

CHAPTER THREE

Myth 3 You Can't Write That and Be Smart

Or, Writing indicates natural intelligence

This time, a freebie for *pick a century*. The passages below are identified for you and are betrayed by wording besides. Still, they show a consistent message over time.

1. **Nineteenth century** (1845, Horace Mann in *The Common School Journal*)

[Writing exams offer] a transcript ... of the state and condition of the pupils' minds ... taken and carried away for general inspection.

2. **Twentieth century** (1925, Cyril Burt in *The Sub-normal School-child*)

The reader... may deduce, both from the physiognomy and from the style of writing, spelling, and expression, what were the intelligence, the temperament, and the educational attainments of the several children.

3. **Twenty-first century** (2021, online, with an intelligence test that takes twenty to thirty minutes)

Have you ever wondered how intelligent you are compared to your friends, your colleagues ... and the rest of the nation? The Great British Intelligence Test ... can reveal all.

Different though they are, these passages reflect the timeworn quest for a single way to measure and rank intelligence. Typing “how intelligent am I?” into Google in the year 2023, I get 454 million links in under 0.4 seconds. Most sources promise that I can “find out where I stand” in less than half an hour.¹ On the first page of results, I have eight options for taking an intelligence test or quiz. By the second page, I can read how to “make my writing sound more intelligent” and how “science explains why people who love writing are smarter.”²

Here are easy some conjectures from this search:

- People are very interested in testing intelligence.
- Intelligence is something some people possess more than others.
- Intelligence can be tested easily.
- Formal written English is part of intelligence.

These ideas pair easily with our first two myths, which limit *correct writing* and writing in school.

In this myth, *correct writing* is further limited, and further empowered. *Correct writing* becomes testable, connected to innate ability, and used to decide who is *intelligent*, and who is not.

Before that happened, people were not all tested the same way. They were not subjected to the inspection described by Horace Mann in passage 1. They would be, however, thanks to an origin story that begins with nineteenth-century science.

3.1 Context for the Myth

3.1.1 *Correct Writing* Becomes a Tool for Ranking and Selecting

Horace Mann once sent a plaster head to his sister in the mail. He was sharing his fervor for phrenology, the nineteenth-century science of measuring head size and shape to determine ability and goodness. At the time, scientific developments were beginning to reach general audiences, and phrenology was fueling public interest in calculating intelligence and morality. By that time, too, we know, *correct writing* already dwelled in popular imagination as a sign of character and capability, thanks to myths 1 and 2.

Enter Mann, who was was a firm believer in myths 1 and 2. Mann merged these beliefs with his conviction that eugenic progress – the notion that racial improvement occurs through selective breeding – was possible through education. Education perfected human nature, he argued, and *correct writing* provided a way to examine where individuals were in that perfection process. Mann was among the first to bring phrenological ideas and *correct writing* together, in written English examinations.

Mann's first step was to blindside students with unexpected tests. In 1845, on a day when Boston primary and secondary students went to school as usual, they were greeted with unannounced, timed, individual tests of written English. Mann promoted the exams as “impartial” testing that avoided teacher “interference.”

This approach to testing was new. Until then, US primary and secondary students demonstrated their learning in annual, interactive exhibitions open to parents and other community members, who could attend and ask questions. Mann's Boston area school exams, by contrast, were unexpected, externally designed, individual, and written.³

Mann's use of the tests was also new. He used overall exam scores as well as the number of *correct writing* spelling, grammar, and punctuation errors to publicly compare and rank Boston schools. For example, Mann showed that the "number of errors committed in Grammar" was 98, and the "number of errors committed in Spelling" was 97, for the students tested at the Adams School, while it was 199 and 91, respectively, for students at the Boylston School.

A few years later, Mann became president of Antioch College in Ohio, where he instituted the first college entrance exams in English composition in 1853. These, too, were evaluated for spelling, grammar, and punctuation errors according to the right side of the continuum.

In other words, first, Mann made *correct writing* errors the basis for public ranking and inspection of schools and students. Then, Mann used *correct writing* errors as the basis for college admission.

We've already seen where that went: Two decades later, *correct writing* entrance exams gained visibility when Charles Eliot began using them at Harvard. Eliot, whom we met in myth 1, really took to human ranking and hierarchy based on mythical ideas about *intelligence*. In particularly horrific examples, he called for racial purity and forced sterilization of the physically disabled as protection from "moral degeneracy." And he believed that only students with testable *innate ability* deserved excellent education. In his very first address as Harvard's president, Eliot made this clear:

The community does not owe superior education to all children, but only to the *élite*, – those who, having the capacity, prove by hard work that they have also the necessary perseverance and endurance.

For Eliot, similar logic meant Harvard could not consider admitting women. "The world knows next to nothing about the natural mental capacities of the female sex," Eliot declared. Accordingly, it would take "generations of civil freedom and social quality ... to obtain the data necessary for an adequate discussion of women's natural tendencies, tastes, and capabilities."

Like Mann, Eliot implemented English writing exams for college admission based on the idea that *correct writing* usage, punctuation, and spelling indicated general aptitude. In particular, Eliot favored *correct writing* in response to questions about English literature. By this logic, in other words, *correct writing* of timed essays on "Mr. Darcy's Courtship" (like those in myth 2) indicated innate fitness for college study.

Thus nineteenth-century eugenic theories fueled interest in uniform ways to rank and select students. Still to come were tools for ranking that went beyond a single geographic area or institution.

3.1.2 *Intelligence becomes innate and testable*

As the nineteenth century turned, interest in evolution and evaluation grew. Charles Darwin's observations of his son excited interest in behavioral development, while studies by Darwin's cousin Francis Galton fueled interest in intelligence and individual variation. In 1905, psychologist Alfred Binet was commissioned by the French government to identify children needing an alternative to typical schooling. Based on their observations at a French boys' school, Binet and his student Theodore Simon created the Binet-Simon test.

As with intelligence tests that followed, the Binet-Simon test was based on the developers' definition of intelligence and how it could be measured. Binet and Simon saw attention, memory, and verbal skill as part of intelligence, and they selected thirty items to measure them, such as touching one's ear, drawing designs from memory, and defining concepts. Students took the test individually and received a score for those items they answered as expected. This score was then divided by the students' age and multiplied by 100, and the resulting number represented the student's "Intelligence Quotient," or IQ.

During the same period, England psychologist Charles Spearman and US psychologist Edward Thorndike were developing theories of testing and mental measurement. In 1904, Thorndike promoted tests for use on large populations, and Spearman promoted tests that "objectively measured general intelligence."

3.1.3 *Writing Becomes Part of Ranking Intelligence*

More interest in general intelligence testing meant more interest in going beyond institution-specific tests like the entrance exams we saw in the last myth. What could provide a single way to measure whether students were producing *correct writing*?

An answer came in 1912 in *A Scale for the Measurement of Quality in English Composition by Young People* by Milo Hillegas. Hillegas had worked with Thorndike to develop the scale, and he published it as a way to sort and rank *correct writing* across US secondary and college English courses.

In his introduction to the scale, Hillegas villified writing measurement that varied across classrooms. Instead, he promised, his uniform scale offered "proper" and "exact" standards for making comparisons across schools and school systems.

The scale included hundreds of short passages from several sources. Some were fabricated, some were by students, and others were by literary

authors. For instance, “specimen 217” was by Nathaniel Hawthorne, while “specimen 221” was by a secondary-school student writer. “Specimen 519,” with “the least merit of any of these,” was also by a secondary-school student writer. The specimens were labeled by author (if by a published writer) or by level (if by a student), or by the label “artificial sample” (if fabricated). They were not labelled according to their writing task or genre (e.g., argumentative essay assignment; literary novel).

To use the scale, educators were told to compare their students' writing to the sample passages, which had been ranked by Hillegas and his team according to *correct writing* usage preferences, spelling, and punctuation, regardless of the task. A short artificial sample representing a value of 0, for instance, twice included the dialect-specific usage *ain't*, which was also used as an example of “inaccuracy” in nineteenth-century Harvard English exams.

Concerns about the scale arose at the time. Educators critiqued the accuracy of the samples, the methods used for the scale, and the insufficient detail about the results. They questioned whether any single scale could capture the complexity of writing, in terms of both discursive (style-based) and propositional (idea-based) content. One critique noted that the scale was of little practical value beyond supporting uniformity, then went on to question whether uniformity was desirable.

The supporters of the scale, however, said it represented writing “merit” and “superiority,” regardless of writing task or classroom context. Similar scales followed that shared the goal of ranking writing according to a uniform instrument.

3.1.4 *Intelligence* Becomes a Tool for Ranking and Selecting

As the Hillegas scale was reaching writing classrooms, intelligence testing was gaining visibility in the UK with the help of British psychologist Cyril Burt. Burt was adapting the Binet-Simon test to measure character, learning, and intelligence, which was labeled “mental ability” and “mental defect.” In 1914, Burt published a review of the Binet-Simon test that promoted three ideas about intelligence:

- Capability is innate and testable.
- Tests can rank individuals and groups according to uniform ability or “mental age.”
- Tests and scales need ongoing work to increase uniformity.

In other words, for Burt, diverse test responses were a problem to be solved, rather than a sign of developmental diversity. For instance, the

fact *normals* could define concepts before they could count thirteen pennies made *defectives* – who counted pennies before defining concepts – developmentally *backward*.⁴ In addition to his ideas about intelligence, Burt promoted three ideas about tests:

- Tests should measure capability in units of their own (not age), according to how completely and quickly an individual performs a task.
- Tests should have separate, standardized schemes according to the race, sex, and social class of test takers.
- Tests should not leave room for diverse responses.

In the US, Lewis Terman at Stanford University promoted similar ideas. Terman adapted the Binet-Simon test to create the Stanford-Binet intelligence scale, which was a tool for measuring *innate capability* rather than capability at one point in time. In 1916, Terman published *The Measurement of Intelligence*, which promoted the Stanford-Binet intelligence scale as a quantifiable way to determine *idiocy*, *feblemindedness*, and *genius* through a single, simple test.

Terman’s own racist and misogynistic ideas – he believed that marriages of white Americans following traditional gender roles would promote eugenically fit children, for instance – informed his efforts. He aimed for an easy, scientifically accepted way to facilitate eugenic sorting. By 1920, the Stanford-Binet test, and its assumptions about innate and testable intelligence, were circulating in and beyond the US.

A prominent application of the Stanford-Binet test was US Army Alpha intelligence testing, which was used to rank soldiers for officer roles. The Army Alpha Test was timed, written in *correct English*, and included multiple choice questions with culturally specific answers. The three questions in Figure 3.1, for instance, were part of the test.

In 1923, Carl Brigham reported Army Alpha Test questions, results, and interpretations in his book *A Study of American Intelligence*. Brigham’s main argument – supported, he said, by “the teeth of the

Christy Mathewson is famous as a	Carrie Nation is known as a	Crisco is a
<ul style="list-style-type: none"> • writer • artist • baseball player • comedian 	<ul style="list-style-type: none"> • singer • temperance agitator • suffragist • nurse 	<ul style="list-style-type: none"> • patent medicine • disinfectant • toothpaste • food product

Figure 3.1 Early IQ test questions, Army Alpha test

facts” – was that the Army Alpha test results showed “the superiority of the Nordic race.” He further interpreted (surely based, again, on facts with teeth) that “In a very definite way,” the results showed “pure Nordic peoples” were “rulers, organizers, and aristocrats,” characterized by “a greater stability and steadiness than are mixed peoples.” Furthermore, in light of test results showing that more time in the US led to better performance, Brigham blamed immigrants rather than the test, arguing that earlier immigrants were smarter. The absence of validity in the test questions is shown Figure 3.1. Knowledge of baseball players, suffragists, and food products is inextricably bound to cultures and places rather than innate ability.

Reviewers of *A Study of American Intelligence* questioned its scientific rigor at the time. Psychologist Edward Boring’s review in *The New Republic* specifically argued that Brigham “discarded the effect of a knowledge of English in stating his differences.” This matters singularly because, as Boring noted, “measurable intelligence is simply what the tests of intelligence test, until further scientific observation allows us to extend the definition.” In this case, Boring noted, *intelligence* was knowledge of formal English.

Brigham did not agree. *A Study of American Intelligence* insisted that knowledge of English did not explain the Army Alpha Test results: the test measured “native or inborn intelligence,” and the ability to use English was “a function of intelligence and education in its broadest sense.” In other words, Brigham interpreted the test results in support of his idea of racial purity, and he paid no mind to the role of *correct English* or culturally specific knowledge.

Together, Burt’s promotion of uniform tests of mental age, Terman’s marketing of test efficiency, and Brigham’s Alpha testing all sold the idea that testing innate ability could be simple and efficient. Plenty of support, in other words, for narrow tests, and narrow interpretations, of *intelligence*.

3.1.5 Intelligence Tests Appeal to Scared Racists

The myth that intelligence is innate and measurable had special allure for fearful racists in the early twentieth century, people concerned about growing immigration and diverse school enrollment. Robert Yerkes, an Army Alpha Test developer and Brigham’s teacher, wrote the foreword to *A Study of American Intelligence*, basically characterizing the book as a primer for xenophobia:

[Brigham] presents not theories or opinions but facts. It behooves us to consider their reliability and their meaning, for no one of us as a citizen can afford to ignore the menace of race deterioration or the evident relations of immigration to national progress and welfare.

In these ideas, we can see the trend of *more access /more regulation* that characterizes several myths. Greater access to education and other national resources came with greater regulation of those resources. Intelligence tests could aid selection, and tests in *correct writing* could favor specific test-takers. Then, to use Eliot's logic, resources could be handed to those most deserving.

3.2 The Myth Emerges

From these origins, we can identify our third myth, in which *correct writing*, already narrowly defined, indicates *intelligence*, also narrowly defined. The characters in this story did not share identical motives, nor were their efforts identical. But phrenological thinking, error ranking, IQ tests, and writing scales came together in terribly complementary ways. The myth emerged, making ability innate and measurable in simple tests. In those tests, *correct writing* was either a clear indication of *intelligence*, or a neutral vehicle for reflecting it.

3.3 Consequences of the Myth

3.3.1 We Narrow *Intelligence*

The overall consequence of this myth is that we limit ideas about *intelligence*. Despite multiple learning domains, in this myth, only the *cognitive* (or discursive) domain, measured through *correct writing*, counts as *intelligence*. Left out are domains that emphasize collaboration, reflection, and well-being.⁵ Figure 3.2 illustrates four domains of human ability, and here's how they apply to the Army Alpha test question about Crisco shown in Figure 3.1:

- The **cognitive** or *discursive* domain is related to reasoning and memory.
Early tests and scales emphasized this domain. Test takers would, for instance, use the cognitive domain in the Alpha tests to recall how they had seen the word Crisco used, and what it was.
- The **interpersonal** domain is related to collaborating with others.
Early tests and scales left out this domain. Had they included it, test taskers might have used the interpersonal domain to consider how to

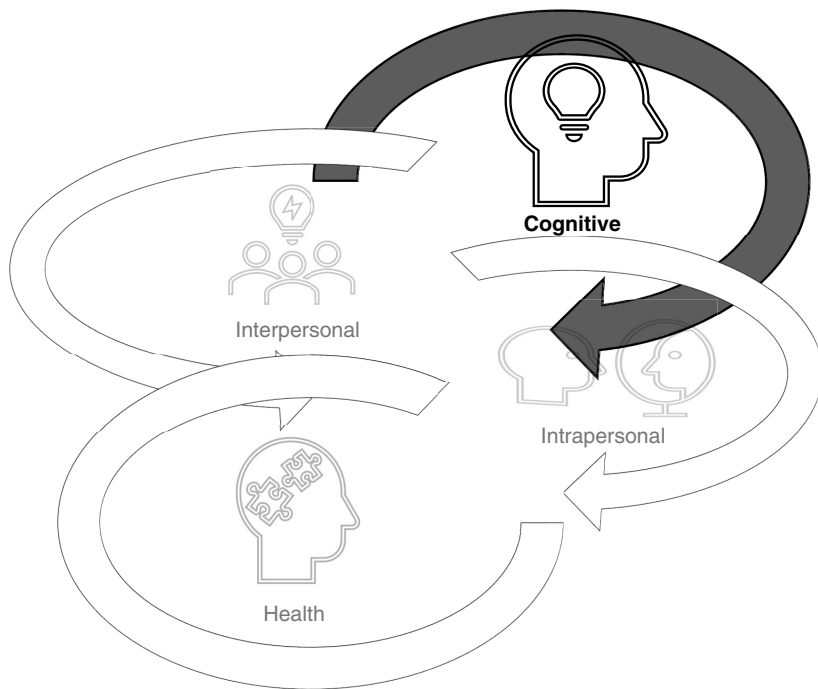


Figure 3.2 A four domain model of writing

ask others for information, or to brainstorm together to narrow down their multiple-choice answers about Crisco.

- The **intrapersonal** domain is related to reflecting and self-moderating. Early tests and scales left out this domain. Had they included it, test taskers might have used the intrapersonal domain to consider what steps they had used in the past when they didn't know an answer, and to determine whether the same steps could be used to determine the meaning of Crisco.
- The **health** domain is related to well-being. Early tests and scales left out this domain. Had they included it, test takers might have used the health domain to ensure they were safe, rested, and fed enough to be able to focus on what they knew about Crisco.

This is the overall consequence of this myth: Instead of a complex, dynamic understanding of knowledge and experience, we limit ideas about *intelligence*. With that comes several more specific consequences listed in Table 3.1.

Table 3.1 Consequences of myth 3

<p>Once we believe</p> <p><i>correct writing</i> indicates natural intelligence, then...</p>	<p>... Only some people are <i>intelligent</i></p> <p>... Writing and ability are 2-D</p> <p>... Efficiency and ideal sameness become values</p> <p>... Trust shifts from teachers to tests</p> <p>... We trust tests without understanding tests</p> <p>... Extrapolation seems fine</p>
--	---

3.3.2 Only Some People Are *Intelligent*

Early intelligence tests privileged the knowledge and experience of some people (including the test designers) and not others. In turn, they rendered only some groups and individuals *intelligent*. In obvious examples, early US IQ tests supported human ranking to the disadvantage of immigrants, women, and races considered mixed or not white.

This unfairness occurred on multiple levels: test design, test interpretation, and test use. Test design was unfair in that certain groups were much less likely to be familiar with *correct writing* or culturally specific details than others. Test interpretation and use were unfair because narrow test scores were used to label entire groups – like Brigham’s “pure Nordics” – *capable* and *moral* or, alternatively, *feeble-minded* and *immoral*.

In cyclical fashion, ideas about *intelligence* were also used to bar people from even attempting to prove themselves through narrow tests. Eliot barred women from Harvard entrance exams because not enough was known (through exams) about their innate ability. It would be more than a century before undergraduate women were included at Harvard, despite the efforts of numerous women throughout that time.

In most of these examples, only the mythmakers told the stories. No one heard from the *nameless feeble-minded girl* in the work of eugenicist Hedy Goddard. They heard from Goddard.

3.3.3 Writing and Ability Are 2-D

Writing and *writing in school* were already narrowed by myths 1 and 2, and with this myth, *intelligence* and *correct writing* are further limited. They become two-dimensional – treated as though writers and language

can operate independent of context and purpose. IQ testing limited *intelligence* to something individual and testable, separated from other people and writing outside of test conditions. Writing scales similarly limited writing, suggesting published literature and student writing could all be ranked on the same scale.

When writing and ability are 2-D, *correct writing errors* do not depend on context and are an efficient focus for uniform tests. This explains why after 1915, errors, particularly in conventions like punctuation and spelling, became a central focus for large-scale writing evaluation.

3.3.4 Efficiency and Ideal Sameness Became Values

Three priorities fuel this myth:

- efficiency (a “simple test”)
- uniformity (a single test or scale for everyone)
- ranking people (according to a single, simple test)

Through the myth glasses, anything that doesn't prioritize these values is suspect, and uniform tests are used even when they don't appear to serve all students, schools, and knowledge domains. Multiple ways to respond to the same task becomes too messy, for instance. Different assignments for different students becomes too time-consuming.

Even after overtly oppressive IQ testing was no longer permitted as a unitary judgement of intellect, many tests today maintain these priorities, valuing efficiency and uniformity over diverse domains and experiences.

3.3.5 Trust Shifts from Teachers to Tests

Valuing efficiency and uniformity inevitably shifts who, and what, has a say in education. More trust in externally designed tests and scales means less trust in specific schools and classrooms. More trust in timed writing means less trust in a teacher's observations of a student's work over a term or a year. And in the event of poor test results, more trust in tests means students and schools, rather than testing instruments, are to blame. Mann's publicized Boston school error rankings, for instance, placed Smith School, serving Black students in Boston's segregated common schools, in last place, based on Mann's interpretation of unanswered questions. But Smith and its students, rather than the uniform test and interpretation, took the blame.

In another example, Burt's 1914 description depicted test scores as the positive alternative to variation and teachers' input. Burt described the correlation between "teachers' estimates" on the one hand and "absolute mental age" on the other, arguing that internally graded tests that didn't align with uniform test scores were to be "rejected, as no tests of intelligence at all." Burt's wording illustrates a key consequence of this myth: A local teacher's evaluation is not only less absolute than that of intelligence tests, but it is also dismissible.

Some twentieth-century educators and leaders regretted this distrust in teachers. The 1943 Norwood Report described problems with uniform UK secondary examinations, saying that they neglected local teachers' knowledge. The report ultimately recommended that exams be set internally in schools by local teachers. But this did not come to pass.⁶ With this myth came the message that teachers were more subjective and less reliable than tests, and that diverse assessment conclusions indicated a mistake rather than inevitable – or instructive – variation.

3.3.6 Learning Culture Shifts to Exam Culture

The shift away from teachers toward efficiency and uniformity was a shift toward exam culture. Exam culture means less emphasis on learning culture, because the needs of test designers are not the same as the needs of specific classrooms and students. If sometimes well intentioned, exam culture promotes consistency (or *reliability*) above all other concerns, no matter how narrow the tests or scale, and no matter what or who is left out. The same 1943 Norwood Report cautioned that students had begun viewing education only in terms of exams, but exams persisted.

3.3.7 We Trust Tests Without Understanding Tests

We've already seen that writing myths prioritize language regulation, even when that comes with language ignorance. The same thing happens in this myth: Limited ideas about intelligence, and limited knowledge of testing, flourish in tandem. In public campaigns for IQ testing, developers like Terman and Brigham emphasized accepting test results, not understanding test design.

A result is that tests and *correct writing* matter, even as they are not well understood. Early efforts to sort and rank immigrants, races, soldiers, students, and other groups didn't invite understanding of

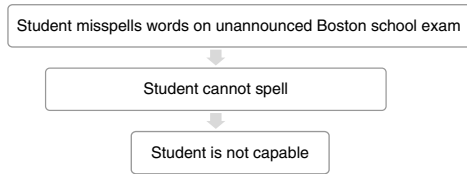


Figure 3.3 Extrapolating from a test to *intelligence*

tests or diverse language use. Early reports emphasized test results and didn't investigate the impact of test design on test taker performance. We will see this consequence persist throughout the coming myths, in headline stories that focus on exam results rather than exam details.

3.3.8 Extrapolation Seems Fine

When *intelligence* is limited and testable, extrapolating from test results seems fine. Extrapolation is a move from an instance (say, an IQ test) and a related instance (skills tested on the IQ test) to the widest possible inference (that a student is not capable). We go from "Jane can't write X" to "Jane can't write," and then from "Jane can't write" to "Jane is not *intelligent*." Figure 3.3 shows the logical failures that occur when causal logic is used without regard to the many factors that contribute to spelling ability, and the cruel leaps that occur when we consider *intelligence* as unitary.

We can illustrate this using the Boston school error rankings in Figure 3.3. Mann's reports moved from (1) a student misspelling on the Boston school examination, to (2) generalizing that the student couldn't spell, to (3) extrapolating that the student was lacking innate capability.

We can call this dynamic *limited test/general use*, because it means using limited test content to draw general conclusions about test takers. *Limited test/general use* implies that a score from a test can indicate abilities beyond that test. We saw this in the case of the Army Alpha intelligence tests. In content, the Army Alpha test emphasized cultural knowledge and *correct writing*. In use, the test results inferred innate capability and fitness for officer leadership.

In this myth, false extrapolation appears necessary and reasonable. In turn, the myth permits sweeping claims without sweeping information.

3.3.9 Uniform Tests Tell a 2-D Story about Writing

Early tools such as Mann's exams and Hillegas' scale offered a two-dimensional story about writing. They intended to do so: A 2-D story is most efficient and consistent. A 2-D story discusses writers, and writing, as though dimensions like context and multiple knowledge domains don't matter.

Still today, many tests imply that writers will write the same way no matter the circumstances: Students still sit down and take specific writing tests, after which people draw general inferences about their ability. In the case of many secondary leaving exams, for instance, examiners infer from individual, timed essay exams whether students can read and produce untimed writing, or whether the student is prepared for college study. In this 2-D story, it is easy to compare and rank writers against one another, and it is easy to extrapolate from *limited tests* to *general use*.

3.4 Closer to the Truth

3.4.1 Uniform Tests Are not Fair

It has been evident for decades that IQ tests are seriously biased and have other limitations. Documented most of all is the clear connection between IQ scores and socioeconomic status (SES), even for children as young as two years old. It is also clear that factors related to test conditions, such as test anxiety, test environment, and examiner effect, make IQ tests unreliable.

3.4.2 Writing Is 3-D

Closer to the truth is that writing is 3-D: it is not just a writer and written language, but writers and language and contexts in dynamic interaction. (Or, in a sentence more to the right of the continuum: Writing is socially constituted meaning-making.) A writing exam, for instance, occurs within an exam context, and it cannot represent a range of writing in non-exam contexts. Only a range of writing with different purposes and audiences illustrates writing for a range of purposes and audiences.

One reason writing is 3-D is that **writing is social**. It responds to audiences and identities across a continuum of writing. My practice at writing formal research articles doesn't get me far when I try to write an informal blog on the same topic, unless I realize what language patterns do and do not apply.

Relatedly, **writing is diverse**, across a continuum of different values and practices. It depends on diverse identities in diverse contexts. Even though Mann and Eliot wanted to believe that a single writing task measured fitness for college, in truth, engagement in and beyond college depends on diverse writing and ongoing practice.

In addition, as we know, **writing engages diverse knowledge domains**, which all interact in a given moment and context. Cognitive, interpersonal, intrapersonal, and health domains are all part of writing.

Finally, **writing is explicitly learned** and is, in this sense, **unnatural**. Unlike spoken language, which most speakers acquire naturally through interaction, writing must be explicitly taught. It is something we acquire with conscious effort. *Correct writing* can be especially unnatural, because it is less common and more full of dense noun phrases than informal writing.

Closer to the truth is that *correct writing* is no one's mother tongue. It is taught and learned through situation-specific practice, observation, and collaboration, and it responds to context, like all of the writing continuum. Closer to the truth is that no part of the writing continuum is innately *correct* or *incorrect*. No part can stand for the whole.

To illustrate what occurs when writing is naturally 3-D but tested as 2-D, we'll explore examples from the Hillegas scale in Table 3.2.

3.4.3 Hillegas Scale Examples Are on a Narrow Writing Continuum

To add to the writing continuum in this chapter, we'll look at two samples from the Hillegas scale.

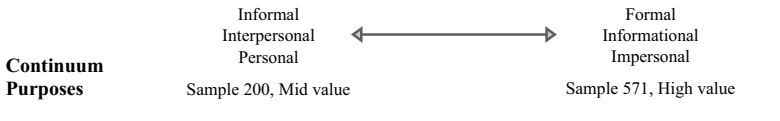
The samples are number 200, ranked as having moderate writing quality, and sample 571, ranked as having high writing quality. Neither sample is described in detail, but the high-value sample is described as written by "a boy in the Freshman class in college."

The mid-value sample is a letter, a bit like a formal email today, and the high-value sample is a description akin to a report or response paper. They represent different parts of the writing continuum, around the formal end of email and the informal end of college writing, respectively. In Table 3.2, each one is summarized in terms of cohesion, connection, focus, stance, and usage. Then, the two samples are represented in full and annotated for language patterns in each one. The patterns range from more interpersonal, personal, and informal in the letter to more informational, impersonal, and formal in the description.

Table 3.2 1912 Hillegas writing scale continuum



1912 Hillegas Writing Scale Continuum Patterns



Continuum Purposes	Sample 200, Mid value	Sample 571, High value
Cohesion	<ul style="list-style-type: none"> • Hourglass organization Clear moves from general greeting and orientation, to specific details, to coda and general closing 	<ul style="list-style-type: none"> • Hourglass organization Clear moves from general topic (<i>the statue, beauty</i>) to specific details (<i>the posture, the limbs</i>) back to general topic (<i>peerless beauty</i>)
Connection	<ul style="list-style-type: none"> • Interpersonal connection Direct address of audience (<i>dear</i>) Text external 1st person emphasizes feelings and experiences (<i>I hope I will</i>) 	<ul style="list-style-type: none"> • Informational connection No direct address Use of 1st person (<i>we</i>) refers to shared experience of statue
Focus	<ul style="list-style-type: none"> • Personal, interpersonal subjects Simple nouns Active verbs 	<ul style="list-style-type: none"> • Informational subjects Nouns and noun phrases focus on statue Passive verbs
Stance	<ul style="list-style-type: none"> • Certain stance Boosters (<i>prettiest, always</i>) show positive stance toward experiences 	<ul style="list-style-type: none"> • Certain stance Boosters show positive stance toward statue (<i>no analysis, peerless</i>)
Usage	<ul style="list-style-type: none"> • Correct writing conventions and usage preferences 	<ul style="list-style-type: none"> • Correct writing conventions and usage preferences
Opening	My dear Fred,--	In looking at this statue we think, not of wisdom, or power, or force, but just of beauty.

In the full samples below, marginal notes and annotations include transitional words **in bold**, connection markers [in brackets], *hedges* in italics, *boosters* and *generalizations* italicized and bolded, and passive verbs [[in double brackets]].

3.4.3.1 Mid-value 1912 Hillegas Scale Sample

[My] *dear* Fred,--

An interpersonal opening:

In the letter greeting and opening sentence, the writer addresses the reader directly

[I] will tell [you] of my journey to Delphi falls, N. Y. There is nice scenery along this route. The *prettiest* scene is in the gulf which is *quite* narrow, a small creek flows down it and the road follows along near its banks.

Interpersonal cohesion:

The writer directly orients the reader to what is to come in the letter

There are woods on either side, these trees look *very* pretty when they are white with snow.

Interpersonal stance and personal focus:

The sentence subjects focus on personal experiences and boosted reactions, with no dense noun phrases

In summer it is *always* shady and cool in them and the small fish may be seen darting back and forth in the water.

Interpersonal audience address:

the writer addresses the reader, expressing a hope to share the experience

[I] hope [I] will have the pleasure of taking [you] over the route some time.

[Yours] sincerely,

Interpersonal closing

3.4.3.2 High-value 1912 Hillegas Scale Sample

In looking at this statue [we] think, not of wisdom, or power, or force, but *just* of beauty.

Informational focus:

Focuses on collective observation

She stands resting the weight of her body on one foot, and advancing the other (left) with knee bent. The posture causes the figure to sway slightly to one side, describing a fine curved line. The lower limbs [are draped] but the upper part of the body is uncovered.

Informational, impersonal cohesion and focus:

The writer moves from the general opening statement to a more specific detail about the statue (posture), using passive verbs to emphasize the statue

The unfortunate loss of the statue's arms prevents a positive knowledge of its original attitude.

Formal focus:

The writer uses a sentence with dense noun phrases to focus on the statue and show a regretful attitude toward the "unfortunate loss of the statue's arms"

The eyes are partly closed, having something of a dreamy languor. The nose is *perfectly* cut, the mouth and chin [are moulded] in adorable curves.

Impersonal stance:

The writer begins to offer evaluation, while the focus of the sentences remains on the statue

Yet to say that *every* feature is of *faultless* perfection is but cold praise. *No analysis* can convey the sense of her *peerless* beauty.

Hourglass cohesion and certain stance:

The writer moves from the specific details back to general statements to close. These statements offer a boosted, certain stance

These samples show how writing across the continuum is both similar and different. Both examples fulfill the five continuum purposes, but each one has some distinct patterns.

The interpersonal letter moves from addressing the recipient, to sharing observations, to closing with a valediction. In so doing, the letter cohesively adds detail. It makes clear the focus of the letter, and it directly and politely addresses the letter reader.

The informational description moves from identifying the topic, to adding and explaining details, back to a general, summative statement. In so doing, the sample adds detail, makes clear the focus of the description and a (positive) stance toward it, and addresses the reader formally, in a shared observation (*we*) and in an evaluation expressed as a shared reaction (*the unfortunate loss...*).

Both the letter and paper share *correct writing* norms for spelling and punctuation. Neither includes flexible usage more common in informal writing, even as the letter is more interpersonal and personal than the description. Though they are ranked differently – with the writing on the left of the continuum receiving a lower rank – both are correct for their context and goals.

Closer to the truth, then, is that only within this myth does it make sense to put these different texts on one scale. Closer to the truth is that writing is not *correct* or *incorrect* on a single scale, but rather a 3-D practice that accounts for its task and context.

Nonetheless, we've inherited 2-D stories about writers and tests, and these continue in our next myth as tests become standardized.