CHAPTER FOUR

Myth 4 You Can't Write That on the Test

Or, Tests Must Regulate Writing

The following passages are from the early to the late twentieth century. See if you can put them in order.

1. An ideal test would be one in which practically everyone could obtain some score and which very few could finish. Then all people would be measured.
2. Experienced and conscientious examiners vary one from another, they cannot all have got the correct and absolute standard.
3. In assessing intelligence (i.e., innate general intellectual ability) teachers are decidedly less reliable than psychological tests.
4. In a climate of growing public interest in public examinations comparability of grading standards is a popular focus of attention.
5. [We] will expect to see far greater comparability in standards between similar examination syllabuses to avoid some papers being seen as "easy".

You guessed right if you thought the first passage was the oldest. Indeed, the passages are already ordered chronologically, and in this way they show a twentieth-century arc of interest in tests. Passages 1 and 2, from the 1920s, illustrate the continuation of myth 3 – the pursuit of a single way to measure and compare people. The first was written in 1923 by Carl Brigham, promoter of pure Nordic peoples and toothy facts. The second was written in 1928 by the Joint Matriculation Board of the universities of Manchester, Leeds, and Liverpool, as they sought an "absolute standard" for test examiners.

Passage 3 reflects where we get by the mid-twentieth century: less trust in teachers; greater trust in standardized exams for measuring ability. The statement appeared in Cyril Burt's 1945 "The Reliability of Teachers' Assessment of their Pupils," at the start of standardized UK exams we will see in the origin story below.

The final two passages reflect public expectations by the late twentieth century. Passage 4 was written in 1978 by the examining boards of the UK General Certificate of Education (GCE) as they faced pressure to publicize exam results. The final passage appeared in 1997, in

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

a press release noting the consolidation of awarding boards in the UK Department for Education and Employment.

All five passages show enduring interest in measuring and comparing *intelligence* through tests, with added concern for standardized ways of doing so. The goal is uniform comparability. The villain is variance, whether in examiners, teachers, or standards.

We've seen similar passages already because this myth builds on all the myths so far. Once we have myths 1, 2, and 3 – and *correct writing* is regulated by schools and narrowly tested as an indication of innate ability – then it is easy to insist that uniform tests must regulate writing.

This myth's origin story begins at the start of written examinations, soon after the start of myth 3.

4.1 Context for the Myth

4.1.1 Exams Begin to Regulate Writing and Students

We know from the last myth that in the early nineteenth century primary and secondary learning was assessed in interactive community events, as was the case in Boston before 1845. Before Horace Mann's unannounced written tests that year, annual interactive events were open to family members, and they focused more on how schools were doing than on comparing individual students.

At universities, learning was also displayed in interactive public events in the eighteenth and early nineteenth centuries. It was uncommon to go to college at the time; essentially, only male, white students from privileged families went.[1] But if you were among them, you would prove your learning by speaking aloud in a classical language, during an interactive period of "verbal jousting." For instance, attending Cambridge University in 1820, you and a peer might debate an author's merits in a public, oral exam in Latin.

By the late nineteenth century, interactive university assessments in Latin were giving way to individual written examinations in English, and school-specific secondary assessments were being replaced by externally-designed written exams. This won't surprise us, given educational shifts we've already seen, away from classical languages in myth 2 and toward Mann's written testing in myth 3. But it entailed a significant change in how learning was evaluated.[2]

4.1.2 Higher Education Expands

Higher education options were expanding around this time. In the UK, Cambridge and Oxford were still socially exclusive and subject to

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

religious tests, and Scottish universities and English dissenting academies began providing alternatives in the 18th and 19th centuries. By 1826, England gained the "godless college of Gower Street," or the University College London, and in the ensuing decades, more colleges were founded, more female students were accepted across social classes, and more leaders called for affordable, accessible institutions for working people.

The same nineteenth-century expansion happened in UK Commonwealth countries and in the US. In Australia, local legislation established the University of Sydney in 1850 and the University of Melbourne in 1853. In Canada, though its first university was established by colonial legislatures in 1789, McGill University and several others followed in the nineteenth century. In the US, the mid- and late-nineteenth-century federal Morrill Acts used Indigenous tribal land, usually obtained through violent seizure or forced cession, as locations and to provide seed money for new public institutions.[3] These new institutions were designed to offer practical training in areas such as agriculture and the mechanical trades. By 1890, private US institutions also began expanding enrollment.

4.1.3 College and Secondary English Writing Exams Begin

With more higher education came more written English exams. In the UK, a local examinations system was established in the 1850s for students leaving secondary school. If you were a male student finishing secondary school near Cambridge in 1858, for instance, you would write a timed essay in English about the queens and children of Henry VIII (in order, mind you) for your history exam. Your response would be evaluated by local examinations evaluators rather than your own teacher. In their evaluation, examiners would look for "correct punctuation, arrangement, spelling, precision, elegance, and completeness."

Students were wholly unused to written examinations like these, as examiner reports make clear. In response, Cambridge 1859 examiners proposed a solution that sounds a lot like today's idea of teaching to the test. "With the stimulus of open competition and the standard of regularly recognised examinations," examiners wrote, "carefulness and ability will receive clearer direction and more open reward."[4]

By the late nineteenth century, students were more accustomed to written English exams, though examiners were not necessarily more content with students' writing. The Cambridge Seniors English Composition Examination 1883 report was cutting:

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

The most usual faults were – statements of utter nonsense, general irrelevance, inexact and pretentious language, carelessness in punctuation and arrangement, and lastly the employment of Scripture texts when the candidate was at a loss for something to say.

Examiners found other written exams lacking as well. Cambridge history examiners lamented students' writing even though they found their historical knowledge satisfactory: "[students'] answers, even when accurate, showed a general uniformity of expression," the examiners reported.

In the US, we know from myth 3, the earliest English entrance exams appeared in 1853, under the direction of Horace Mann at Antioch College. These included an English grammar exam, a history exam, and a geography exam. All three were evaluated for *correct writing*: In a testament to myth 1, Mann saw the study of pure English grammar as a way to purify thoughts.

Harvard wasn't far behind, because by 1869 Charles Eliot was promoting *correct writing* as a way to rank and select students. Beginning in 1872, all of Harvard's student entrance exams were evaluated for "correct spelling, punctuation, and expression" in English. By 1874, Eliot hired one of his former Harvard peers, newspaperman Adams Sherman Hill, to design Harvard's English composition entrance exams. Applicants were furthermore advised that "writing on any entrance exam may be regarded as part of his examination in English."

Late nineteenth-century Harvard reports also documented the rise of English composition exams at other US colleges, including Princeton, the University of Michigan, and the University of Pennsylvania. English writing exams appeared at the US secondary level in the late nineteenth century, including the New York State Regents Examination for graduating secondary students in 1878.

4.1.4 College Exams Emphasize Timed Writing, Literature, and *Correct Writing*

Early English exams emphasized writing under time constraints. Oxford and Cambridge exams lasted around three hours; Harvard's lasted around one hour. Naturally, these exams measured whether students could write quickly, without substantial time for reading or revising.

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

Concerns about timed writing appeared at the time. In 1873, a Cambridge examination student named Amy wrote her parents that she ran out of time during her exams while "A fellow of the university, cap-a-pie, very severe looking, sat at the head of the room, or walked up and down, and frightened me."[5] In 1890, English composition examiner LeBaron Russel Briggs described similar conditions at Harvard. He bemoaned timed exams and rushed exam grading, and he described what today we would call test anxiety: "Again and again I have seen the untrained youth, however cultivated for his years, flinch before every searching test." Another early Harvard examiner, Byron Satterlee Hurlbut, argued in 1892 that students should practice timed writing in class, even though "The more elaborate advanced work must, of course, be done outside the class." This argument illustrates an enduring paradox of timed English writing exams: They are used in high-stakes evaluation, yet not viewed as students' best work.

In addition to being timed, we've seen that early writing exams often focused on literature. Harvard couched their early English entrance exams as responses to "standard authors"; early Oxford, Cambridge, and Harvard exams regularly focused on Shakespearean texts and characters. There were exceptions; an 1858 Cambridge English Composition exam, for instance, asked students to *Write a letter to a friend in Australia, announcing your intention to emigrate*.[6]

Writing knowledge is not the same thing as literary knowledge, and written responses to literature are different than literature itself. As imaginative writing, literature has different purposes and patterns than the writing continuum we are exploring throughout this book. Still, we know from myths 1 and 2 that literature, like *correct writing*, served nationalist goals of standardizing and celebrating English. Harvard president Charles Eliot put it this way: "It is enough to say of the English language that it is the language of English literature."

Finally, early exams were subject to *correct writing* expectations, even on exams in subjects other than English composition. We saw, for instance, that spelling, capitalization, and punctuation were checked across exams at Antioch and Harvard, and that "uniformity of expression" disappointed the Cambridge history examiners despite accurate historical information.

4.1.5 Standardized Exams Begin

The last myth and this one fueled confidence in uniform tests, and it was only a matter of time before local exams gave way to national,

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

standardized exams. At the start of the twentieth century, for example, the US College Board tailored exams to individual colleges. By 1923, it commissioned Carl Brigham (of myth 4 Army Alpha Test fame) to develop a single Scholastic Aptitude Test (SAT). His charge was to create a single test for all US students, to determine which ones had "the power to think clearly."

The SAT was first administered for college entrance in 1926. Promoted as a test of innate ability, it was comprised of ten timed subtests, including definitions, arithmetical problems, classification, antonyms, analogies, and paragraph reading. Nine of the ten subtests included sentences or paragraphs in *correct writing*. For instance, a classification section included the following directions:

> In each of the lines below, the first two words are related to each other in some way. You are to see what the relation is between the first two words, and find the one word in the parentheses that is related in the same way to the third word, writing the number of that word in the margin at the end of each line.

Some exam questions emphasized grammatical terms, such as, *An____is a word used to limit or qualify the application of a noun or a nominal phrase*. Other questions emphasized culturally specific information such as: *Three of the following words are related to each other in some definite way: Columbus, Socrates, Beethoven, Wagner, Verdi, Corneille. Which three words are most closely related?*.[7] As these examples show, to understand and respond to the 1926 SAT, the first requirement was to parse *correct writing*.

In the UK, secondary exams also became more centralized during the twentieth century. In 1918, the UK local examinations were consolidated in the British Board of Education's School Certificate Examinations. These were used across England and then the British colonies, based on the idea that uniform external examinations set an international standard. By the 1940s, UK standardized exams extended to primary schools. The 1944 Education Act, led by Cyril Burt, introduced the eleven-plus examination as a sorting tool. Only those students receiving high exam scores were selected for grammar schools, which emphasized university preparation.

The eleven-plus exam included sections on "general English," comprehension, and arithmetic. As in the SAT, students needed to comprehend *correct writing* throughout the exam, and in specific questions, they needed to know *correct writing* spelling and usage preferences, such as when to use *which* versus *whom*.

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

By the mid-twentieth century, standardized exams were common. By 1951, the UK had the General Certificate of Education (GCE) exam in England, Wales, and Northern Ireland, which expanded to secondary education certificates in the 1960s. By the 1960s, the US had the SAT, the American College Test (ACT), and several state exams for students graduating secondary school and applying to college. In these developments, we can again see *more access/more regulation.* As college access increased, so did entrance examinations, which emphasized *correct writing* within and beyond English exams.

4.2 The Myth Emerges

With standardized, high-stakes tests in the early and mid twentieth century, this myth emerged. Uniform, externally developed tests could now regulate writing. They could reward the right side of the continuum only, and select only *correct* writers for particular educational opportunities.

4.3 Consequences of the Myth

4.3.1 We Scale up Limited Definitions of *Writing* and *Intelligence*

In this myth, the writing and regulating valued in the first three myths becomes standardized in large-scale tests for secondary and college students. Regulating English was already *manifestly desirable*; in this myth, standardized tests become *manifestly desirable*, too.

Standardized tests depend on and propel all the myths so far. They scale up the limited mold of *correct writing* equated with ability, along with several related consequences noted in Table 4.1.

Table 4.1 Consequences of myth 4

Once we believe	… Every exam is an English exam
Tests must regulate writing, then…	… Exams emphasize exam writing
	… Exams emphasize English literature
	… Exams imply writing tasks don’t matter
	… Exam culture overshadows learning culture
	… Efficiency and ideal sameness prevail
	… Exams become the only obvious option
	… Extrapolation seems fine

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

4.3.2 Every Exam Is an English Exam

Once exams were written in English, a history exam was no longer a history exam. It was a test of *correct writing*, too. Parsing *correct writing* was necessary to score well on the SAT, and *correct writing* errors could hurt your chances for college entrance whether they appeared on your English composition exam or not.[8]

4.3.3 Exams Emphasize Exam Writing

Today, we have a veritable alphabet soup for regulating *correct writing*. Well-known examples include the UK's General Certificate of Secondary English (GCSE) and Advanced Subsidiary and Advanced levels (AS and A Levels), the International Baccalaureate (IB) diploma exam; Australia's Special Tertiary Admissions Test (STAT) and National Assessment Program – Literacy and Numeracy (NAPLAN); and the US Collegiate Learning Assessment (CLA), Advanced Placement (AP), and Accuplacer exams. All of these standardized exams depend on producing and/or parsing *correct writing* under timed circumstances. All of them reinforce earlier myths, by conflating *correct writing* and *intelligence* and bolstering two-dimensional ideas about writing, in which circumstances and tasks don't influence what students write.

For one thing, standardized exams have scaled up timed writing, despite ongoing concerns about it. Research shows that writing quickly leaves no time for regressions, multiple knowledge domains, or complex processing. For another thing, by leaving no time for revision and by emphasizing *correct writing* errors, standardized exams have scaled up error-hunting and error-reporting, reinforcing what writing historians Robert Connors, Lisa Ede, and Andrea Lunsford call the "cult of correctness" dawning as tests began to regulate written English.

Contemporary exam support resources reinforce language regulation mode. I was able to take two practice tests online, the US ACT Writing Test and the College Board Accuplacer. The US ACT Writing Test is used for college admission, and the College Board Accuplacer is used for writing placement. Both exams include multiple choice sections focused on *correct writing* errors. Questions include distinguishing between *that* and *which, there* and *their,* and *whose* and *who's,* as well as identifying comma usage and subject-verb agreement prescribed on the right of the continuum. In their test tips, Accuplacer encourages students to use grammar and spell check apps, which suggests that test criteria especially concern conventions and usage preference errors.[9] This advice brings us

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

full circle back to myth 1, because grammar-checkers are often informed by eighteenth- and nineteenth-century usage preferences.[10]

4.3.4 Exams and Courses Emphasize English Literature

Early English exams and courses emphasized literature more than language. There were exceptions: Some early university students received English language instruction in Scottish universities,[11] and before the 1970s, several US composition programs incorporated insights from linguistics. But it was not until a conference held at Dartmouth College in 1966 that educators began to challenge the exclusive emphasis on literature and lack of emphasis on written language instruction. By then, departments, courses, and exams had focused on literature for a century, and English composition courses were often taught by instructors trained in analyzing or producing literature.

Still today, many English instructors are trained in literary studies rather than (also) English linguistics, rhetoric, or composition. The university writing program I direct, for instance, is housed in an English department and largely follows this approach. In other words, many writing instructors are trained in literary studies, a discipline that favors particular genres (such as essays), evidence (e.g., literary forms), and other specific writing choices (e.g., emphasis on writers' interpretive reasoning versus empirical results). The instructors are rarely trained in language development or how writing in literary studies does and does not apply to other kinds of writing. In turn, many secondary and early college students end up practicing responses to literature, rather than studying a range of written English.

4.3.5 Exams Imply Writing Tasks Don't Matter

Early exams implied that exam writing tasks – what students were asked to do on a test or assignment – didn't matter. A good writer was a good writer no matter the task, a bad writer bad regardless, and so forth. In other words, if you were a good writer taking Oxford's 1884 English exam, you would write a good timed essay about "signs of the immaturity of Shakespeare's genius in Richard II," whether or not you had read and discussed *Richard II* in school before the exam. You would be the same *correct* or *incorrect* writer if you wrote a report on farming.

Fast forward seven decades, and Cyril Burt was promoting the same idea. After arguing that writing speed was an index of ability, Burt's eleven-plus exam offered writing tasks with no detail, such as the mid-twentieth-century exam task that was one word: School. Burt

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

further specified that the exam topic "should not be stated until the last moment, when … the test-period is about to commence."

It is still common to withhold exam topics until the start of exams, as was recommended by the UK Department of Education for the 2022 GCSEs:[12]

> …for subjects in which a choice of topic or content is provided (English literature, history, ancient history and geography), advance information about the focus of exams should not also be given. We believe the combination of the two measures would have the effect of giving students taking those subjects an unfair advantage and making the qualifications less rigorous.

More generally, the implication persists that writing tasks don't influence writing performance. That is why the University of Cambridge Local Examinations Syndicate (UCLES) exam directions, for instance, state, "This examination paper tests how well you can read, write, and present information," rather than something like, *this examination paper tests how well you can read, write, and present information in a timed essay addressing literature on this exam.*

4.3.6 Exam Culture Overshadows Learning Culture

When we prioritize the values of exam culture, we overlook a century of documented problems with uniform tests, several of which are documented by assessment historian Andrew Watts. An early twentieth-century criticism of the UK's local examination system was that it was out of touch, "ruled and regulated by middle-aged and even elderly gentlemen, who now have little to do with the education of the young, and in many cases have never had to do with it." The local examinations were also described as inequitable: "The achievements of a few were purchased at the expense of many," because the curriculum was designed around the few deemed able to sit the exams.

A similar criticism was leveled in 1928 at the "distorting and harmful effects" of using external examinations in British Overseas Territories. The external exams, noted the UK parliamentary under-secretary, constrained local secondary education by emphasizing English language and culture, and they tended to "favor a small class of selected students." When the local examinations were further consolidated in the School Certificate Examinations, the UK Board of Education offered a "cardinal principle" for standardized exams: "The examination should follow the curriculum and not determine it." Later still, in the 1940s, UK school reports argued that standardized examinations should be stopped, because they threatened the independence of schools and teachers' freedom.

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

When exam culture overshadows learning culture, the priorities of test developers prevail despite these problems and cautionary tales. This is why timed writing and spelling errors persist – not because they relate to learning priorities, such as sustained inquiry, revision, and reflection on writing choices – but because they are efficient to administer and evaluate. This is also why the US Accuplacer exam uses commissioned writing for the passages that students read and correct during the exam. Authentic writing supports student learning, because it is what students will encounter and produce in the real world. But the commissioned writing, narrowly designed and efficient, fulfills the priorities of exam culture.

4.3.7 Efficiency and Ideal Sameness Prevail

In myth 3, we saw Cyril Burt argue that tests were more reliable than teachers, so we won't be surprised to hear that he promoted the eleven-plus exam along these lines. Burt described that two standardized exam scores tended to show close agreement, while those of two teachers often differed – and therefore could not both be correct. For Burt, variance meant inaccuracy. There was no room for diverse responses to student work.

A similar message appeared decades later, when league tables ranked schools according to students' performance on the UK GCSE and GCE A-level examinations. These 1990s rankings were based on exam scores without attention to 3-D details such as test conditions, and there were documented concerns about the rankings at the time. However, these concerns led to more, rather than less, uniformity, because the ranking reports ultimately called for a standard that would not change over time or tasks. As Rebecca Zwick wrote in *Who Gets In?* thirty years later, exam scores give the "an illusion of exactitude" even though scores can be affected by test takers' moods, testing conditions, writing tasks, and lucky guesses.

4.3.8 Exams Become the Only Obvious Option

It is no mystery that standardized tests offer extremely limited information about students. Teachers have reported this for decades – that standardized tests offer little to no useful information about students' writing or broader literacy. But once we believe the myth that *writing* and *intelligence* need to be regulated in efficient and uniform ways, standardized exams become the only obvious option. Today, school funding is regularly linked to test scores, and teachers often have to focus on standardized test writing at the expense of varied writing tasks. Schools

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

are pressured to make reading and writing measurable and evaluative, meaning they have little choice but to focus instruction on exams.

4.3.9 Extrapolation Seems Fine

We saw false extrapolation in the last myth: Early intelligence tests went from "Jane can't write X" to "Jane is not capable." In this myth, we see similar moves here: going from "Jane can't write X exam" to "Jane can't write." This happens when we use specific tasks to draw general conclusions.

Standardized tests can affirm one another in cycles of generalization and extrapolation. This happens when similar standardized test results are seen as evidence of validity, rather than what they are: evidence of consistency across similar tests. This was the case in a much-cited study we will see in the next myth, in which students who did well on the Collegiate Learning Assessment (CLA) also had high SAT/ACT scores. The authors of the study argued that similar CLA and SAT scores proved their conclusions about student writing. A different interpretation is that students will perform similarly on similarly narrow testing instruments.

To label a student an *incorrect writer* because of spelling or usage on a timed exam generalizes and extrapolates beyond that timed exam, but this is how exam results are often used.

4.4 Closer to the Truth

4.4.1 Standardized Test Scores Measure Socioeconomic Status and Test Preparation

Closer to the truth is that – like IQ test scores – standardized exam scores (including SAT, ACT, and GCSE scores) correlate with socioeconomic status (SES). SES is also associated with choices in A-level subjects, which impacts college admission.

Also closer to the truth is that SES often determines whether students have money or time for test preparation, and test preparation impacts test scores. One-on-one test tutoring, for instance, has been shown to significantly improve standardized test scores. Because it is clear that test preparation makes a difference to make timed standardized writing exams more fair, all students would need regular practice with timed standardized exam writing. But this would mean even more exam culture.

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

4.4.2 Problems Are Well-documented, but Efficiency Rules

Standardized tests tend to be uniform and narrow, while writing and students are diverse and expansive. Anya Kamenetz, author of the book *The Test: Why Our Schools are Obsessed with Standardized Testing – But You Don't Have to Be*, puts it this way: "The way much of school is organized around these tests make little sense for young humans developmentally. Nor does it square with what the world needs."

Problems range from test design, to performance, to use. In terms of design, closer to the truth is that standardized tests leave out important knowledge we saw in the last myth, including interpersonal and intrapersonal knowledge related to leadership, collaboration, adaptation, and creativity. Another test deisgn concern is that when standardized tests emphasize culturally specific knowledge, they are not equally fair for all students. And when exams test language knowledge even when they are not language exams, they are not valid: There is a mismatch between what they are measuring and what they are claiming to measure.

In terms of test performance, closer to the truth is that students respond to test situations. Everything from misunderstandings to cognitive overload can influence students taking an exam on a given day. This is why for any test performance, there are myriad explanations aside from ability.

In terms of test use, standardized test scores are often used in college admissions, but closer to the truth is that test scores do not predict how most students will do in college. Standardized exams emphasize narrow tasks and narrow domain knowledge, while college learning is less narrow. The SAT Writing and Language Test, for instance, has a low correlation to students' first-year college grades, and an even lower correlation with their first-year course grades in English and writing. Perhaps the reason for this failure of concurrent and predictive value is the fact that the test requires no student writing whatsoever.

Closer to the truth is that standardized tests emerged after several writing myths had already emerged. Most of today's contemporary tests have been guided by past test instruments and have not been sufficiently tested for fairness. Educators have therefore called for alternatives, including portfolios and collaborative assignments. These alternatives are more like writing outside of test situations, meaning they are more varied and less efficient.

4.4.3 Tests Must Be Tested

Closer to the truth is that writing is complex, and testing tests is tricky. How can we tell a test is accurate? Test results can be compared with

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

students' grade point average (GPA), but GPA is often partial and inconclusive. Secondary writing exams can be tested against how students do in their college GPA, but as we saw above, these measures have highlighted the poor predictive power of standardized test results.

Writing exam scores and GPA can be tested against student writing success throughout and after college, but such studies are challenging and rare. Even so, it is notable that rare research like this – following postsecondary writers for multiple years, based on how well they do on different kinds of writing – paints a rosier picture than standardized test results. Closer to the truth is that more diverse writing offers a more robust picture of how students write, and the robust picture is a more positive one. The bleakest picture we get is from narrow testing of narrow writing.

4.4.4 Tests Only Test What Is on the Tests

Closer to the truth is that no single writing test can test writing ability. Specific writing tasks directly shape writing, soliciting certain writing choices and not others, a phenomena that I've found fascinating to research over the past several years. I like the assessment term "constructed response tasks" for this reason. Reading and writing tasks construct the responses to them.

For one thing, as demonstrated in Table 4.2, students' familiarity with the **topic** matters. If students know something about a writing topic, they have more working memory for their writing choices, like cohesion and usage conventions. To use an earlier example, a student who had read and discussed *Richard II* before the Oxford 1884 exam would have more bandwidth for writing choices than a student who was less familiar with *Richard II*. This explains why students with relevant prior knowledge tend to produce what evaluators consider more fluent writing.

For another thing, **how** students are asked to write matters. Writing is influenced by whether it is timed or untimed, or takes the form of an essay or a report, and so on. When a test parameter changes, writing changes.

Let's start with the fact that students write differently if they have ample time to write. There are obvious reasons for this: Untimed writing means students have time to revise their spelling, punctuation, and other usage conventions. But there are less obvious reasons, too. Untimed writing means students engage more with other texts and perspectives, which is an expectation of most college writing.

Most subtle is that timed and untimed writing have different language patterns. When students write quickly, they tend to use language patterns on the left of the continuum, probably because they are the most

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

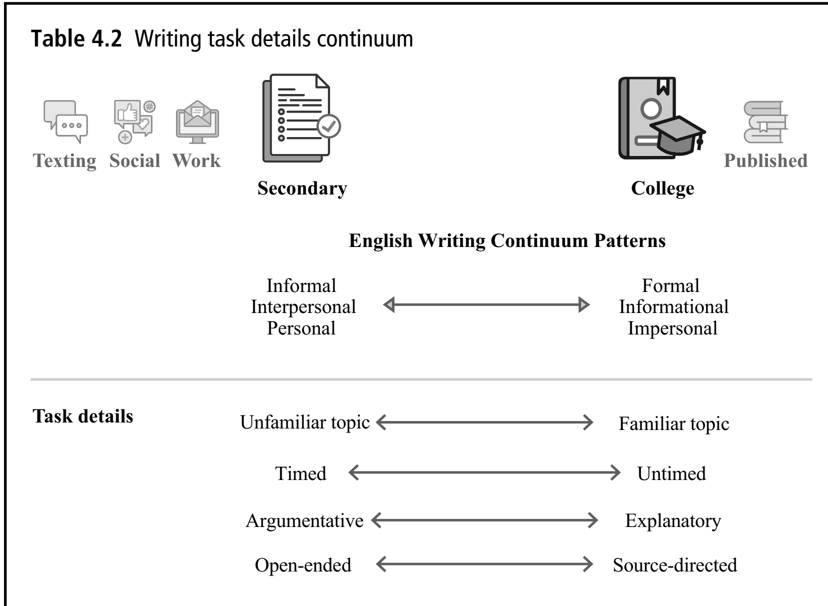

Table 4.2 Writing task details continuum

practiced and familiar. Timed writing includes significantly more personal and interpersonal language patterns, such as boosted statements and text-external first-person pronouns. This matters because *correct writing* tends to include impersonal, informational language patterns instead, as we've seen since myth 1.

Students also write differently based on genre – a personal narrative, an argumentative essay, and a summary report all have different language patterns, even by the same student. Personal narratives tend to include storytelling moves, interpersonal connection (*you won't believe what happened*), and personalized stance patterns (*I'm so excited*). Argumentative essays and other persuasive writing include significantly more generalizations (e.g., ***everyone** has cheated at some point*) and boosted claims (e.g., *cheating is **clearly** wrong*). Explanatory writing is less likely to include boosters and more likely to include noun phrases (e.g., *increasing reports of cheating are a concern for educators*). Here again, these patterns matter, because *correct writing* is usually informational and impersonal, or more toward the right of the continuum, meaning readers tend to expect few generalizations and more noun phrases.

Thus topic, timing, and genre all make a difference. But it doesn't stop there. Research shows that students responding to open-ended questions (e.g., *Why do students cheat?*) write differently than those responding to

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

Table 4.3 Writing task details continuum

Continuum Purposes	**English Exam Writing Continuum Patterns**		

Cohesion	• Greetings and closings, narrative moves	• Explicit paragraphs	• Explicit moves, transition words, paragraphs
Connection	• 2nd person pronouns, text external 1st person	• 1st person references to collective needs or experiences (*we, our*)	• Rare 1st person, text internal 1st person
Focus	• Sentence subjects emphasize people, experiences, events	• Sentence subjects emphasize people, observations (*there is*)	• Dense noun phrase subjects emphasizing concepts, phenomena, research
Stance	• Boosters and generalizations, personal reactions including strong evaluations	• Boosters and generalizations, some hedges	• Hedges, some boosters, qualified generalizations
Usage	• *Correct writing* conventions and usage preferences	• *Correct writing* conventions and usage preferences	• *Correct writing* conventions and usage preferences
Example tasks:			
The University of London Entrance Exam, 1838 **Composition Exam**		Give an analysis of the part of the first book of Thucydides, which relates the causes and occasions of the Peloponnesian War, with the substance of the arguments used in the speeches. How far did the character and actions of Pompey entitle him to the epithet of The Great?	

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

Table 4.3 (cont.)

Cambridge Local Examination, 1858 English Composition Junior Exam	**Choice 3** Write a letter to a friend in Australia announcing your intention to emigrate, and asking for information	**Choice 4** Discuss the change produced in the habits of the people by Railways
Cambridge Examination for Women, 1870 English Composition Exam	**Choice 3** The question of compulsory emigration as a means of relief to national destitution **Choice 4** The question of opening to women professional careers, especially that of medicine.	**Choice 1** The political position of Greece, with a review of its history from the beginning of the century. **Choice 2** A comparison of French and English tragedy, illustrated by special comparison of Racine with Shakespeare.
Oxford University 1884 Women's English Examination	[Are there] any parts of *Macbeth* which seem so unworthy of Shakespeare as to justify a doubt as to their being genuine? [Are there] any signs of the immaturity of Shakespeare's genius in Richard II? [Are there] any traces of a failure of dramatic power in the *Tempest*?	
	Summarise Milton's arguments against the censorship of the press. Which do you consider the most convincing, which the most rhetorically effective?	Explain, by reference to this or any other of his prose writings, Milton's idea of Liberty.
Australia STAT 2009 Written English **Part A**	**Comment 1** Education helps individuals grow and has a civilising and humanising influence on society as a whole.	

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

Table 4.3 (cont.)

		Comment 2 Too much of current education is concerned with rote learning that has little relationship to real problems and real life.	
Australia STAT 2009 Written English **Part B**		**Comment 6** Romances come and go, but it is friendship that remains. **Comment 7** It is important that we learn to be confident within ourselves rather than dependent on the good opinion of others.	
Cambridge International A Level English Language 2016 Exam	Write a 150-250-word diary entry as though you are the author of the speech you read.	Comment on the style and language of a speech made by Australian Prime Minister Julia Gillard	
New Zealand IELTS, 2022 Academic Writing Exam			Explain pie charts representing energy use across countries.

source-oriented questions (e.g., *Do you agree or disagree with what the author says about cheating?*). Specifically, students responding to open-ended questions use more informal, interpersonal, and personal patterns, including generalizations and references to personal experiences. By contrast, students responding to source-oriented questions use more formal, informational, impersonal patterns, including more references to source texts and fewer references to personal experiences. I've mapped all of these general writing task trends onto the continuum in Table 4.2.

Closer to the truth is that every writing task constructs *correct* and *incorrect* writing in specific ways, and only the writing done for a particular task is measured in its evaluation. Any writing exam score tells one story about a writer – based on the test conditions, topic, genre, and other test parameters – and not another.

This doesn't mean that all writing exams are bad. It means that what they can tell us is limited. Closer to the truth is that no writing exam can tell us whether a student can write. Each one tells us how a student writes on that exam.

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press

4.4.5 Writing Exam Tasks Are on a Continuum

We can now use Table 4.3 to add secondary and college exam tasks to the writing continuum, according to the language patterns associated with them. More informal, interpersonal, personal exam writing such as diary and letter writing will tend toward the left end of the continuum, while summary and synthesis writing will tend more toward the formal, informational, impersonal end of secondary and college writing. Because it is both personal and informational, persuasive essay writing tends to fall in between.[13]

As you can see on this task continuum, several exams on this task continuum offered students a choice, as though the tasks were all the same. Closer to the truth is that the task matters, and there are writing patterns associated with particular tasks. Thus students choosing different options will have different chances of success, particularly if examiners expect only the formal, impersonal, and informational patterns of *correct writing*.

For all of these reasons, the design and use of most standardized exams keep the myth glasses firmly on. We'll see similar themes in the next myth, that most students can't write.

https://doi.org/10.1017/9781009231299.005 Published online by Cambridge University Press