

Revising the A2 Key and B1 Preliminary Writing exam

Anthi Panagiotopoulou

Assessment, Cambridge Assessment English

James Lambie

Assessment, Cambridge Assessment English

Kevin Y F Cheung

Research and Thought Leadership, Cambridge Assessment English

CONTENTS

Introduction	3
Key: What has changed and why?	3
Preliminary: what has changed and why?	6
How were the changes received?	9
What is the expected impact on teaching and learning?	11
Conclusion	11
References	11

Introduction

The revised A2 Key, A2 Key for Schools, B1 Preliminary and B1 Preliminary for Schools will be introduced in 2020.¹ The major drivers behind changes to the assessment of writing for these Cambridge English Qualifications were to ensure the Writing test constructs remain fit for purpose, and to better support upward certification, which was introduced in 2011.

In the pre-revision Key exam, writing was assessed as part of the Reading and Writing papers, and performance for these two skills was reported together. Candidates were asked to produce little of their own writing, both in Key and Key for Schools. The most noticeable change to how writing will be assessed post-revision is the reporting of Writing scores separately from scores for reading. For Preliminary, Reading and Writing will be separated into two distinct papers. Key will also report Reading and Writing scores separately, but will continue to combine the tasks in a single paper for practical purposes.

This approach aligns Key and Preliminary more closely with other Cambridge English Qualifications, such as B2 First (Lim 2015), increases the proportion of testing that focuses on writing and provides more information about a candidate's writing performance. The latter outcome is particularly important because providing useful information for learners is at the heart of the renewed Cambridge English focus on Learning Oriented Assessment (LOA).

Key and Preliminary are potentially 'forms of large-scale evidence which could contribute evidence of learning during rather than at the end of the process' (Jones and Saville 2016:79). Candidates taking Key are often near the beginning of their learning journeys and these revisions are an opportunity to put LOA principles into practice with a large-scale exam. By increasing the proportion of testing time spent on writing, we also intended to create positive washback on the teaching and learning of writing skills at these levels.

In preparation for the revision, a series of structural equation modelling (SEM) studies were conducted with exam data from Key and Preliminary to inform the revision process and explore the constructs assessed across the three papers (Listening, Speaking, Reading and Writing). SEM is a measurement technique used to investigate the underlying constructs assessed by an exam and the relationship between skills targeted by different components. These analyses showed that some tasks targeting writing ability in the pre-revision exams, particularly those in Key, were more closely associated with reading than performance on productive writing tasks. For the pre-revision exams, this had no impact on results because Reading and Writing scores were reported together. In order to report these scores separately, as planned for the exams post-revision, we needed to separate the tasks and ensure that those contributing to the Writing score covered the writing ability construct more adequately. In addition, a panel of experts in writing assessment reviewed the writing papers to establish how well tasks were eliciting writing at the target levels. Therefore, some tasks have been discontinued and others have been specified as assessments of reading ability. To ensure that the writing construct was fully represented, new tasks were designed and trialled; one was added to Key and two new tasks are included in Preliminary.

Key: What has changed and why?

The pre-revision Key and Key for Schools included a guided writing task (Part 9), which asked candidates to write a short message of 25 words or more, as a note or an email. Cambridge English assessment specialists reviewed

1. A2 Key and B1 Preliminary will hereafter be referred to as Key and Preliminary, respectively; A2 Key for Schools and B1 Preliminary for Schools will hereafter be referred to as Key for Schools and Preliminary for Schools, respectively.

versions of this task and candidate responses, with reference to the Common European Framework of Reference for Languages (CEFR, Council of Europe 2001). By reviewing responses from live sessions alongside A2 CEFR descriptors, it was established that the task was eliciting responses at the target CEFR level. In particular, candidate submissions were aligned to the following Can Do statements, which are an important part of the writing construct as conceptualised in the CEFR. Therefore, this guided writing task was retained for the exam.

Can write short, simple formulaic notes relating to matters in areas of immediate need.

Can write short, simple notes, emails and text messages (e.g. to send or reply to an invitation, to confirm or change an arrangement).

However, this was the only productive writing task in the pre-revision Key exam and it exclusively focuses on CEFR illustrative descriptors from the writing correspondence scale. The task was also reviewed by the panel of experts alongside ALTE 'Can Do' statements for A1, A2 and B1. This showed that the guided writing task focused on the following A2 statements:

Can write simple letters of the "Thank you" type.

Can convey personal information of a routine nature to, for example, a penfriend.

Can express opinions of the "I don't like" type.

The panel of experts in writing assessment noted that the task did not cover B1 statements for writing and recommended that a new productive writing task be considered as part of the revisions. Assessment specialists responsible for setting pre-revision papers also recommended trying to elicit more writing. For example, when asked to provide feedback, they reported that:

'Continuous writing is not given enough importance. I think there could be a place for a longer writing task that carried more of the overall marks.'

'Replace [Part 8] with an additional writing task, e.g. story or simple essay to stretch the more able candidates.'

'The current [Key] Part 9 is very similar to Part 2 on the current PET [former name for Preliminary] Writing paper. The number of words that candidates have to produce for Part 9 is, at 25–35 words, very low. Since there should be scope for candidates to demonstrate ability at B1 level on the [Key] paper it may be worth considering including another writing task.'

In addition to increasing the variety and amount of writing required of candidates, the panel recommended that a new task should better support upward certification to B1 level.

The other parts specified as writing in the pre-revision Key exam (6–8) were a spelling task, an open cloze task and information transfer tasks. As previously mentioned, SEM analyses indicated that these tasks were assessing skills more associated with reading than with writing. Other analyses also identified further issues to consider as part of the revisions. For example, we investigated how similar-ability candidates from different groups perform on particular tasks. This analysis, known as differential item functioning (DIF), indicated that some groups (particular L1s or age groups) performed differently on the spelling task. Some of these findings may result from different teaching practices across geographical regions, particularly where some countries focus on particular topic areas for learning vocabulary. Although this is not problematic, and there are other possible explanations for this finding, the panel of experts in writing assessment recommended reducing these group differences in the revised exam, if possible.

Teachers and candidates were also surveyed to gather their perspectives on the paper and the tasks; they identified the information transfer task as unpopular, because it was difficult to match to commonly used curricula. The spelling and information transfer tasks were removed from the exam based on the reviews by assessment specialists and the research studies. Also informed by the SEM studies, the open cloze task was re-specified as an assessment of reading ability for inclusion in the revised Key exam. These changes made sufficient space for us to consider another productive writing task.

According to the CEFR, A2-level language learners should be able to write simple connected texts, so new tasks to elicit extended writing were proposed, designed and trialled. One of these tasks was selected for inclusion in the revised exam. The new 'picture story' task requires test takers to write a short 35-word story based on three pictures. This allows the candidate to demonstrate they can perform functions from the creative writing illustrative descriptors of the CEFR (see Table 1), complementing the aspects of writing targeted by the guided writing task.

Table 1: Creative writing descriptors from CEFR A2 and B1 (Council of Europe 2018:76)

B1	Can narrate a story. Can clearly signal chronological sequence in narrative text.
A2	Can tell a simple story (e.g. about events on a holiday or about life in the distant future). Can write very short, basic descriptions of events, past activities and personal experiences.

In most writing tasks at this level (including the guided writing task retained as Part 6), content points are provided in bullet points to scaffold test takers and support them to produce text. This design feature reduces the cognitive load of generating ideas, freeing up resources needed to access lexical and syntactic knowledge (Shaw and Weir 2007). However, the bullet points can make it difficult to assess organisation at lower levels, because test takers inevitably match the presentation order of the bullet points precisely in their written response. An alternative approach to providing content points was adopted for the new writing task, to give test takers the opportunity to construct their own organisational structure. Instead of written bullet points, pictures are provided to help test takers generate ideas. These pictures give test takers something concrete to base their responses on, but they also allow for more creativity to be used in generating ideas and structuring responses. Development of this task was informed by recent work on the assessment of writing in the Cambridge English Young Learners exams (Davies and Dunlop 2018). Following a series of trials, the new task was specified and developed for inclusion in the Writing component. The result is a new Key Writing test that includes two productive writing tasks, covering key functions of writing included in the CEFR (Council of Europe 2018). These additions, and previously discussed changes, meant that there were some alterations to the overall format of the Key Writing paper (see Table 2).

The new picture story task encourages test takers to engage in some simple organising of their responses. Instead of only testing candidates' ability to produce simple transactional messages, we are now testing their ability to produce simple narratives. The task and the accompanying mark scheme enable us to distinguish between candidates who 'can write very short, basic descriptions of events, past activities and personal experiences' (A2) and those who 'can narrate a story' (B1). Trials of the revised Key Writing paper showed the new format successfully achieved this aim.

Table 2: Changes in the Writing paper of the Key and Key for Schools Reading and Writing papers

	<i>Task type</i>	<i>What is required of the candidate</i>	<i>Rationale for decision</i>
Removed	Part 6: Spelling task	Read five definitions and write the words they define – the first letter and number of letters are given.	1. SEM studies indicated that the task did not cover sufficiently the sub-construct of writing ability. 2. DIF studies showed some unexpected group differences in task performance.
	Part 8: Information transfer task	Read two short input texts and complete gaps in output text.	1. SEM studies indicated that the task did not cover sufficiently the sub-construct of writing ability. 2. Survey studies showed that this task was unpopular with teachers and candidates, due to being difficult to match with curricula.
Revised for the Reading paper	Part 7: Open cloze	Completing gaps in text.	SEM studies showed that this task was assessing skills more associated with reading than with writing.
Retained	Part 9: Guided writing (Part 6 in revised Key)	Write a short email or note of 25 words or more.	Analysis of candidate responses by expert reviewers indicated that this task was appropriate for eliciting A2 writing.
Introduced	Picture story (Part 7 in revised Key)	Write a short story of 35 words or more based on three picture prompts.	Introduces a longer productive writing task to Key, allowing candidates to display a wider range of writing skills that cover some CEFR B1 descriptors.

Preliminary: What has changed and why?

Preliminary reading and writing tasks were administered in a single paper pre-revision even though the two skills were reported separately. The Writing paper consisted of a sentence transformation task, a short communicative message task, and an extensive writing task, where candidates were able to choose between writing a story or an informal letter.

In order to explore whether the exam remains fit for purpose, a construct investigation study was carried out to examine whether the underlying constructs of the exam support a componential aspect of language proficiency where each component assesses a distinct aspect of language proficiency (Elliott, Docherty and Benjamin 2015). Exploratory Factor Analysis (EFA) was carried out on the papers to map the task types to the skills measured and then Confirmatory Factor Analysis (CFA) was used to investigate the appropriacy of different plausible construct models.

The analyses showed that the Writing paper would benefit from being revised as the sentence transformation task was tapping more into the reading construct than into the writing one, and the short communicative message task did not clearly measure writing ability. The extensive writing task was found to clearly represent the intended writing construct and the use of assessment criteria allowed for accurate score interpretation. For this reason, it was decided to replace the sentence transformation task and the short communicative message task with one or more extensive writing tasks that would represent the writing construct at B1 more clearly, and which would be assessed using the same assessment criteria as the short communicative message task, thus increasing both the validity and the reliability of the paper.

The email task in Preliminary and Preliminary for Schools is one of the new tasks and has replaced the pre-revision optional informal letter task. It requires the candidate to demonstrate the ability to handle the language of functions, retaining the testing focus of the current short email message but at the same time eliciting a fuller range of functions expected at B1. The candidate is supported through the provision of annotations on the email, which provide the necessary scaffolding for the level. The required output being at 100 words, though, enables the candidate to display a wider range of writing sub-skills, thus allowing for more accurate assessment of candidates' performance and upward certification. Table 3 gives a full summary of the task changes.

Table 3: Changes in the Writing paper of Preliminary and Preliminary for Schools

	<i>Task type</i>	<i>What is required of the candidate</i>	<i>Rationale for decision</i>
Removed	Sentence transformations	To linguistically manipulate part of a sentence to make it semantically equivalent to an input sentence.	Often constrained to one-word answers, which are not part of the B1 writing construct.
	Short communicative message	To produce a short communicative message based on three prompts.	Second longer writing task more suited to level and upward certification.
	Informal letter (optional)	To produce an informal letter based on a written prompt.	Too similar in style to new compulsory email, so retesting the same part of the writing construct.
Retained	Story writing (optional)	To produce a story using a given title or first sentence.	
Introduced	Email response to annotated input (compulsory) Change from 30 words to 100	To respond to an email based on annotations on the input email.	Retains functional focus of current short message, but allows stronger candidates to show B2 level.
	Article (optional)	To produce an article on a topic of personal interest, describing events, feelings and giving opinions.	Allows candidates to display a wider range of writing skills, improving validity of upward certification.

The email genre was chosen to be the compulsory task in Preliminary on the basis that non-native speakers are most likely to have to cope with emails in English in their future lives and most candidates already have genre knowledge, i.e. knowledge of the genre conventions which help candidates shape possible responses (Matsuda and Silva 2010). Having the email task as a compulsory one provides a basis for comparison between candidates and brings the structure of the test in line with the structure of B2 First and B2 First for Schools.

Part 2 of the Writing paper in Preliminary and Preliminary for Schools is an extensive writing task which is not supported by annotations as Part 1 is. Candidates are required to draw on their linguistic resources in order to respond to either an article task, which is new to the paper, or a story task, which is present in the current version of the paper. At this level of writing, candidates are asked to create pieces of extensive writing that demonstrate coherence. This enhances the performance authenticity (Brown and Abeywickrama 2010) and widens the range of lexical resources the candidate can use. This will better enable candidates whose writing ability is above CEFR Level B1 to demonstrate it and be awarded marks accordingly.

Having the email task as a compulsory Part 1 meant that a new task was needed as an option for Part 2 of the paper. Several possible genres were considered, but it was found that the article task was a better match to the Preliminary candidates' experiences and CEFR B1 creative writing descriptors. The candidates are provided with prompts to

guide their writing and they are asked to draw on their own experiences in order to write the article addressing the prompts. They are asked not only to use descriptive language, but also to write accounts of their feelings, reactions and opinions, thus broadening the construct. Creative writing is an important part of the writing construct at B1 level which can allow candidates to show their linguistic range and control.

The story task has remained as an option for Part 2 as the construct investigation analyses (Elliott et al 2015) showed that this specific task not only represented the construct well but also allowed for valid inferences to be made about the candidates' proficiency level.

As Table 4 shows, CEFR descriptors for creative writing apply to both tasks of the Writing paper Part 2 and would thus contribute to the generalisability of reported scores and reliability of the exam, regardless of which task a candidate would opt for.

Table 4: CEFR creative writing descriptors (Council of Europe 2001:62)

B2	<p>Can write clear, detailed descriptions of real or imaginary events and experiences, marking the relationship between ideas in clear connected text, and following established conventions of the genre concerned.</p> <p>Can write clear, detailed descriptions on a variety of subjects related to his/her field of interest.</p> <p>Can write straightforward, detailed descriptions on a range of familiar subjects within his/her field of interest.</p>
B1	<p>Can write accounts of experiences, describing feelings and reactions in simple connected text.</p> <p>Can write a description of an event, a recent trip – real or imagined.</p> <p>Can narrate a story.</p>

Reliability is further increased by the fact that all tasks in the Writing paper are now marked using the same assessment scales that were themselves developed with reference to the CEFR. These scales were in place pre-revision and were used to mark the optional tasks (informal letter and story) and were found to serve the construct appropriately.

Including a compulsory extensive writing task meant that the candidate output as well as the time required for the test would be increased. Taking into consideration the exam's cognitive load as well as the average age of the candidature, the panel recommended that it would be for the candidates' benefit to separate the Reading and Writing paper. The revised Writing paper includes two 100-word tasks for the candidates to complete in 45 minutes and is administered after the Reading paper has been completed. This allows for the candidates to fully focus on the Writing paper and manage their time efficiently, and timing was explored in trials of the test.

The changes at both Key and Preliminary also mean that writing – defined as the production of written text for a communicative purpose (as opposed to notes or single words) – is receiving greater prominence among the four skills. As a result, Writing scores for Key and Key for Schools can now be reported separately from the Reading scores, and on Preliminary and Preliminary for Schools, Writing is now a separate paper. Table 5 summarises how the revision process has changed the Writing component.

Table 5: Increased prominence of productive writing

		<i>No. of parts involving the production of written text</i>	<i>No. of words candidates should write</i>	<i>Recommended time candidates will spend writing</i>	<i>Marks available for communicative writing</i>	<i>Reporting of the Writing marks</i>
Key and Key for Schools	Pre-revision	One (out of nine on the Reading and Writing paper)	25–35	10 mins out of 70	5 (out of 60 for the Reading and Writing paper)	Reading and Writing reported together
	2020 revision	Two (out of seven on the Reading and Writing paper)	60–80	20 mins out of 60	30	Reading and Writing reported separately
Preliminary and Preliminary for Schools	Pre-revision	Two (out of eight on the Reading and Writing paper)	135–145	40 mins out of 90	15 (out of 25 for the Writing paper)	Reading and Writing reported separately
	2020 revision	Two (but Writing is now a separate paper from Reading)	About 200	A separate paper taking 45 mins	40	Reading and Writing reported separately

How were the changes received?

We have seen that the revised papers are intended to better reflect the writing constructs for their level and to increase the prominence of writing at both levels, but now we need to look at whether candidates and teachers agree. Trialling showed that the new Key and Key for Schools picture story task is perceived by candidates as one of the more difficult parts of the paper, while still being manageable for most candidates, and is perceived as more difficult than the email task. The task type of each part of the test is given in Table 6 and Figure 1 breaks down how each part was rated for difficulty by candidates.

Table 6: Format of revised Key and Key for Schools exam

Part 1	3-option multiple choice with messages and signs
Part 2	Multiple-matching reading
Part 3	3-option multiple-choice reading (long text)
Part 4	Multiple-choice cloze
Part 5	Open cloze
Part 6	Guided email writing
Part 7	Picture story

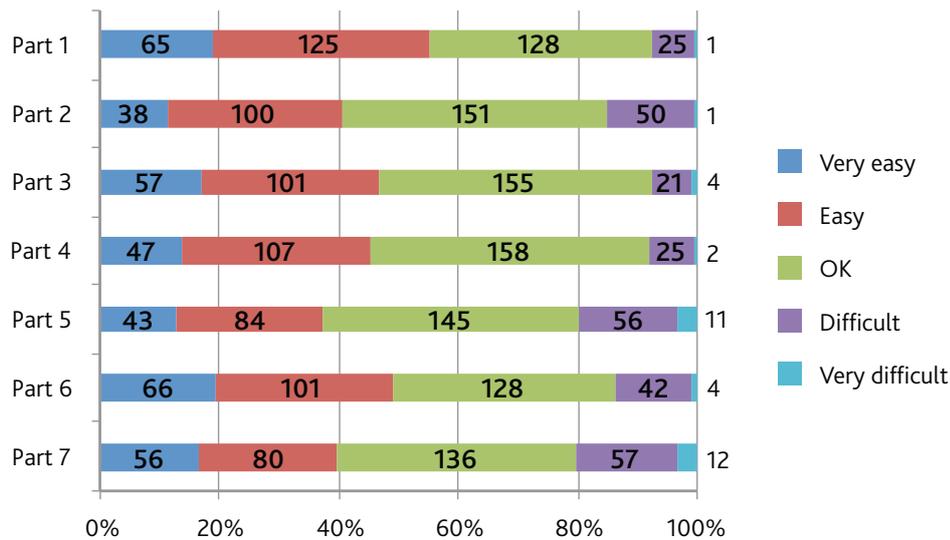


Figure 1: Revised Key and Key for Schools tasks rated for difficulty by candidates

Comments from teachers also suggested that the picture story made the paper more difficult, but again they did not suggest this was inappropriately difficult.

'The writing is longer, however I think it could benefit students.'

'Writing a story implies better handling of tenses and cohesive devices.'

'[What I like about this test is] story telling – it makes the exam more challenging.'

Similarly for B1 Preliminary and B1 Preliminary for Schools, 95% of candidates who took the trial tests agreed that the test allowed them to show their writing ability and that task instructions were clear (see Figure 2). Candidates, teachers, invigilators, stakeholders and focus groups provided feedback on paper difficulty, writing amounts and timing through a variety of channels. Candidates reported that the level of the new tasks was deemed appropriate, the amount of writing was neither too little nor too much, and the proposed timing of the paper at 45 minutes was considered sufficient. Other test users reported that 'the tasks were important and worth students' time to prepare'. Test users also acknowledged the improvements made to the test and reported that the revised paper allowed candidates to better show their language ability (Vidaković and Elliott 2018).

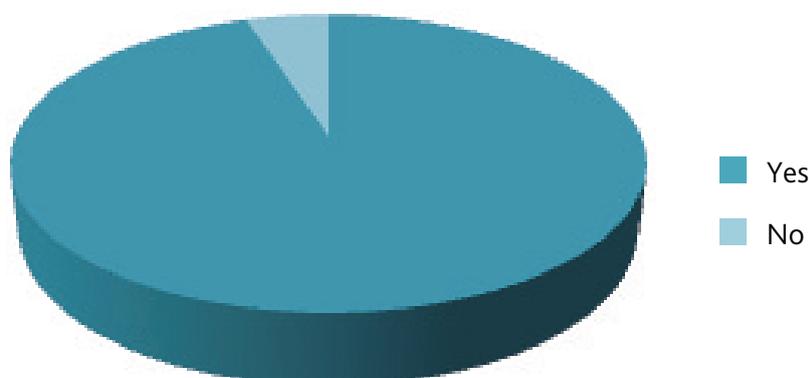


Figure 2: Candidate feedback, Trial 2 – The test allowed me to show my writing ability

As with Key and Key for Schools they also found the second writing task more difficult – 12% responses at 'difficult' or 'very difficult' for Part 1, but 21% responses at 'difficult' or 'very difficult' for Part 2.

What is the expected impact on teaching and learning?

We expect that the changes will have a positive effect on teaching and learning. With productive writing now asking candidates to do more independent writing on Key and Key for Schools, and being made a separate paper on Preliminary and Preliminary for Schools, we expect it to feature more prominently than before in teachers' course and lesson planning. Learners will be required to produce more texts (emails, stories, articles) and will receive more feedback on the aspects of their writing that are tested (lexis, grammar, organisation, task achievement). This should encourage teachers to devote more class time to productive writing, and particularly to practising extended writing. In contrast to short writing exercises, which typically focus on lexis and grammar, extended writing exercises provide important opportunities to raise awareness of other aspects, such as cohesion.

Research is currently underway to establish baseline data for A2 Key for Schools. We will be interviewing teachers to find out how much classroom and homework time is devoted to productive writing, and how they focus on it. Once the revision is introduced we will go back to ascertain what effect the new picture story task and the overall increased prominence of productive writing have on the test as a whole, have had on teachers' approach to course and lesson planning, and to classroom practice.

Conclusion

The revision of Key, Key for Schools, Preliminary and Preliminary for Schools resulted in important changes to the way writing skills are assessed. The key consideration at both A2 and B1 levels was to test more of the writing construct. The driver behind this was ensuring that the test constructs remain fit for purpose, especially in relation to upward certification. This has been achieved through the inclusion of new tasks that target CEFR descriptors which were previously not tested. As a result, candidates are being asked to do more independent writing than on the pre-revision tests, which allows for the reporting of separate scores for Reading and for Writing on Key and Key for Schools, and for Writing to become a separate paper on Preliminary and Preliminary for Schools. This, along with the new task types, brings both papers more into line with other papers from B2 First upwards. It is anticipated that the increased prominence of writing at A2 and B1 will be reflected in a proportionate increase in the amount of study time devoted to the skill, which will provide candidates with a good grounding as they progress towards B2 First.

References

Brown, H D and Abeywickrama, P (2010) *Language Assessment: Principles and Classroom Practices*, White Plains: Pearson Education.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.

Council of Europe (2018) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with New Descriptors*, Strasbourg: Council of Europe.

- Davies, K and Dunlop, M (2018) New writing tasks for young learners, in Cambridge Assessment English (Ed) *The 2018 Pre A1 Starters, A1 Movers and A2 Flyers Revisions*, Cambridge: Cambridge Assessment English, 25–27, available online: www.cambridgeenglish.org/Images/461823-young-learners-revision-publication.pdf
- Elliott, M, Docherty, C and Benjamin, T (2015) *PET(fS) Construct Investigation (SEM)*, Cambridge: Cambridge English internal report.
- Jones, N and Saville, N (2016) *Learning-oriented Assessment: A Systemic Approach*, Studies in Language Testing volume 45, Cambridge: UCLES/Cambridge University Press.
- Lim, G S (2015) Aspects of the revision process for tests of Writing, *Research Notes* 62, 32–35.
- Matsuda, P K and Silva, T (2010) Writing, in Schmitt, N (Ed) *An Introduction to Applied Linguistics*, London: Routledge, 232–246.
- Shaw, S D and Weir, C J (2007) *Examining Writing: Research and Practice in Assessing Second Language Writing*, Studies in Language Testing volume 26, Cambridge: UCLES/Cambridge University Press.
- Vidaković, I and Elliott, M (2018) *Preliminary (fS) Revision Sep 2016 – Across skills*, Cambridge: Cambridge English internal report.