History by the Numbers: A Quantitative Approach to Teaching the Importance of Conflicting Evidence

Peter Burkholder
Fairleigh Dickinson University

THE STRUGGLES between Christians and Muslims in medieval Spain, as recorded in the epic, *The Song of the Cid*, were a function of religion. How do we know? Because there are ample, explicit references in the poem to characters being motivated by religious factors. Then again, this anonymously authored poem also displays considerable evidence of economic motives, and even shows cooperation across religious lines. So, scratch the above—it's clearly monetary intentions undergirding the characters' actions. Or maybe it's some sort of combination of factors. It's confusing.

Niccolò Machiavelli was a Renaissance thinker. How do we know? Because he often drew on writings by classical and contemporary authors in his seminal work, *The Prince*. On the other hand, the frequency with which he invoked examples from different eras, and the ways he used them, vary. But perhaps better not to delve into such things and just stick with the obvious: Machiavelli was a Renaissance thinker, full stop.

The two examples above illustrate the difficulty students encounter when entering into the fog of historical analysis, a place where evidence rarely lines up neatly and contradictions abound. Too

often, novices conveniently ignore any sort of counterevidence that could muddy a clean explanation, thus reverting to safe truisms that sidestep key problems. Meanwhile, professional historians revel in uncertainty, understanding that it is evidentiary ambiguity and conflict that offer opportunities to ponder, to grapple, and to explain. How does one bridge the gap? How can instructors start their classes down that admittedly long, difficult road toward a more genuine understanding of past peoples and events—one that, most importantly, embraces conflicting evidence?

This article reports on a unique approach to these problems, one that forced learners to come face-to-face with contradictory information. Students did not just confront incongruous evidence; they actually measured and weighed it by applying a quantitative method to textual analysis. Their findings, as well as their experiences with the technique, indicate that this unusual research approach has the potential not only to better deal with ambiguous evidence, but also to change learners' perceptions of history itself.

A Problem Worth Exploring

Like so many other pedagogical inquiries, this one was born of a problem.[1] Across a wide variety of assignments, my students had shown a consistent inability to recognize, let alone deal with, the inherent complexities of the past, especially when faced with incongruent or contradictory evidence. Even a cursory perusal of history pedagogical materials indicates this is a major stumbling block to learners' appreciation of the past.[2]

The assignments we devise for history courses play a crucial role—perhaps *the* crucial role—in moving students toward a more sophisticated understanding of the past. This is because, no matter how clever our class activities, no matter how dazzling our multimedia, no matter how riveting our readings and lectures, students primarily learn what they are assessed on.[3] In recent years, the American Historical Association has underscored this basic fact by promoting "assignment charrettes" at its annual meetings: historians share their assignments in small groups, then receive multiple points of feedback from peers. Thereafter, a debriefing session generalizes valuable lessons learned, so that these critical assessments can more fully and effectively achieve their learning goals.[4]

In my case, two texts used in two different mid-level courses had yielded lackluster student papers. *The Song of the Cid*, assigned in a Mediterranean history class, contains rich information on motives of, and relations among, a pluralistic society in medieval Spain. Based on the life of the historical, eleventh-century Rodrigo Díaz, but highly fictionalized in its epic form, this action-packed story, like any piece of great literature, can be read in multiple ways.[5] In the parlance of subject matter experts, we were asking whether the epic suggested a state of *convivencia* (peaceful co-existence between religious groups) or *conveniencia* (cooperation across religions when convenient for personal benefit).[6]

Machiavelli's *The Prince*, a unit reading in a Renaissance and Reformation course, can serve as a gateway into the world of an important early modern thinker. Like *The Song of the Cid* (or really any historical text, for that matter), *The Prince* is marked by complexity and contradictions, both internally and in comparison with secondary readings. Close analysis indicates that Machiavelli vacillates on age-old questions (e.g., Is it better for a leader to be loved or feared?) and, more importantly for our purposes here, he supports his arguments with a hodge-podge of examples drawn from classical antiquity and the contemporary world. For, as Machiavelli states in his opening dedicatory letter, he hopes to educate his reader "through long experience of modern things" as well as "constant reading about ancient things."[7] Our central question was thus how Machiavelli used the past and present, and whether his uses shed light on the substance of the Renaissance itself.

As preparation for understanding these texts, both courses were frontloaded with a historiography unit, at the end of which students wrote papers on disparities between historians' assessments of the medieval Mediterranean or Renaissance context. Editors' introductions to each text and in-class reviews thereof provided more specific information on historical milieux and authors' possible motives for writing. Thereafter, small-group and whole-class discussions of each text emphasized their intricacies. In the case of *The Song of the Cid*, Christian and Muslim (and, to a limited extent, Jewish) characters variously fight and cooperate with one another, both within and across religious lines. Meanwhile, motives vacillate between religious and economic. Over the course of multiple class meetings, student groups were tasked with locating evidence for a

given position in a reading segment, while other groups gathered counterevidence on the same point. After general discussion, the groups would switch roles for a different portion of text, thereby arguing against what they had previously supported. A similar approach with *The Prince* had teams finding support for Machiavelli showing (or not showing) attributes of a Renaissance thinker within the context of earlier secondary readings. For example, Jocelyn Hunt describes the Renaissance as a "golden age" when authors wrote about "classical learning being reborn, while at the same time describing their world as new and revolutionary."[8] Do such attributes apply to chapters of Machiavelli's work? In theory, these approaches should have been conducive to recognizing inconsistencies in the readings.

At least, that was the theory. In practice, the tensions and contradictions that had readily arisen in class discussions rarely materialized in unit papers. Tasked with determining whether the western Mediterranean was, in the estimation of historian Michel Balard, a "battleground between Christians and Muslims" in the era of the Cid,[9] students usually defaulted to one of two camps. The first was the "binary school," where papers either agreed or disagreed with the "battleground" statement, but conveniently ignored the very counterevidence the students had so capably adduced in class discussions. The second approach was one of accommodation, where, like the fabled King Solomon, students split the difference: they recognized the evidence was unsettled with regard to the "battleground" portrayal, but that was as much as they could say. The situation, in these latter students' minds, was a little of this, a little of that. All arguments were assessed as equally valid.

Papers addressing *The Prince* likewise demonstrated a failure to apply findings from class activities. Asked whether Machiavelli showed characteristics of a Renaissance thinker, the vast majority of students pointed to a small, seemingly random selection of the author's references to classical antiquity and modernity to satisfy such a labeling. Whether those references were typical of the whole book, let alone how Machiavelli actually used that material, was beside the point; rather, its very existence satisfied a low-bar criterion for Renaissance thought, thereby obviating any need for closer examination. In short, both paper exercises had turned into cases of confirmation bias: pick a side, marshal a few examples to support it, and pretend any sort of complexity or counterevidence simply does not exist.

Obstacles to Understanding

Why is this type of problem so pervasive in history classrooms? There seems to be a number of interrelated causes. The first is one of conditioning. Over many years, learners have gathered that, in history classes, like in other subjects, there is a "right answer" to any given problem. Such environments are often the result of "coverage" approaches, where the past is presented as clear and objective, and where a learner's goal is to master that past and faithfully reproduce it on command. There are few to no gray areas in these settings, so that students who succeed in them inevitably bring facile skills and flawed expectations to their subsequent classes.[10] None of this is to place the blame on K-12 education. Joel Sipress and David Voelker, building on the work by Lendol Calder, observe that "coverage" is the "signature pedagogy" even at the college level, while ambitious adjustments are being made at the primary and secondary levels.[11] Broader findings about undergraduates' poor abilities in critical thinking are likely, in part, a result of coverage approaches across the entire curriculum at all levels.[12]

A second contributing factor is a matter of intellectual development. Entry-level college students, explain Susan Ambrose and her colleagues, are often "dualists": something is or it isn't, with no room for ambiguity. These learners eventually move on to the stage of "multiplicity": they recognize the existence of different explanations, but all accounts are seen as equally valid. If they progress to the third stage, these people become "relativists": they perceive that not all ideas are legitimate, and that there are ways of teasing out differences and shortcomings. The final stage is when learners reach "commitment": like dualists, they make choices about interpretations, but unlike dualists, those choices are nuanced and substantiated.[13] In my case of the two assignments under examination here, most of the students had been writing at the dualist level, with perhaps a smattering of learners at the multiplicity stage.

A third issue is a function of the field of history itself, and it, too, ties in with the preceding factors. While some fields such as mathematics can yield a single, agreed-upon answer to problems (these fields are called "well-structured"), others such as history do not. Universal consensus and certainty about complex issues of the past are rare. Methods and evidence are still paramount. Nothing

here says there are no such things as facts, or that all explanations are equally valid: historians often proffer contrasting arguments, even while acknowledging common underlying evidence. Thus, ambiguity and complexity are part and parcel of the domain, resulting in history's "ill-structured" field classification.[14] If students approach the past authentically, as I was attempting to have them do with my in-class exercises, they face an unfamiliar and daunting landscape.

It is inherently difficult to move from one stage to the next, to have expectations for an entire domain upended, and to perform the necessary work to achieve a deeper understanding of a historical problem. Some students may reject the entire premise, especially if their extant skills and knowledge have hitherto served them adequately.[15] Making the necessary commitment is a threat to the considerable "sunk costs" already invested in their views and understandings of a simplistic, either/or past. As Jeffrey Burton Russell explained about the pervasive myth of the medieval "flat earth," it is much easier to hold on to faulty ideas than it is to admit error and rearrange basic understanding. Old beliefs die hard.[16]

Rectifying the above deficiencies in one assignment or even a whole course probably is not feasible, as adjustments to learning, skills, and outlooks take place slowly and incrementally. But a task that would force students to come to grips with the intricacies and contradictions of the past could begin them down that path. After all, Angela Duckworth reminds us that "a high level of performance is, in fact, an accretion of mundane acts."[17] Along the way, students would confront "expectation failure" and "learning bottlenecks," points of confusion and inherent difficulty that are absent from coverage-style classes. For my part, I would have to "decode" the process of untangling contested history, and devise "deliberate practice" techniques for the learning to take hold.[18]

Decoding the Conflicting Evidence Problem

We saw above some reasons why students fail to deal with inconsistent evidence effectively. But how do I, as a professional historian, handle such cases? Only by understanding the many, often subconscious steps I perform could I possibly help my learners through this vexing learning bottleneck. Such a metacognitive exercise is indispensable for instructors. Expertise is something to

be celebrated, and we expect it in college teachers. But expertise has a downside, insofar as it creates blind spots to the learning problems of novices. Economists coined this "the curse of knowledge," a state where it is hard to imagine somebody not knowing what you know. And as Steven Pinker reminds us, "The better you know something, the less you remember about how hard it was to learn."[19]

Upon careful reflection—and with the aid of historians David Pace and Lendol Calder in an exercise at the American Historical Association conference in 2016—I realized that when evidence does not line up nicely, I "play" with various possibilities. That is, I can construct, modify, and discard complex scenarios to explain evidentiary situations, and I can do so very rapidly, not worrying about whether my models are workable or even silly. "Play" is not a cognate of "fun" here; rather, it describes the private, low-risk ways I tap into my "mock reader" to interrogate historical information, and then see what I can do with it.[20] This is not anything I was ever taught explicitly, but an ability that developed slowly, over time, as I moved from being a history student to a historian.[21] However useful this heuristic may be, I could not teach it as such to novices, though I had made rough attempts to do so with the small-group debate exercises described earlier.

Another technique was available, one that was more objective, consisted of discrete steps, was amenable to quantification, and would force learners to confront conflicting evidence in their texts. This involved the wholesale coding of specified information from the books, then determining the relative abundance of that information. For students to do this, I would have to create user-friendly coding templates, and carefully guide the classes through the process of filling them out and calculating frequencies.

Moreover, the process would have to be mandatory. That was made clear to me when, years before the events discussed in the present article, I had debuted my quantitative approach to reading *The Prince*. In one class period, I demonstrated how evidence from the text could be broken down and measured, according to prescribed variables. Students seemed to follow along with interest, leading me to believe that some of them would utilize the method in their unit paper. To my disappointment—and, in hindsight, I should not have been surprised—no students did. I performed a straw poll to see why (respondents could offer multiple answers, as applicable), and learned the following: nearly 70% thought the approach would

	Year in College				Major			Gender	
Text/Course	1st	2nd	3rd	4th	History	Science	Other	Male	Female
The Prince (N = 19)	6	4	6	3	6	4*	9**	10	9
The Song of the Cid (N = 20)	5	7	4	4	7	3†	10††	13	7

Figure 1: Demographic breakdown of the two intermediate-level history courses. Includes only students who completed the assignments. *Science majors: biology (3), biochemistry (1). **Other majors: psychology (3), creative writing (2), criminology (2), communication (1), accounting (1). †Science majors: biology (1), chemistry (1), pharmacy (1). ††Other majors: criminology (4), creative writing (2), accounting (1), business (1), political science (1), psychology (1).

be too much work, while 40% simply wanted to avoid mathematics in a history course.[22] In retrospect, I had not offered them the actual tools and practice they really needed to mimic my work. The "curse of knowledge" had gotten the better of me.

Student Demographics and Preconceptions

The sizes and demographics of the two courses used in this study were remarkably consistent. Roughly one-third of each class were history majors; a small proportion were students studying sciences and could be expected to have a facility with numbers; and the remainder (about half) were majoring in neither history nor a science, though some of their fields (e.g., accounting) would require numerical acumen. Each course likewise enrolled a spectrum of first-year students through seniors, with male students outnumbering female students to lesser or greater degrees (see **Figure 1**).

Students' perceptions of, and experiences with, a quantitative approach to history, both before and after the unit, were gathered in various forms. Because the methods used were likely unfamiliar to most learners, because math anxiety is a well-known phenomenon,[23] and because students' math skills may even be faltering in recent years,[24] anonymous surveys were distributed to each class at the beginning of the project units to ascertain attitudes and preconceptions. Everyone in the *Prince* course reported prior training in basic arithmetic and algebra, and all but one had formal

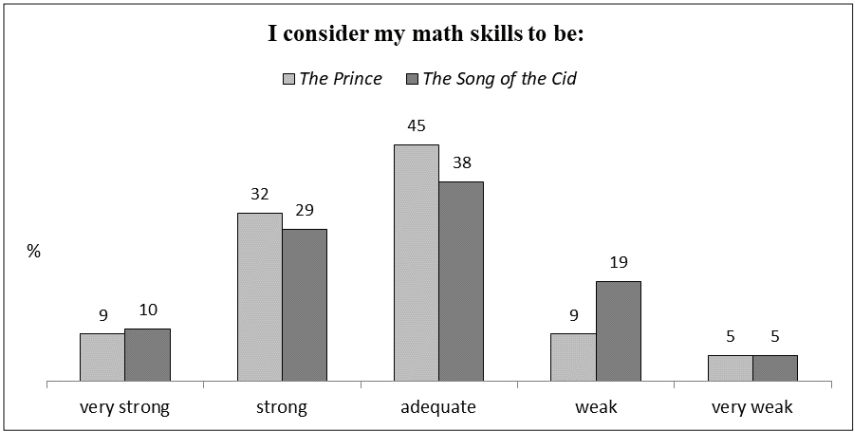

Figure 2: Students' self-assessed math skills, pre-assignment (N = 22 and 21, respectively). An additional response of "no opinion" received no votes.

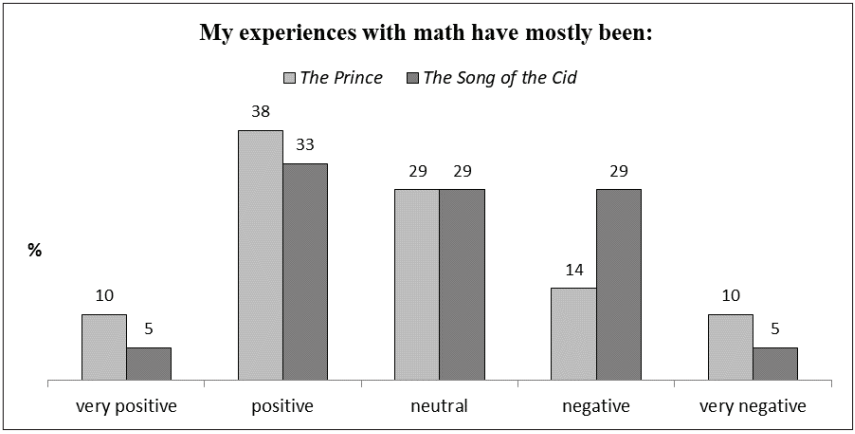

Figure 3: Students' self-reported past experiences with mathematics (N = 22 and 21, respectively). An additional response of "no opinion" received no votes.

experience in geometry. These numbers all skewed lower for the *Cid* course. Nearly half of each class had taken calculus, while about one-third had experience in a pre-calculus or statistics course. Self-assessed strengths in math skills varied somewhat between the two courses, with stronger confidence being registered in the *Prince* course than in the *Cid* course (**Figure 2**). This pattern played out as well when students expressed their past experiences with math, ranging from very positive to very negative (**Figure 3**).

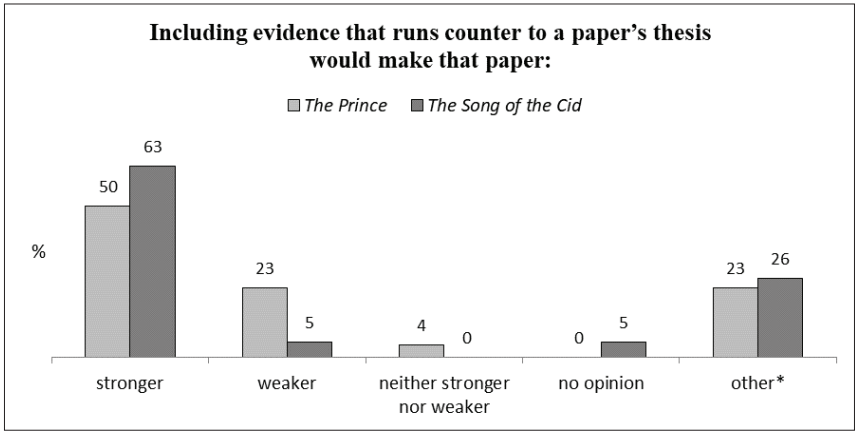

Figure 4: Students' pre-assignment views on the effects of including conflicting evidence (N = 22 and 19, respectively). *All "other" responses included written qualifiers falling under the rubric of "it depends."

Additional questions on the pre-unit surveys pertained to methods for dealing with conflicting historical evidence. When confronted with such evidence, around one-half in each class said they would attempt to determine the relative merits of the inconsistent data, with slightly lower percentages indicating they would attempt to accommodate the conflict somehow. Significantly, none or only a small minority said they would simply ignore the evidence running counter to an argument they were making. On the contrary, half or more of each class felt that including data disagreeing with a paper's thesis would make the end-product *stronger*, while only small minorities thought such inclusion would have a negative impact (**Figure 4**).

Yet, the self-reported information in the preceding paragraph was wholly inconsistent with what I had seen in past work. As indicated earlier, students historically had avoided discussion of conflicting evidence (the "binary" learners), with only a select few acknowledging it (the "relativists"). This points out yet another well-attested metacognitive roadblock: the tendency of learners, especially weaker ones, to overestimate their abilities while underestimating the difficulty of a situation or assignment.[25] Indeed, one recent study ascertained that undergraduates, whether steeped in college-level history courses or not, viewed the assessment of scholars' conflicting

interpretations of the past as a fairly straightforward affair. In that investigation, over 72% of respondents who had taken fewer than three college-level history classes agreed or strongly agreed they were capable of such work; meanwhile, fewer than 2% questioned their abilities. Fifty-seven percent of the same cohort expressed confidence in their ability to extract implicit evidence from primary sources, while a mere 3% showed reservations. The intrinsic difficulty of engaging in historiography and source analysis was lost on most of the respondents.[26] Other research exposes students' inflated views of their own writing abilities. In a 2015 survey commissioned by the American Association of Colleges and Universities, 65% of undergraduate respondents rated themselves as being "well prepared" in writing skills, whereas a mere 27% of employers agreed.[27]

Introduction to the Quantitative Method and Its Rationale

In applying this new quantitative method, some aspects of a more standard approach to textual analysis were certainly retained. The pertinent activities on historiography and authors' possible motives, both mentioned earlier, preceded the methods described below. But because the approach to upcoming analyses of the texts would be a bit unusual this time around, students needed to buy into the importance and utility of the quantitative method. This was facilitated by relating use of evidence, in general, to the students' own experiences. Through a series of questions about a generic assignment, the classes agreed that invoking evidence judiciously would strengthen an analytical paper. They further agreed that the type and quantity of evidence invoked would play a role in determining the merits of a paper's argument. Should this logic pertain to Machiavelli's use of evidence in his persuasive treatise under consideration? Could the nature and weight of evidence in *The Song of the Cid* give insights beyond the epic's storyline? Both course sections collectively agreed to these possibilities.

So far, so good. But at the same time, in considering the evidence presented, was Machiavelli a slave to the ancient past, or did he simultaneously make use of examples from contemporary or near-contemporary times to persuade his audience? There was abundant evidence of the latter, the class agreed. So, did Machiavelli seem to draw more heavily on distant or proximate events and persons in

his treatise? And did he deploy those examples in ways that betray his attitude toward the ancient past and his modern present? Might the answers to these questions shed light on what it meant to be a Renaissance thinker, as defined in the secondary literature already read? Parallel questions were posed to the *Cid* course. In each case, there was general agreement on the merits of such questions, but students were unsure of how one would proceed systematically in this line of inquiry. It was time to introduce a new protocol.

The optimal amount of guidance to provide the students was something I wrestled with. On the one hand, a constructivist, problem-based approach would entail laying out the issue and having the class work towards a solution. On the other hand, I knew I could introduce a protocol that would allow students more time to work with the evidence, and thus opted for the latter. Yet to do so effectively would entail decoding the issue—that is, working through how I, as a professional historian, would go about analyzing and arranging evidence from these two texts in a rigorous fashion. Crucial to this would be devising a systematic organizational structure that would overcome students' tendency to disregard (or simply forget) inconvenient evidence, a tendency that had led to confirmation bias in earlier traditional approaches to analysis. This organizational structure is crucial, since the ways that experts and novices arrange information are quite dissimilar.[28]

Fortunately, a model for the approach already existed in the form of diplomatics. This method, utilized by historians and archivists for centuries, requires researchers to break down sources into their constituent parts, to make reasoned assumptions about their meaning and content, and to code them according to the types of events transpiring in them.[29] Diplomatics has also been used to great effect in history teaching settings.[30] The main difference here was that the coded evidence from the two texts would be quantified.

I introduced mathematical language to the classes with some trepidation. Despite the pre-unit surveys showing familiarity and general comfort with math, the method's symbolic language in the form of variables could spell trouble. Yet the costs would probably be worth it: mathematical symbols have power and efficiency that standard, descriptive prose lacks, and if a goal of the project was to get students seeing the past differently, then some symbolic language could be leveraged to that effect.[31]

The key underlying questions were as follows: First, how can we identify and code historical evidence in these two texts in a concise, accurate, and consistent way? Second, how can we make the data amenable to quantitative analysis? In answer to the first question, I proposed various categories for how Machiavelli brought examples from history into play in *The Prince*, as follows:

Let n = the total number of historical examples found in any given chapter. That being the case:

Let n_a = the number of cases from *classical antiquity* in the chapter, further subdivided as follows:

- Let n_{a+} = the number of *positive* cases from antiquity. In other words, these are examples used to reinforce what a successful ruler *should* do.
- Let n_{a-} = the number of *negative* cases from antiquity, i.e., examples of what a successful ruler should *not* do.
- Let n_{a0} = the number of *neutral* cases from antiquity, i.e., examples that are neither prescriptive nor proscriptive.

A parallel set of variables denotes examples from modern history (i.e., contemporary or near-contemporary with Machiavelli), as follows:

Let n_m = the number of cases from *modern history* in a given chapter, further subdivided as follows:

- Let n_{m+} = the number of *positive* cases from modern history, i.e., examples of what a successful ruler *should* do.
- Let n_{m-} = the number of *negative* cases from modern history, i.e., examples of what a successful ruler should *not* do.
- Let n_{m0} = the number of *neutral* cases from modern history, i.e., examples that are neither prescriptive nor proscriptive.

In all cases, a point of emphasis was that our coding was not driven by presentist attitudes. That is, we were not imposing our own definitions of "positive," "negative," and "neutral" actions, but determining whether Machiavelli thought such actions were prudent or not for rulers in his world. A second important point is that students were not simply looking for specified terms for mechanical coding. Rather, they had to read the whole text carefully, always on the lookout for evidence that fell within the domain of one of the prescribed variables. When such evidence appeared, readers had

to determine which coding category was most appropriate, with the explicit understanding that some examples could not be readily coded, resulting in a "neutral" classification. Such procedures follow well-established qualitative and quantitative research methods.[32]

Once the historical examples used by Machiavelli in a selected chapter are coded according to the variables above, one can determine the relative frequency of the types of evidence appearing. This gives potential insight into how Machiavelli thought of and used the past and present, especially insofar as he drew upon classical and contemporary history to help support his arguments in *The Prince*.

There was no expectation that students could immediately apply these abstractions to the task at hand, so I first modeled a selected chapter for the class. In chapter twelve, "How many kinds of soldiers there are, and concerning mercenary soldiers," Machiavelli invokes fourteen historical examples (this is our n), three of which are from ancient history (n_a), and eleven from modern (n_m). Looking at these more closely, we could see that the cases from antiquity are split between one positive (n_{a+}) and two negative (n_{a-}), whereas the modern examples subdivide into two positive (n_{m+}), six negative (n_{m-}), and three neutral (n_{m0}). In standardized form, the evidence analysis for this chapter breaks down as follows:

n = number of historical examples in chapter twelve = 14

n_a = number of examples from *classical antiquity* = 3 = 21% of all chapter twelve examples

- n_{a+} = number of *positive* examples from classical antiquity = 1 = 33% of n_a
- n_{a-} = number of *negative* examples from classical antiquity = 2 = 67% of n_a
- n_{a0} = number of *neutral* examples from classical antiquity = 0 = 0% of n_a

n_m = number of examples from *modern history* = 11 = 79% of all chapter twelve examples

- n_{m+} = number of *positive* examples from modern history = 2 = 18% of n_m
- n_{m-} = number of *negative* examples from modern history = 6 = 55% of n_m
- n_{m0} = number of *neutral* examples from modern history = 3 = 27% of n_m

For this chapter, a preliminary analysis of the data shows that, for all his knowledge of the distant past, Machiavelli leaned far more heavily on contemporary or near-contemporary examples (79%). More than half of all cases are negative, in that they warn rulers against the use of mercenary troops.

The methods and variables for analyzing *The Song of the Cid* were similar, though in this case, students would be considering two main categories: the nature of relations between Christians and Muslims in medieval Spain; and the motives, be they religious or economic, of the characters in the epic.[33] As such, the variables for coding purposes were as follows:

Let n = number of historical examples in any given stanza. Then:

Let n_m = number of examples showing characters' *motives*, further subdivided as follows:

- n_{me} = number of *economic* motives examples, i.e., where characters perform acts for economic gain.
- n_{mr} = number of *religious* motives examples, i.e., where characters perform acts in furtherance of a religious agenda.

In addition, let n_{ic} = number of *interactions* of Christian actors with fellow Christians. Then:

- n_{ic+} = number of *positive* interactions between Christians.
- n_{ic-} = number of *negative* interactions between Christians.
- n_{ic0} = number of *neutral* interactions between Christians.

Finally, let n_{im} = number of *interactions* of Christian actors with Muslims. Then:

- n_{im+} = number of *positive* interactions with Muslims.
- n_{im-} = number of *negative* interactions with Muslims.
- n_{im0} = number of *neutral* interactions with Muslims.

Because Jews appear only infrequently (though importantly), their motives and interactions would be handled on an ad hoc basis using the same coding principles outlined above.

Just as in the case of Machiavelli's *The Prince*, the variables for *The Song of the Cid* were rolled out slowly, with demonstrations of how they applied to selections of text, followed by students practicing rudimentary coding in class.

Chapter: 12		**Title**: How many kinds of soldiers there are, and concerning mercenary soldiers							**Group(s)**: Professor
Ex.	**Pg(s)**	**Brief Description of Example**	**Classical Antiquity**			**Modernity**			**Coding Justification**
			Pos.	*Neg.*	*Neut.*	*Pos.*	*Neg.*	*Neut.*	
1	76-77	Charles VIII of France invades Italy, easily sweeps aside mercenaries employed there					X		Mercenaries ineffective
2	77	Rome and Sparta used citizen-soldiers	X						Arming own citizens resulted in successes
3	77	Swiss use own soldiers, yet retain freedom				X			Arming own citizens resulted in successes
4	77	Carthage almost destroyed by own mercenaries, even though led by own citizens		X					Why mercenary is untrustworthy
5	77	Philip of Macedon hired as mercenary by Thebes, but turned against employer		X					Why mercenary is untrustworthy
6	78	Francesco Sforza hired as mercenary, turns against employer					X		Why mercenary is untrustworthy
7	78	F. Sforza's father, hired as mercenary, turns against employer					X		Why mercenary is untrustworthy
8	78	Venice's/Florence's use of mercenaries; got lucky: mercenaries defeated (e.g., John Hawkwood) before they could turn on employers						X	Mercenaries successful, but only due to luck
9	78	F. Sforza and Bracceschi (both mercenaries) keep each other in check, so both turn aggressions elsewhere						X	Opposing mercenaries check each other's aggressions
10	78-79	Florence's use of Paolo Vitelli as mercenary; Florence lucky he failed as soldier, or would've been beholden to him						X	Mixed results, but due to luck

Chapter: 12		**Title**: How many kinds of soldiers there are, and concerning mercenary soldiers							**Group(s)**: Professor
Ex.	**Pg(s)**	**Brief Description of Example**	**Classical Antiquity**			**Modernity**			**Coding Justification**
			Pos.	*Neg.*	*Neut.*	*Pos.*	*Neg.*	*Neut.*	
11	79	Venice flourishes as power when employing own people as soldiers				X			Using own people for soldiers successful
12	79	Venice switches to mercenaries; kills own mercenary captain, lest he rule them. Later mercenaries lose battles, undo years' worth of gains					X		Mercenaries untrustworthy, ineffective
13	80	Church's frequent use of mercenaries in wake of HRE's power ebbing; Alberico of Cunio, followed by others (Braccio, Sforza); easily defeated by French, Swiss, Spanish invaders					X		HRE = Holy Roman Empire; mercenaries ineffective
14	80-81	Modern mercenaries are risk-averse, undisciplined					X		Mercenaries as ineffective
		TOTALS	**1**	**2**	**0**	**2**	**6**	**3**	

Figure 5: Example of coding worksheet completed by professor for chapter twelve of Machiavelli's *The Prince*.

Still, it would be difficult for students to organize their coded data based on the variables above. For this reason, a standardized, user-friendly worksheet of my own creation was distributed to students in both hard-copy and electronic form (see **Figure 5** and **Figure 6**). Such a common form would help ensure that students gathered and analyzed their data in a consistent and rigorous fashion. Fill-in-the-blank data analysis sections followed each worksheet, again with the aim of uniform examination. To further illustrate how the worksheets should be used, I provided the classes with worked examples of completed forms for selected segments of text, and had small groups practice analyzing additional passages using

Stanza(s):			**Page(s)**:					**Group(s)**:			
Ex.	**Stanza**	**Brief Description**	**Motives**		**Interactions with fellow Christians**			**Interactions with Muslims**			**Coding Justification**
			Econ.	*Relig.*	*Pos.*	*Neg.*	*Neut.*	*Pos.*	*Neg.*	*Neut.*	
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
		TOTALS									

Figure 6: Blank coding worksheet for a chapter of *The Song of the Cid*. Rows can be added to or deleted from the electronic version, as needed.

prescribed protocols. Informal group presentations of their findings were then compared to ascertain whether the methods made sense, and to debate or troubleshoot any discrepancies that arose. For all of the novelty and complexity of the variables delineated above, students—with the help of their peers—picked up the textual coding and analysis methods remarkably fast. If problems did arise, they mostly pertained to identifying or separating historical examples in the text, as opposed to how to code them.

Going forward, students would work in groups to analyze assigned segments of text, then pool their findings with the rest of the class in order to achieve full coverage of the reading. There were specific reasons for the group approach. First, as a matter of practicality, having individuals or even each group try to code evidence from an entire book would probably be overwhelming.[34] Second, each group had a parallel team that would cover the same

Group	Members	Stanzas, Mtg. 1	Stanzas, Mtg. 2	Stanzas, Mtg. 3
---	Professor	1-9	---	---
A1	Students 1, 2, 3	10-17	40-57	112-122
A2	Students 4, 5, 6	10-17	40-57	112-122
B1	Students 7, 8, 9	18-23	58-75	123-132
B2	Students 10, 11, 12	18-23	58-75	123-132
C1	Students 13, 14, 15	24-30	76-93	133-142
C2	Students 16, 17, 18	24-30	76-93	133-142
D1	Students 19, 20, 21	31-39	94-111	143-152
D2	Students 22, 23, 24	31-39	94-111	143-152

Figure 7: Protocol for dividing the 152 stanzas of *The Song of the Cid* between groups. Each group has a parallel group (e.g., A1 and A2, C1 and C2) with which it confers to ensure consistency of data coding and analysis over a three-day period. A similar protocol was used to divide the twenty-six chapters of Machiavelli's *The Prince*.

portions of text (**Figure 7**). After reaching consensus within a group, the two paired teams could then convene to compare findings and make adjustments, as necessary, to arrive at inter-group unanimity. As such, this was an important component of quality control. Groups were set up by the instructor to ensure a spectrum of history, science, and other majors in each, and to consist of students who had experienced success or struggled on the previous unit's historiography paper.[35]

Results 1: What the Methods Revealed about the Texts

What did students find after putting all this effort into coding and tabulating the evidence in each text? After considerable in-class troubleshooting and then back-reading on my part for consistency and clarity, the reconciled worksheets and data analysis sections were made available electronically to each class via the learning management system, but it was up to individual students to compile the data for an overall understanding of the reading. In each case, the evidence revealed information that had been there all along, but due to its complexity and overwhelming nature, had effectively been hidden from view. Now, trends began to

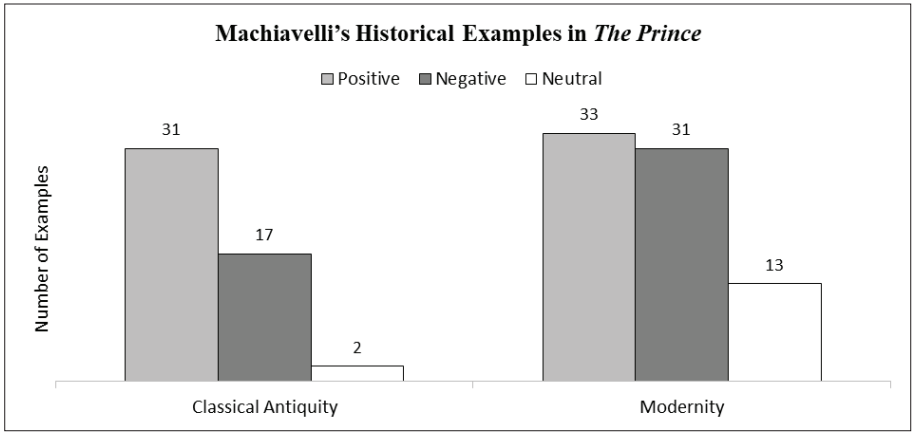

Figure 8: Students' findings on Machiavelli's positive, negative, and neutral uses of examples from classical antiquity and the contemporary world in *The Prince*, all chapters (N = 127).

emerge to highlight and help make sense of conflicting evidence and multiple agendas.

Results pooled from all groups' worksheets for *The Prince* resulted in 127 historical examples across the twenty-six chapters (**Figure 8**). Where did Machiavelli draw his evidence from? Mostly, it was his contemporary or near-contemporary world, with 77 cases (61%) filed in this category, while the remaining 50 cases (39%) derived from classical antiquity. Thus, although Machiavelli clearly fulfilled the Renaissance expectation of turning to the ancients, the majority of his thinking was grounded in events nearer to home—a fact that had never surfaced via more traditional reading methods. Yet, how the writer utilized examples pulled from the distant and more immediate past was noticeably different. The coding of the author's evidence as prescriptive, neutral, or proscriptive showed that classical antiquity served mostly as a source of inspiration for what a leader should do in Machiavelli's own day. Nearly two-thirds (62%) of his cases from classical societies were deemed positive, whereas only a third (34%) served as examples of what to avoid (the remaining 4% were neutral). Things were more evenly split for Machiavelli's contemporary evidence: 43% were exemplars, 40% were cautionary, and 17% were neutral. Does this hint at the writer's nostalgia for the ancient, while signaling ambivalence

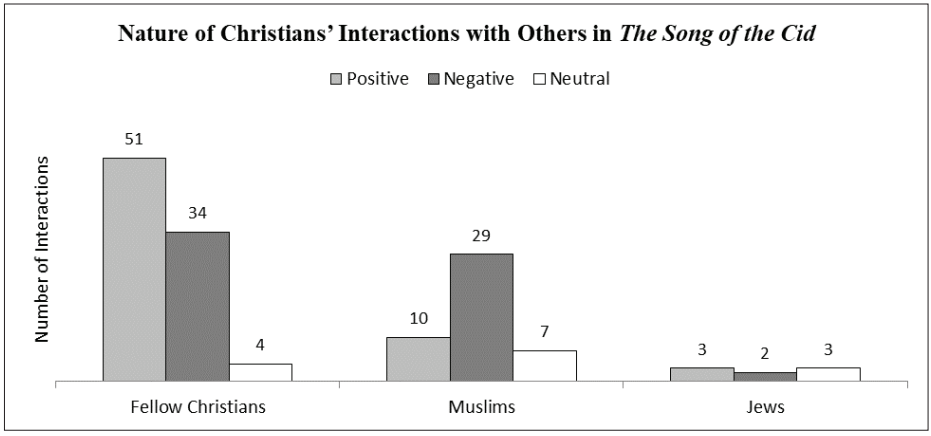

Figure 9: Students' findings on the nature of Christian interactions with other Christians, Muslims, and Jews in *The Song of the Cid*, all stanzas (N = 143).

or even suspicion about the world he lived in? Is it enough for a sixteenth-century writer to make a single reference to antiquity to be labeled a Renaissance thinker? Students now had plentiful, weighted evidence if they wished to pursue such questions.

In the case of *The Song of the Cid*, the class was able to determine characters' motivations, be they religious or economic, in seventy-six instances. Of those, 82% fell under the economic rubric, while only 18% were coded as religious, an important detail that had not arisen from standard reading techniques used in prior iterations. This was perhaps a bit surprising: amidst the Reconquista, where one might reasonably expect strong religious impulses, there is an imbalance of nearly five-to-one favoring economic drivers. On the basis of this evidence, it appeared not to be a case of all explanations for motives being equally valid. If the western Mediterranean was, in fact, a "battleground" between Muslims and Christians at this time, it seemed to be more of an economic zone of conflict than a religious one.

The complex nature of relations within and between religious groups in *The Song of the Cid* likewise emerged from the students' research (**Figure 9**). The coding approach resulted in 143 discrete interactions between characters of three faiths. Christian figures usually (57% of the time) related positively with members of their own religion, but not all such interactions were amicable; 38% of exchanges between Christians were negative, the remainder being

Text	Thesis/ Argument		Methods Discussed			Statistics Used		Secondary Framework	Conflicting Evidence Considered		Graphs Used
	Exists	*Quant.-Based*	*Strong*	*Weak*	*None*	*Local*	*Global*	*Exists*	*Prim.*	*Second.*	*Present*
The Prince	60%	15%	25%	5%	70%	60%	35%	65%	60%	15%	15%
The Song of the Cid	89%	32%	37%	26%	37%	0%	84%	84%	84%	74%	16%

Figure 10: Results from student papers. Students in the *Cid* course (N = 20) clearly performed better, possibly as a result of lessons learned from the earlier *Prince* assignment in a previous course (N = 19).

neutral. The situation was nearly reversed when involving Christian and Muslim characters: 22% of such meetings were positive, whereas 63% were not. In sum, these data seem to point in certain directions, but there is enough counterevidence to suggest that intra- and interfaith relations were varied and complex in the Cid's world.

One should not lose sight of what the classes had accomplished here as a result of the new method. Arthur Marwick refers to the "witting and unwitting testimony" of historical documents, the latter consisting of the unintentional yet very real evidence that "gives historians fascinating insight into the structure, attitudes and life" of past peoples.[36] By systematically coding and tabulating the evidence in each text according to prescribed variables, students had moved well beyond surface reading to unveil embedded evidence, even possible subconscious motives on the part of the authors. Obviously, a quantitative methodology is not the only way to get at such issues, but it was far more conducive to unearthing, let alone weighing, implicit evidence than a more standard approach to these texts had ever been.

Results 2: Student Papers

If coding was an effective way to analyze the historical texts, it could also prove useful for deciphering students' learning as reflected in their papers. There were a number of variables I looked for, a task made easier by using the same paper coding rubric for each assignment (**Figure 10**). Those variables included the following:

- Did the student have a thesis or argument? If so, was it an argument based on the quantified evidence?
- Were the unique methods of the assignment discussed?
- Were statistics used in the paper? If so, were those statistics local in nature (i.e., figures for a limited selection of chapters or stanzas) or global (i.e., combined figures for the whole text)?
- Was work placed within the context of secondary literature?
- Did the student explicitly discuss conflicting evidence? If so, did that evidence point out disagreements with secondary source material and/or illustrate discrepancies within the primary source?
- Were charts or graphs included to help illustrate points?

Because the *Prince* papers were written first, I had no learning artifact data going in and was thus not able to steer those students as effectively toward targeted learning goals as students in the later *Cid* class. In the case of the *Prince* papers, 60% of them made a passable argument; of those, a subset (15% of all papers) had a thesis built around the quantified evidence—that is, they made reference in some way to the weight of computed data undergirding their argument. One such paper argued that Machiavelli "exaggerated [sic] his use of [historical] examples to support his point." The writer continued: "He uses large numbers of positive, classical examples coupled with modern, negative examples in order to express his preference" for Greek and Roman military leaders as exemplars. Students in the *Cid* course had the benefit of more refined training and guidance, resulting in 89% of their papers making an argument, with 32% being quantitative-based. Discussion of unique methods, an explicit requirement of the assignments, also tilted in favor of the *Cid* students: whereas 70% of the *Prince* class omitted any mention of methods, only 37% did so in the *Cid* case. Meanwhile, 37% of *Cid* students had strong methods sections, with only one-quarter doing comparable work in the *Prince* paper.

An ability to harness the power and significance of the quantified evidence also showed a learning curve from the *Prince* case to the *Cid* one—again, a learning curve probably related to my own, in that I was able to avoid pitfalls and give added guidance in the second iteration of this approach. The *Prince* papers showed a hesitancy to utilize global statistics from the entire book (just 35%

did so), while 60% invoked statistics only from selected chapters. In doing so, these latter students showed a tendency to cherry-pick evidence that may not be representative of the whole. Having seen this problem, I was in a better position to train the *Cid* class to use global data more effectively: none of these latter students turned to statistics from only a selection of stanzas, whereas 84% of them placed their arguments and observations within the context of data collected over the entire text.

Did learners turn to secondary materials to help frame their papers? This would seem to be inevitable in the *Cid* case, since the question prompt explicitly invoked one of their assigned secondary readings. Although most (84%) did so, the remainder omitted secondary context. Two-thirds of *Prince* students made connections between primary evidence and secondary materials, despite direct reference to the latter not being part of the assignment prompt (though it featured prominently in class discussions). That said, only 15% of the latter cohort called attention to discrepancies between the *Prince* text and a secondary reading. *Cid* papers were clearly stronger in both regards: 84% recognized conflicts within the text, and 74% remarked how the text diverged from a secondary author. As one student wrote, "[The historian] Balard's statement on the western Mediterranean does not embody the full extent of the [evidence]." She continued, "The situation is too complicated to be summarized [in] a single statement" as Balard had done. Both classes were on the same page when it came to representing their quantified evidence in the form of charts or graphs: despite a crash-course tutorial in how to translate their numerical findings into visuals (which admittedly was not a requirement), only a small minority in either class did so.

Getting students to acknowledge the presence of conflicting evidence, and then attempt to make sense of it, was the primary driver of the assignment, so it merits special attention here. Again, results varied between the two assignments. *Prince* students seemed to view the matter primarily as an internal one: ambiguity and counterevidence, if encountered at all, were to be found within the *Prince* text itself. Thus, 60% pointed out inconsistencies in Machiavelli's invocation of historical cases. Wrote one student, "While Machiavelli does not exclusively use classical examples, he does favor them in several chapters," followed by a breakdown

of classical vs. modern in two chapters. Such observations were typical: select a handful of chapters (perhaps at random) and note the trends. The high-bar goal was whether students saw overall patterns, a goal that only around one-third achieved.

Here, too, *Cid* students performed noticeably better than their *Prince* counterparts. Eighty-four percent pointed out internal inconsistencies in how Christians interacted with co-religionists or those of a different faith, or how religion and economics drove characters' actions. After walking the reader through the data on intra- and interfaith relations, one student concluded that "the data point toward a more complicated relationship between religions of the western Mediterranean." She then posited that, based on the weight of the data, Christians were driven primarily by economic rather than religious motives. Those observations reflected in a fellow student's paper, which stated that the characters in *Cid* "were neither divided by insurmountable religious differences, nor were their interactions completely hostile." Such statements were often girded with statistical evidence drawn from their coded data. Christians' interactions with co-religionists could not be assumed peaceful, wrote one, because "Christians fought against other Christians…with negative interactions occurring with 37% frequency." Similarly, wrote another, one could not presume that all dealings between Christians and Muslims were negative: "While there is tension [between Christians and Muslims], with 63% of encounters between them being negative," he wrote, "not every cross-faith relationship is marked by dispute." A majority of the class (74%) even had the temerity to point out discrepancies between the text and an assessment of a professional historian—a rare occurrence in papers, but justified here, given the weight of evidence.

In such cases—which, class-wide, were the rule rather than the exception in the *Cid* instance—students were not simply disregarding evidence that did not line up as expected, nor did they automatically defer to the assessment of an expert. Instead, they confronted head-on the fact that multiple, even contradictory motives can guide human affairs, and that historians' arguments are prone to amendment. Without the systematic evidence analysis and quantified data staring right at them, such conclusions had proven much more difficult to reach. The significance of this shift in thinking should not be overlooked. Learners, and even the general public, are prone

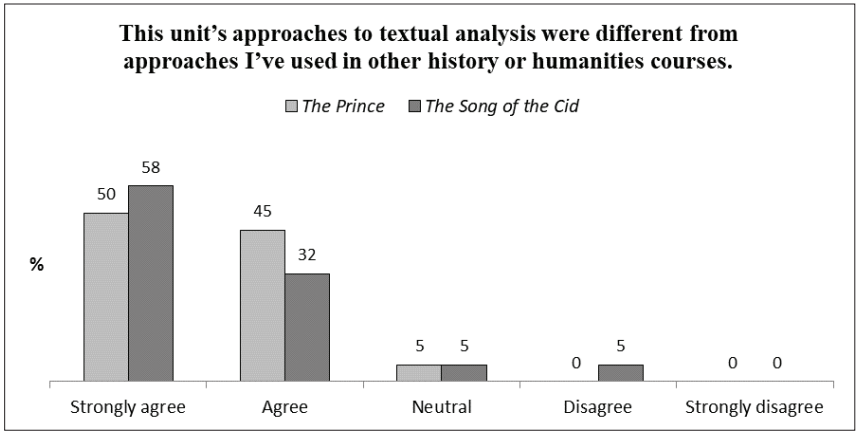

Figure 11: Student perceptions of novelty of quantitative analysis methods (N = 20 and 19, respectively). An additional answer of "no opinion" received no votes.

to seeing "history" merely as a collection of facts to memorize as opposed to an evidence-based explanation of the past.[37] That students felt confident enough to challenge a historian's interpretation speaks to the power of the quantified evidence they had so diligently gathered. It would be rash to see such learners as expert-level, committed thinkers, but they had clearly moved beyond the dualist and even multiplicity stages. By detecting a problem in an earlier explanation of the past, such writers had entered into the realm of relativists: they perceived that not all ideas are equally valid, and that there are means of determining deficiencies.

Results 3: Anonymous Student Surveys and Reflective Essays

If a goal was to shake learners out of their comfort zones and view historical inquiry differently, the assignments succeeded admirably according to anonymous end-of-unit surveys and optional reflective essays (also anonymous).[38] In the *Prince* class, 95% of students agreed or strongly agreed that the methods utilized differed from those typically used in history or humanities courses, while 90% felt likewise in the *Cid* course (**Figure 11**). All anonymous essays from both classes underscored the fact that the quantitative approach to textual analysis was novel. "Normally for papers like this one I would just examine the texts we used and base my ideas on the facts

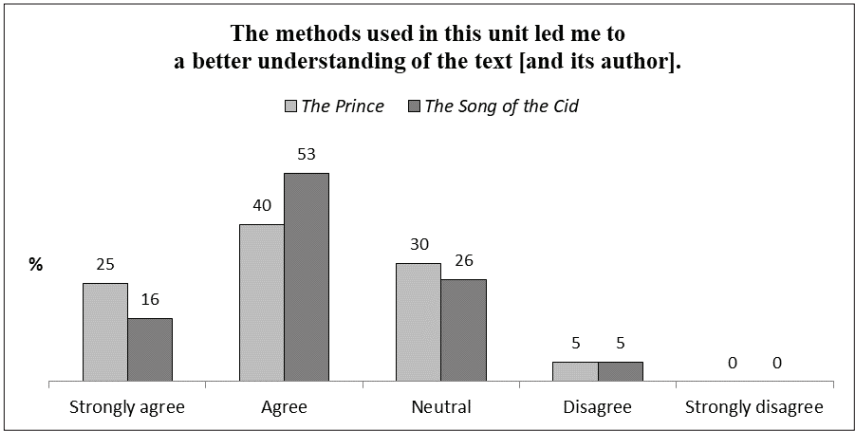

Figure 12: Student perceptions of textual analysis method efficacy (N = 20 and 19, respectively). "Author" was used only in the case of Machiavelli, since *The Song of the Cid* is an anonymous work. An additional response of "no opinion" received no votes.

without considering if there was a pattern of data for the events," wrote a *Cid* student in an anonymous reflection. "I have never had to use data analysis to look at history before."

That students viewed these exercises as departures from other, more typical approaches to analyzing texts may not be surprising, but "different" here was clearly good from a perceived learning standpoint: asked whether the coding and quantification approaches resulted in a better understanding of the respective texts (**Figure 12**), 65% in the *Prince* section and 69% in the *Cid* one agreed or strongly agreed that they had (only 5% in each class disagreed with that prospect). Perceptions diverged somewhat between the two courses when considering whether the method forced learners to grapple with conflicting evidence. Whereas a full 80% in the *Prince* section agreed or strongly agreed that it had, only 58% of *Cid* students felt likewise (**Figure 13**). An anonymous essay from the *Prince* section encapsulates this conflict. "An inherent strength of this approach is that raw numbers are more telling than just reading normally," the student wrote. But there was a perceived downside as well, insofar as "the numbers sometimes do not follow the patterns present in most of the book and can throw off an argument." Such tension likewise appeared in an anonymous *Cid* essay:

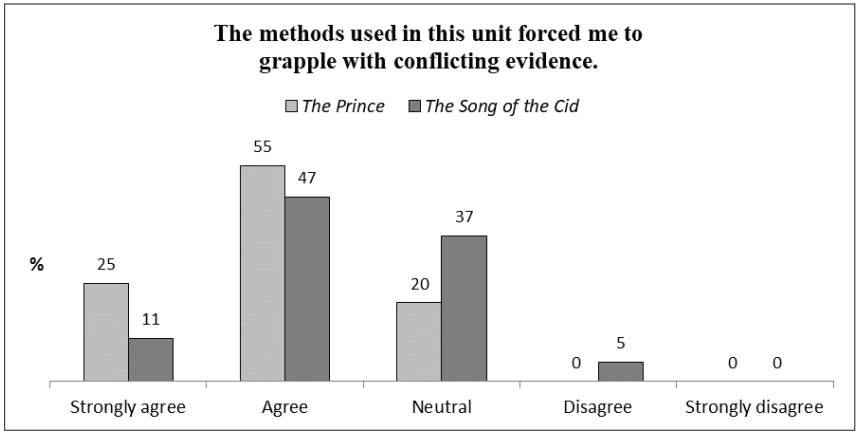

Figure 13: Student perceptions of being forced to deal with conflicting evidence as a function of assignment methods (N = 20 and 19, respectively). An additional response of "no opinion" received no votes.

> I had to alter the way I was looking at my data and approaching my paper so that there would be room to address the conflicting data. I not only had to think about my argument and the data I was using to prove my point, I also had to take into consideration how I would handle the evidence that conflicted with my evidence.

Far from being a detriment, an awareness of such conflicts was a primary driver of the assignment.

Potentially lasting impacts are also seen in statements from anonymous essays. Half of those essays in the *Prince* section and two-thirds of those from the *Cid* cohort expressed that the assignments had changed the way learners viewed the domain of history or the study thereof. A *Prince* essay writer indicated that "[t]his project allowed the college student to step briefly into the real world of what a true historian goes through," a sentiment endorsed by a fellow learner. "We are not just reading the information, but…trying to understand the actual meaning." The writer continued: "This assignment really made me understand how intense the subject of history is and how much more there is to it that I had never realized before." Such observations carried over to the other assignment. A *Cid* student wrote that the quantitative approach caused him/her to see the past "in a much different way," while a classmate admitted to "a much better understanding of historical research." Exposure to this sort of

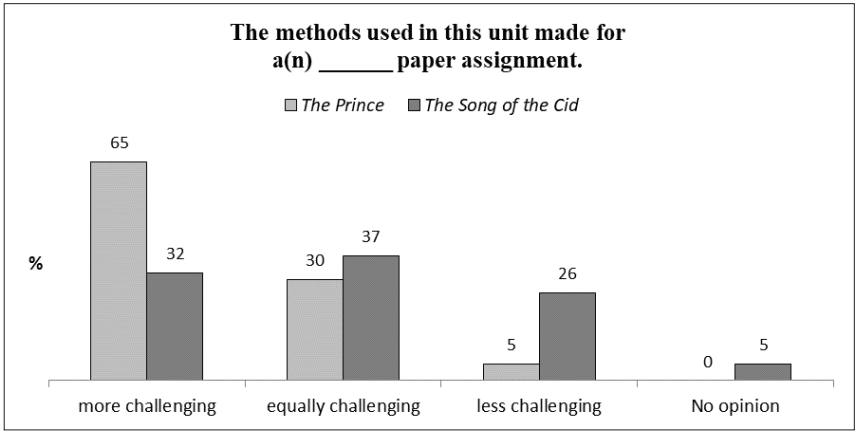

Figure 14: Student perceptions of assignment difficulty as a function of methods (N = 20 and 19, respectively).

work, the hidden "middle stage" of thinking and writing that rarely makes it into the final product, is vital to helping learners understand what lies behind historians' research.

Opinions varied when it came to assessing whether the unique method made for a more challenging assignment: 65% of *Prince* learners thought it did, while only 32% of *Cid* students agreed with that premise. Those views held at the other end of the spectrum as well, where a mere 5% in the *Prince* class felt the approach resulted in a less challenging assignment, a sentiment that a full 26% of *Cid* students shared (**Figure 14**). These sentiments were thus highly variable; moreover, there was no expectation that more or less difficult was somehow preferable. In fact, individual learners could perceive the assignment as being especially challenging yet beneficial, as one anonymous essay from the *Prince* section revealed. "[The method] required a lot more time and effort," s/he wrote, "but with the statistical data my paper [was] much stronger than it would have been without it." Such observations—that the method could be a lot of work but yielded tangible benefits—echo those of students trained in the novel and meticulous techniques of diplomatics in a historical methods course.[39]

Yet, not all steps in the assignments were viewed as equally difficult. According to the post-project surveys, the actual coding of evidence was the easiest step for the *Prince* project, while *Cid* students rated finding patterns in the coded data as the least difficult. Still, individuals

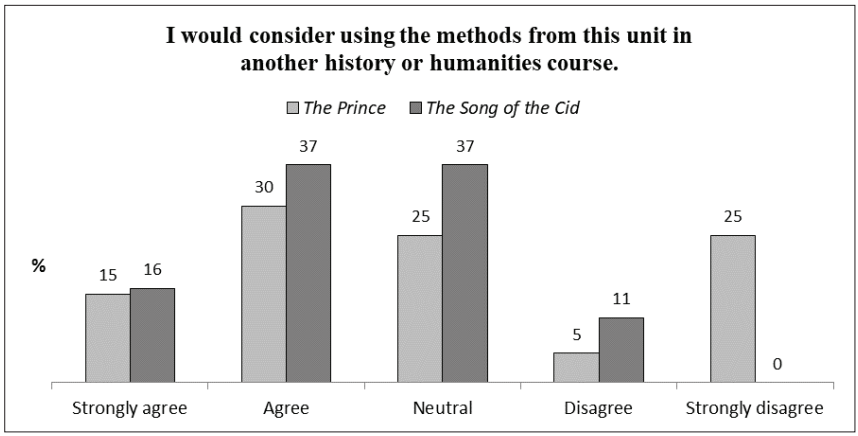

Figure 15: Student views on utility of assignment method in future settings (N = 20 and 19, respectively). An additional response of "no opinion" received no votes.

raised legitimate issues with the coding process. One student in the *Cid* section protested that treating all coded evidence as equal was problematic, because some examples could be deemed more important than others. He thus proposed assigning a multiplier to each piece of evidence to give it its weighted due. Other learners voiced some difficulty in identifying discrete pieces of evidence for examination, or coming to a consensus about positive, negative, or neutral lessons and motives. Intra- and intergroup discussions with peers overcame such obstacles, showing how important the collective approach to textual analysis was, and how conducive the method was in getting students to see and discuss the texts differently. Meanwhile, the two cohorts viewed the problem of dealing with the conflicting evidence that resulted from the coding as challenging, as quotes from anonymous essays above illustrate. Perhaps unsurprisingly, the expression of their ideas in the unit paper was the most demanding task of all.

Given such difficulties, would learners use, or consider using, these unique methods in other settings calling for textual analysis? Tracking students over time for longitudinal impacts is inherently difficult, but two measures give a glimpse of possible prolonged effects. The first simply asked students, by way of an anonymous survey soon after project completion, whether they could foresee using a coding/quantitative approach in another history or humanities course (**Figure 15**). While a slight minority (45%) in the *Prince*

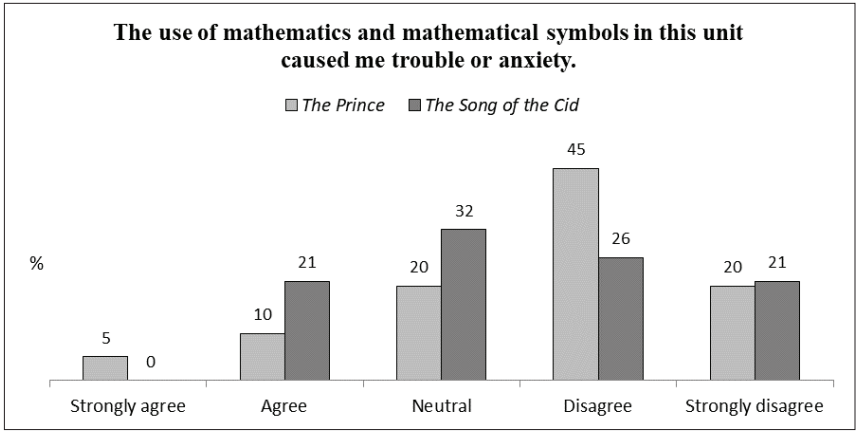

Figure 16: Incidence of math troubles or anxieties triggered by assignments (N = 20 and 19, respectively). An additional response of "no opinion" received no votes.

course saw this as a possibility, a little over half (53%) in the *Cid* cohort thought they would. Another 25% and 37%, respectively, were neutral, but at least held out the possibility.

A second measure of broader learning effects came at the end of the term, when students wrote reflective essays as part of a portfolio assignment on their semester's worth of learning. In particular, learners responded to their own preconception essays written in the first week of the course, and to whether they felt they had begun thinking and writing like historians. Students were not prompted to address the *Prince* or *Cid* projects per se, and the classes were equally removed from having completed these assignments (both had come in the second unit out of four in the courses). In the *Prince* case, 44% of the portfolio essays referred to the assignment; of these, the majority (75%) were positive in nature, the remainder being neutral. All of the portfolio papers in the *Cid* section mentioned the assignment, with 59% being positive, 35% neutral, and 6% negative. It is hard to draw strong conclusions from these data, but the results suggest that learners—especially those in the *Cid* class—did not simply put the exercises behind them upon completion. Some students clearly saw utility in the unique methods employed. Said one science major in the *Cid* course:

> I felt at home when configuring the data for the assignment. Before this, I would argue just by using words to convince the reader that one

point is stronger than another because of the quality of the supporting details. Now, however, I know how to use data to argue my point and I will be using this later on in my future arguments.

Some final good news is that the mathematical approaches caused minimal trouble for learners as a whole. Despite evidence of negative prior experiences with math (see **Figure 3** above), the majority of students in each class registered no math anxiety being triggered (**Figure 16**), suggesting that symbolic language and basic arithmetic were not limiting factors in these assignments.

Results 4: Perceived Difficulty and Learning in Broader Perspective

How did students perceive these quantitative projects within the context of all papers written for the courses? In addition to some smaller assignments, each class wrote four unit papers, three of which were more traditional in nature. At the end of the semester, students were asked to rank all four paper assignments according to relative difficulty and perceived learning. The results here differed by course. For students reading *The Prince*, the quantitative nature of the assignment made it the most difficult, and roughly on par with the final paper of the semester. From these students' standpoint, however, the challenge of the assignment did not translate to greater learning: the Machiavelli paper stood as the assignment they felt they learned the least from, while the third paper, a combined historiography and primary source assignment following more traditional contours, was considered the most beneficial.

By the same criteria, the experiences of the *Cid* students were markedly different. That class saw the quantitative assignment as the least difficult of the four papers, while also ranking it—by far—as the assignment they learned the most from. Why such drastically different perceptions from the two cohorts using similar methods to examine sources of the past? And in the present context: How does one explain this contradictory data?

Here, one can only speculate, as the results may be idiosyncratic, a function of the particular cohorts.[40] Certainly, the fact that students in the *Cid* class benefited from lessons learned in the earlier *Prince* assignment cannot be discounted. That is no small matter in unusual assignments like these. But the nature of the texts themselves may hold

powerful clues as well. Machiavelli's *The Prince* is arguably the more challenging read. As a how-to manual for Renaissance rulers, there is no story arc, no central characters to relate to, and more abstract links between the text and previously read secondary materials. Meanwhile, *The Song of the Cid*, as a great medieval epic, is a riveting, action-packed story with a well-defined *dramatis personae* of heroes and villains. Simply put, the latter reading more readily fulfills the narrative expectations students harbor for historical texts, while the connections with secondary readings from earlier in the course are more immediate. Their seeing it as a more user-friendly source, relatively speaking, is probably inevitable, regardless of the methods used to examine it.[41]

One should always be on guard against reading too much into students' self-assessments. The survey responses and anonymous essays from the classes constitute *indirect* measures of learning, and researchers know that students, especially weaker ones, can be very poor at assessing their own progress.[42] This is akin to the notorious student evaluations of teaching, which have been shown to have little, no, or even inverse correlation with direct measures of actual learning.[43] Yet, if changing hearts and minds is at issue—if helping students see the merits of a rigorous yet unusual approach to historical analysis is at stake—then perceptions are certainly part of the overall equation. In that context, the fact that both classes clearly viewed the quantitative approaches as helping gain a better understanding of the texts, relative to traditional analysis methods, should not be overlooked (see **Figure 12** above).

Conclusion

In his recent book, *The Reading Mind*, cognitive psychologist Daniel Willingham throws cold water on the idea that our brains have been altered by digital technologies to avoid sustained reading efforts. That said, Willingham does suspect that readers' expectations for texts have changed. "It's impatience with boredom," he writes. "[W]hatever experience the technology offers, you get it immediately…Furthermore, producing this experience costs you very little—minimal effort."[44] If that is true—and Willingham admits he cannot prove his stance empirically—then the reading methods described in this article go against the grain of what students now do naturally with texts. Instead of an immediate ability to glean

significant information from complicated readings, learners must slog through each and every word of a text, carefully scrutinizing and coding information to reveal patterns that may challenge preliminary assumptions. Having perhaps had their preconceptions about a text challenged, and then having gone through the difficult mental gymnastics of rearranging their beliefs, students then must figure out how to make sense of their data, knowing full well that discordant evidence is staring them in the face. This is perhaps the antithesis to instantaneous gratification, such work being impossible without great investments of time, diligence, and effort.

Rather, it is all a very unnatural process for most learners, and that is precisely the point: to shake them out of a state of complacency, to cast them into the roiling waters of historical ambiguity, to have them see for themselves that the past is a complicated arena where humans often do not act according to present-day expectations. For many students, this is all unknown and potentially frightening territory for which they must be thrown a lifeline. And that is why exercises like this are all the more needed to help transition dualist thinkers to relativists, moving them down that long path toward committed thinking. The process is not easy either for the learner (who can view it as tedious and needlessly complicated) or teacher (who must deal with a host of new instructional, logistical, and quality control issues). Yet the payoff is in the final product: papers that had been devoid of complexity, that had accepted expert assessments as holy writ, that had conveniently ignored the controversy of inconsistent evidence—these did not altogether vanish, but they were fewer, and the nature of most had improved.

Change is never easy, no less so in the case of transforming students' often simplistic notions of the past into understandings more closely approximating those of their teachers. Perhaps Niccolò Machiavelli should have the last word here, his observations on the necessity of confronting change ringing as true today as they did 500 years ago:

> [O]ne who governs himself with caution and patience, if the times and his circumstances run in such a way that this course of action is good, becomes happy. But if the times and his circumstances change, he is ruined, because he does not change his way of proceeding… because when a man has always prospered by walking in one path, he cannot be persuaded to depart from it. For this reason the cautious man, when it is time to be impetuous, does not know how to do it, whence he is ruined.[45]

Notes

I would like to offer my great thanks to David Pace, Maryellen Weimer, and the anonymous referees for this journal for generously reading and commenting on drafts of this article.

1. On this common impetus, see Randy Bass, "The Scholarship of Teaching: What's the Problem?" *Inventio* 1, no. 1 (February 1999), <https://my.vanderbilt.edu/sotl/files/2013/08/Bass-Problem1.pdf>.

2. See for example American Historical Association, "AHA History Tuning Project: 2016 History Discipline Core," <https://www.historians.org/teaching-and-learning/tuning-the-history-discipline/2016-history-discipline-core>; and the collection of Tuning essays by Daniel McInerney, Sarah Shurts, Nancy Quam-Wickham, Elaine Carey and Tracey-Anne Cooper, and Andrew Stuart Bergerson and Nathan Lindsay with Leah K. Gensheimer and Dan Stroud, in *The History Teacher* 49, no. 4 (August 2016), <https://www.thehistoryteacher.org/tuning>; Sam Wineburg, *Why Learn History (When It's Already on Your Phone)* (Chicago, IL: The University of Chicago Press, 2018), chs. 4-5; Lendol Calder and Tracy Steffes, "Measuring College Learning in History," in *Improving Quality in Higher Education: Learning Outcomes and Assessments for the 21st Century*, ed. Richard Arum, Josipa Roksa, and Amanda Cook (San Francisco, CA: Jossey-Bass, 2016), 37-86.

3. Maryellen Weimer, *Learner-Centered Teaching: Five Key Changes to Practice* (San Francisco, CA: Jossey-Bass, 2002), 16.

4. See Nancy Quam-Wickham, "Peer Reviewing History Assignments at the AHA Teaching Workshop," *AHA Today*, October 13, 2016, <http://blog.historians.org/2016/10/peer-reviewing-history-assignments-aha17-undergraduate-teaching-workshop/>. I was a facilitator to these sessions at AHA meetings in 2016 and 2017.

5. *The Song of the Cid (A Dual-Language Edition with Parallel Text)*, trans. Burton Raffel (New York: Penguin, 2009).

6. See for example Brian A. Catlos, *Kingdoms of Faith: A New History of Islamic Spain* (New York: Basic Books, 2018), 428-429, 445.

7. Niccolò Machiavelli, *The Prince (with Related Documents)*, ed. and trans. William J. Connell (Boston, MA: Bedford/St. Martin's, 2005), 40. A second edition of Connell's translation was published in 2016, but all references in this article are to the first edition.

8. Jocelyn Hunt, *The Renaissance* (New York: Routledge 1999), 6.

9. Michel Balard, "A Christian Mediterranean: 1000-1500," in *The Mediterranean in History*, ed. David Abulafia (Los Angeles, CA: J. Paul Getty Museum, 2003), 183-217 at 216.

10. On disparities between students' and professors' readings of historical texts, see Peter Burkholder, "Teaching Historical Literacy within a SOTL Framework," *Teaching History: A Journal of Methods* 44, no. 2 (Fall 2019): 44-50.

11. Joel M. Sipress and David J. Voelker. "From Learning History to Doing History: Beyond the Coverage Model," in *Exploring Signature Pedagogies:*

Approaches to Teaching Disciplinary Habits of Mind, ed. Regan A. R. Gurung, Nancy L. Chick, and Aeron Haynie (Sterling, VA: Stylus, 2009), 19-35; Lendol Calder, "Uncoverage: Toward a Signature Pedagogy for the History Survey," *The Journal of American History* 92, no. 4 (March 2006): 1358-1370; Susannah Walker and Gustavo Carrera, "Developing A Signature Pedagogy for the High School U.S. History Survey: A Case Study," *The History Teacher* 51, no. 1 (November 2017): 65-88.

12. Richard Arum and Josipa Roksa, *Academically Adrift: Limited Learning on College Campuses* (Chicago, IL: The University of Chicago Press, 2011), 35-36. A more forceful indictment of the utility of college as a whole is Bryan Caplan, *The Case Against Education: Why the Education System Is a Waste of Time and Money* (Princeton, NJ: Princeton University Press, 2018).

13. Susan A. Ambrose, Michael W. Bridges, Michele DiPietro, Marsha C. Lovett, and Marie K. Norman, *How Learning Works: Seven Research-Based Principles for Smart Teaching* (San Francisco, CA: Jossey-Bass, 2010), 163-165.

14. James Voss and Jennifer Wiley, "Expertise in History," in *The Cambridge Handbook of Expertise and Expert Performance*, ed. K. Anders Ericsson, Neil Charness, Robert R. Hoffman, and Paul J. Feltovich (Cambridge, United Kingdom: Cambridge University Press, 2006), 569-584.

15. See Ambrose, Bridges, DiPietro, Lovett, and Norman, *How Learning Works*, 23-24, on how prior knowledge and expectations can have negative consequences.

16. Jeffrey Burton Russell, *Inventing the Flat Earth: Columbus and Modern Historians* (Westport, CT: Praeger, 1991), esp. 77. More recently, see Pete Burkholder, "What You Know That Just Ain't So," *The Teaching Professor*, April 6, 2020, <https://www.teachingprofessor.com/topics/student-learning/what-you-know-that-just-aint-so/>.

17. Angela Duckworth, *Grit: The Power of Passion and Perseverance* (New York: Scribner, 2016), 38.

18. On "expectation failure," Ken Bain, *What the Best College Teachers Do* (Cambridge, MA: Harvard University Press, 2004), 28; on "learning bottlenecks" and "decoding," David Pace, *The Decoding the Disciplines Paradigm: Seven Steps to Increased Student Learning* (Bloomington, IN: Indiana University Press, 2017); on "deliberate practice," see Anders Ericsson and Robert Pool, *Peak: Secrets from the New Science of Expertise* (Boston, MA: Houghton Mifflin Harcourt, 2016), 97-100.

19. Steven Pinker, *The Sense of Style: The Thinking Person's Guide to Writing in the 21st Century* (New York: Viking, 2014), 59.

20. For the "mock reader," Samuel S. Wineburg, *Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching the Past* (Philadelphia, PA: Temple University Press, 2001), 70.

21. See Leah Shopkow's similar experience learning the art of citing sources: "How Many Sources Do I Need?" *The History Teacher* 50, no. 2 (February 2017): 169-200 at 170-172.

22. "Work-avoidance goals" are well attested; see Ambrose, Bridges, DiPietro, Lovett, and Norman, *How Learning Works*, 72-73.

23. Erin A. Maloney and Sian L. Beilock, "Math Anxiety: Who Has It, Why It Develops, and How to Guard Against It," *Trends in Cognitive Sciences* 16, no. 8 (August 2012): 404-406.

24. Kate Zernike, "Test Scores Show a Decline in Math Among High School Seniors," *The New York Times*, April 27, 2016, <https://www.nytimes.com/2016/04/27/us/math-test-scores-decline-high-school-seniors.html>.

25. See note 42 below.

26. Pete Burkholder, "Metacognitive Roadblocks: How Students' Perceived Knowledge and Abilities May Hinder Performance in Undergraduate History Courses" (May 2015).

27. Hart Research Associates, "Falling Short? College Learning and Career Success," 12, <https://www.aacu.org/sites/default/files/files/LEAP/2015employerstudentsurvey.pdf>.

28. Ambrose, Bridges, DiPietro, Lovett, and Norman, *How Learning Works*, 49-54.

29. Leonard Boyle, "Diplomatics," in *Medieval Studies: An Introduction*, ed. James M. Powell, second ed. (Syracuse, NY: Syracuse University Press, 1992), 82-113. On the utility of diplomatics to archivists, Heather MacNeil, *Trusting Records: Legal, Historical, and Diplomatic Perspectives* (Dordrecht, Netherlands: Kluwer Academic Publishers, 2000), 20-24, 86-112.

30. Peter Burkholder, "Getting Medieval on American History Research: A Method to Help Students Think Historically," *The History Teacher* 43, no. 4 (August 2010): 545-562.

31. As my departmental colleague writes, "[N]atural language has the advantage of familiarity but the disadvantage of wordiness," while mathematical symbols carry "the advantage of being compact but the disadvantage of esoteric abstraction." John W. Schiemann, *Does Torture Work?* (New York: Oxford University Press, 2016), 121.

32. Carol A. B. Warren and Tracy Xavia Karner, *Discovering Qualitative Methods: Ethnography, Interviews, Documents, and Images*, third ed. (New York: Oxford University Press, 2015), 210-215. On the symbiotic relationship between qualitative and quantitative methods, Patrik Aspers and Ugo Corte, "What Is Qualitative in Qualitative Research," *Qualitative Sociology* 42, no. 2 (June 2019): 139-160 at 156.

33. This is not meant to be an exhaustive list of themes that one could code and quantify, but a sufficient selection of important subjects for instructional purposes. For example, "honor" motives also permeate *The Song of the Cid*. At the same time, one must be on guard against seeing patterns in things that are not really there, akin to John Nash in the film, *A Beautiful Mind* (dir. Ron Howard, 2001).

34. Nevertheless, familiarity with the entire text, not just a group's assigned sections, was necessary. Daily annotation checks ensured that students had read the entire book.

35. Each course also enrolled a small number of non-credit-seeking students, all of whom held undergraduate degrees (or higher) and were especially engaged and helpful in the projects. Although they did not write papers or partake in surveys, these students were embedded in groups of traditional undergraduates.

36. Arthur Marwick, *The New Nature of History: Knowledge, Evidence, Language* (London, United Kingdom: Red Globe Press, 2001), 172-173.

37. On this tendency, see Norman J. Wilson, *History in Crisis? Recent Directions in Historiography*, second ed. (Upper Saddle River, NJ: Pearson Prentice Hall, 2005), 1-3. Actual data to this effect were recently gathered through statewide polling (N = 801); see Peter Burkholder and Krista Jenkins, "What Are Our Fields About? Survey Suggests Disconnect between Professionals and the Public," *The Teaching Professor*, November 18, 2019, <https://www.teachingprofessor.com/topics/student-learning/what-are-our-fields-about/>.

38. Students in each section were given the option of writing short, anonymous essays about their experiences with their respective projects for token extra credit. In the *Prince* section, ten students (53% of the class) wrote such essays, whereas six (30%) did in the *Cid* section.

39. Burkholder, "Getting Medieval on American History Research."

40. See Maryellen Weimer, "Will the Same Approach Get You the Same Results?" *The Teaching Professor*, July 1, 2019, <https://www.teachingprofessor.com/topics/for-those-who-teach/will-the-same-approach-get-you-the-same-results/>.

41. In fact, *The Song of the Cid*, similar to the medieval French epic *The Song of Roland*, plays out very much like a Hollywood movie; see Peter Burkholder, "How to Read a Historical Film," *World History Connected* 16, no. 2 (June 2019): 1-14, <https://worldhistoryconnected.press.uillinois.edu/16.2/forum_burkholder.html>.

42. Donald R. Bacon, "Reporting Actual and Perceived Student Learning in Education Research," *Journal of Marketing Education* 38, no. 1 (April 2016): 3-6; Susan A. Ambrose and Marsha C. Lovett, "Prior Knowledge Is More than Content: Skills and Beliefs Also Impact Learning," in *Applying Science of Learning in Education: Infusing Psychological Science into the Curriculum*, ed. Victor A. Benassi, Catherine E. Overson, and Christopher M. Hakala (Washington, DC: Society for the Teaching of Psychology, 2014), 7-19 at 15-16; John Girash, "Metacognition and Instruction," in *Applying Science of Learning in Education*, 152-168.

43. There is extensive bibliography on this point; for a recent overview, see Bob Uttl, Carmela A. White, and Daniela Wong Gonzalez, "Meta-Analysis of Faculty's Teaching Effectiveness: Student Evaluation of Teaching Ratings and Student Learning Are Not Related," *Studies in Educational Evaluation* 54 (September 2017): 22-42. A more reader-friendly summary of the topic is Jason Brennan and Phillip W. Magness, *Cracks in the Ivory Tower: The Moral Mess of Higher Education* (New York: Oxford University Press, 2019), ch. 4: "On Reading Entrails and Student Evaluations."

44. Daniel T. Willingham, *The Reading Mind: A Cognitive Approach to Understanding How the Mind Reads* (San Francisco, CA: Jossey-Bass, 2017), 173.

45. Machiavelli, *The Prince*, 118.