to ensure written assessment integrity, that the work submitted is the candidate's own response, and to mitigate the possibility that a candidate pre-prepares a response based on Source 1 and Source 2:

1. Source 3 is unseen and must be included in the written response
2. Five possible prompts outlined in section 2.3 above
3. 600 – 650-word extract selected from the full approximate 2000-word length of Source Bank Text 1 and of Source Bank Text 2 – the specific extract is unknown to the candidate
4. Change of research topic with each delivery of the course