TC: A Journal of Biblical Textual Criticism 29 (2024): 183–204

Data Democratization and Biblical Manuscript Studies: A Caution for the Age of Access

W. Andrew Smith, Shepherds Theological Seminary

Abstract: As manuscript transcriptions, images, and metadata become more widely available to the public through digital media, (nearly) gone are the days when a Richard Bentley would have to send a Johann Wettstein to a far-off library to acquire collations for his research. The modern Bentley can sit at his or her laptop and immediately access much of what was unavailable prior to the 1990s. While this remains a remarkable boon to scholarship and draws back the veil for the nonexpert, this paper explores the problems attending data democratization at a time when authority and expertise are devalued.

1. Introduction

Scholars in biblical studies, textual criticism, and history have certainly benefited from the wealth of biblical manuscript images and metadata that are being made increasingly available online. Access to these data not only benefits research in the traditional sense but also opens opportunities for unexpected discovery.[1] However, there is the risk that the same images and data that aid knowledge production for experts is also leading to error production from inexpert analysis, which may be quickly and widely disseminated through online platforms. The purpose of this paper is to explore an unintended consequence of making these data openly available and to offer advice to scholars who can mitigate the damage of counter-knowledge.

1. At a recent panel discussion at Rochester Institute of Technology, Will Noel (Special Collections Librarian at Princeton University) related that making digital images of curated objects available online has resulted in public user comments that have revealed surprising information on the history or meaning of those objects; in his words, "serendipity is now a huge part of discovery" (Will Noel, "On Promiscuous Data" [presentation, Promiscuous Data: Museum, Archive, and Library Collections and the Digital Age, Rochester, NY, 19 October 2023]).

In their 1991 introduction to New Testament textual criticism, León Vaganay and Christian-Bernard Amphoux provide a helpful four-era model for the periods of modern textual criticism, identifying them as: "the Rise of the 'Textus Receptus' (1514–1633)," "the Reign of the 'Textus Receptus' (1633–1831)," "the Fall of the 'Textus Receptus' (1831–1934)," and "the Era of Documentation (1935–1990)."[2] To this, I would add a fifth period extending through today: the Age of Access. With the creation of the World Wide Web (WWW) by Tim Berners-Lee at CERN in 1989 and its public launch in 1993, this ideal platform for both publicly and privately sharing images and other data has transformed how we access information.

In the past, scholarly research of biblical manuscripts was limited by distance, time, and access. While there are certainly exceptions, gone are the days when a Richard Bentley must send a J. J. Wettstein to a distant library to collate a well-known manuscript for study. Today the modern Bentley can access many manuscripts images, transcriptions, and collations from his or her laptop. In 2012, Keith Elliott noted the importance of digitization for text-critical studies and referred to it as the "democratization of scholarship," commenting that it has led "to a genuine transparency and sharing of knowledge, plans and resources."[3] Paywalls restrict access to some useful text-critical resources—and certainly printed materials continue to be expensive—but the field of biblical manuscript studies has taken great steps toward data democratization. Futurist Bernard Marr provides a helpful definition of data democratization:

> Data democratization means that everybody has access to data and there are no gatekeepers that create a bottleneck at the gateway to the data. It requires that we accompany the access with an easy way for people to understand the data so that they can use it to expedite decision-making and uncover opportunities for an organization. The goal is to have anybody use data at any time to make decisions with no barriers to access or understanding.[4]

2. Léon Vaganay and Christian-Bernard Amphoux, *An Introduction to New Testament Textual Criticism* (Cambridge: Cambridge University Press, 1991), 129–71.

3. J. K. Elliott, "Recent Trends in the Textual Criticism of the New Testament: A New Millennium, a New Beginning?," *BABELAO* 1 (2012): 125.

4. Bernard Marr, "What Is Data Democratization? A Super Simple Explanation and the Key Pros and Cons," *Forbes* (24 July 2017): https://www.forbes.com/sites/bernardmarr/2017/07/24/what-is-data-democratization-a-super-simple-explanation-and-the-key-pros-and-cons/.

Making digital representations of biblical manuscript data freely available online has transitioned *material religion* into *digital religion* and has generated a stewardship responsibility of digital curation as well as educational opportunities in digital humanities.[5] Numerous institutions have been funded to make their resources available online, with a wealth of material emerging from providers such as the Digital Vatican Library, the British Library, the Bibliothèque nationale de France, the Dead Sea Scrolls Digital Library, and others.[6] Important centers have made images, bibliographies, manuscript transcriptions, and other analyses widely available as well. Responsible use of this remediated manuscript data can and has already led to great advances in our knowledge of the history, reception, and transmission of the Bible and extrabiblical works.

However, irresponsible use and misinterpretation of these data can now proliferate in ways never dreamed of prior to the age of the internet. While the maxim of "if it's on the internet it must be true" is clearly fallacious, an impressive website, ephemeral credentials, and well-packaged data can appear very convincing to the average reader. As Daniel Levitin comments:

> The promise of the Internet is that it is a great democratizing force, allowing everyone to express their opinions, and everyone to have immediate access to all the world's information. Combine these two, as the Internet and social media do, and you have a virtual world of information and misinformation cohabiting side by side, staring back at you like identical twins, one who will help you and the other who will hurt you. Figuring out which one to choose falls upon all of us, and it requires careful thinking and one thing that most of us feel is in short supply: time. Critical thinking is not something you do once with an issue and then drop it. It's an active

5. Arjun Sabharwal, "Digital Scriptures, Material Religion, and the Digital Humanities: An Interdisciplinary Framework for Curating Digitized Sacred Texts Online," in *Digital Humanities and Material Religion*, ed. Emily Suzanne Clark and Rachel McBride Lindsey, Introduction to Digital Humanities—Religion 6 (Berlin: de Gruyter, 2022), 53–68.

6. http://digital.vatlib.it/, https://www.bl.uk/manuscripts/ (currently disabled for over a year now due to a devastating cyber-attack), https://archivesetmanuscrits.bnf.fr/, and http://www.deadseascrolls.org.il/explore-the-archive, respectively.

> and ongoing process. It requires that we all think like Bayesians, updating our knowledge as new information comes in.[7]

Levitin then notes that "it's difficult to trust your own knowledge if half of it turns out to be counterknowledge."[8]

Certainly, this is not an issue specific to biblical manuscript studies or textual criticism. Yet this field has certain vulnerabilities:

1. Regarding data acquisition, no single scholar can process all the manuscript data, and thus we must rely on the work of others in the field to continue construction of the edifice of text-critical study. When the field was limited to published works and papers in a relatively well-known group of contributors, go-to resources were easy to locate, request, and confirm. Now, with a flood of voices contributing data—whether vetted or not—to a global, digital conversation, resources must be subjected to more rigorous scrutiny.
2. As researchers and educators, we wish to communicate our findings to the public, to peers, and to potential patrons and/or collaborators. Removing obstacles to that communication, especially at a time when internet media outputs can easily drown out the voices of legitimate scholarship, can only help us meet that end.
3. Anyone with an internet connection feels qualified to evaluate data for themselves, whether they have any domain experience or not. The high amount of *misinformation* for our field is only compounded by the high volume and accessibility of that misinformation.

2. Misunderstanding Data

One concern about data democratization in business is that data must be understood to be used correctly. Marr notes, "There is still concern by some organizations that misinterpretation of the data by non-technical employees could occur and these employees would then make bad decisions based on their bad interpretation of the data. In addition, the more

7. Daniel J. Levitin, *Weaponized Lies: How to Think Critically in the Post-truth Era* (Toronto: Penguin, 2017), n.p.

8. Levitin, *Weaponized Lies*.

users who have access to the data the bigger the data security risk and the more challenges to maintaining the data integrity."[9] In the business world, data integrity is an issue if users with free access to data can also modify data. In biblical manuscript studies, data integrity remains an issue not because end users can corrupt a central repository (which will typically be read-only) but because the rapid proliferation of corrupted data through digital media results in increased opportunity for an inquirer to pull the *wrong* dataset.

The risk of data democratization leading to misunderstood data being used incorrectly and then passed along digitally should not be underestimated. As more than one contributor to the recent *Myths and Mistakes in New Testament Textual Criticism* volume has aptly demonstrated, even those within the guild of biblical studies have not always handled these data appropriately. Apologists who compare extant witnesses for classical works with those for the Greek New Testament, for example, have routinely underrepresented the classics because they have not understood that the two fields count manuscripts differently.[10] How can one expect those outside the academy with less training and experience to use these data wisely as they become more publicly available?

Note that the purpose of this article is not to turn a critical eye toward the misunderstanding of these data within academia. The process of scholarly dialogue that took place prior to data democratization should continue despite ongoing debates regarding failures of the peer-review process and the difficulty in identifying retracted articles for a primarily humanities-based discipline.[11] The rigors of scholarly research march on. Additionally, this is not an invective against the marketplace of ideas, though even that seems to be in jeopardy in an age of mis- or disinformation. This is instead a look at the risks created by widely disseminating

9. Marr, "What Is Data Democratization?"

10. James B. Prothro, "Myths about Classical Literature: Responsibly Comparing the New Testament to Ancient Works," in *Myths and Mistakes in New Testament Textual Criticism*, ed. Elijah Hixon and Peter J. Gurry (Downers Grove, IL: IVP Academic, 2019), 70–89.

11. A recent study of 129 Arts and Humanities papers, in which retractions were issued and identified using Retraction Watch and Scopus, concluded that the primary reasons for retraction were plagiarism and recycling and that retracted articles continued to be downloaded, read, and cited (Gali Halevi, "Why Articles in Arts and Humanities Are Being Retracted?," *Publishing Research Quarterly* 36 [2020]: 55–62). As biblical manuscript scholarship continues expanding cross-disciplinary research with statistical analyses and the material sciences, this profile is likely to change significantly.

biblical manuscript data in a post-truth[12] intellectual climate, where objective data are subordinated to subjective emotion.

2.1. Example #1: Wikipedia and Good Intentions

A significant example of "bad decisions based on bad interpretation of data" is readily demonstrated through Wikipedia. The Wikipedia article on academic use of Wikipedia (which is an essay and neither an encyclopedia article nor a policies or guidelines document) openly states, "Wikipedia is not a reliable source for academic writing or research."[13] It is "not a reliable source" because it is user-generated, mutable, vulnerable to vandalism and error, and is at best a tertiary source that may serve as "starting point for research, not an ending point."[14] Wikipedia issues a general disclaimer that acknowledges that the site "cannot guarantee the validity of the information" it hosts, that its content "is not uniformly peer reviewed," and that all of its information "is without any implied warranty of fitness for any purpose or use whatsoever."[15] Contrasting the reliability of digital and print media for academics, L. W. C. van Lit comments, "Scholarship needs fixed points of reference, which print can give and digital cannot."[16]

Despite these cautions, Wikipedia is currently (as of April 2022) the seventh most visited website worldwide.[17] When searching for a named biblical manuscript in Google using a private window, the first result will nearly always be Wikipedia; as a quick test, I searched on the first forty majuscule codex names (from Codex Sinaiticus through Codex Zacynthius), and codexsinaiticus.org was the only exception, presumably because a matching domain name would outrank a wikipedia.org page (the Wikipedia page was listed second). Searching on a handful of general search strings for biblical manuscripts, Wikipedia dominates the results pages:

12. The *Oxford English Dictionary* defines the modern use of *post-truth* as "relating to or denoting circumstances in which objective facts are less influential in shaping political debate or public opinion than appeals to emotion and personal belief."

13. "Wikipedia: Academic Use," Wikipedia, last modified 21 March 2022, https://en.wikipedia.org/wiki/Wikipedia:Academic_use.

14. "Wikipedia: Academic Use."

15. "Wikipedia: General Disclaimer," Wikipedia, last modified 11 February 2022, https://en.wikipedia.org/wiki/Wikipedia:General_disclaimer.

16. L. W. C. van Lit, *Among Digitized Manuscripts: Philology, Codicology, Paleography in a Digital World*, HOS 137 (Leiden: Brill, 2019), 53.

17. "Top Website Ranking," SimilarWeb, https://www.similarweb.com/top-websites/.

Search String	Wikipedia Result
Greek New Testament manuscripts	2nd
Greek New Testament	5th
New Testament papyri	1st
New Testament majuscules	1st
New Testament minuscules	1st
New Testament lectionaries	1st
Canons of criticism	1st

It is not unreasonable to assume that Wikipedia is a resource that will be used by someone (at least a nonspecialist) wanting to look up a biblical manuscript.

The Wikipedia entry for Codex Alexandrinus is an interesting test case for two reasons. First, the entry is currently marked as a "good article" by Wikipedia. A good article

> is an article that meets a core set of editorial standards, the good article criteria, passing through the good article nomination process successfully. They are well written, contain factually accurate and verifiable information, are broad in coverage, neutral in point of view, stable, and illustrated, where possible, by relevant images with suitable copyright licenses. Good articles do not have to be as comprehensive as featured articles (FA), but they should not omit any major facets of the topic: a comparison of the criteria for good and featured articles describes further differences.[18]

The description goes on to note that, "out of the 6,499,819 articles on Wikipedia, 36,335 are categorized as good articles (about 1 in 179)."[19] Second, the entry is interesting because of a historical intervention. Ten years ago on the Evangelical Textual Criticism blog, Peter Head wrote a post titled "How Bad Is Wikipedia? Codex Alexandrinus as a Test Case."[20] In this post, Head reviewed the first half of the entry and immediately documented ten factual problems in its contents. He concluded that:

18. "Wikipedia: Good Articles," Wikipedia, last modified 16 May 2022, https://en.wikipedia.org/wiki/Wikipedia:Good_articles.

19. "Wikipedia: Good Articles."

20. Peter M. Head, "How Bad Is Wikipedia? Codex Alexandrinus as a Test Case," Evangelical Textual Criticism, 15 February 2014, http://evangelicaltextualcriticism.blogspot.com/2014/02/how-bad-is-wikipedia-codex-alexandrinus.html.

> Wikipedia is quite bad. Facts are wrong, correct facts are placed in the wrong context, incorrect conclusions are drawn. Some of these errors would seem to have been deliberately inserted (either that or very stupid people are getting things badly wrong and adding them in). The best and most recent scholarship is cited the least. Evidence is not routinely provided. And the overall style is dreadful.[21]

According to the page's revision history, several updates were made in 2014 as a response to Head's post.[22] Some new citations were added and reference links updated in the intervening years but without significant change to the article contents. Where Head did not make comments (i.e., the latter half of the article), the contents and citations remain factually error-ridden, the scholarship is severely outdated, and the style retains its dreadfulness.[23] The article may appear scholarly, however, as it uses field-specific jargon and technical terms with high frequency. A statistical analysis comparing the use of academic vocabulary in Wikipedia articles (over a range of fields) with published research articles found that Wikipedia articles mimic the frequency of academic vocabulary use found in the

21. Head, "How Bad Is Wikipedia?"

22. "Codex Alexandrinus: Revision History," Wikipedia, last modified 16 May 2022, https://en.wikipedia.org/w/index.php?title=Codex_Alexandrinus&action=history.

23. For an amusing example of content error, the article states, "At the end of each book the colophon is ornamented by pretty volutes from *prima manu*," citing Scrivener's S*ix Lectures on the Text of the New Testament and the Ancient Manuscripts Which Contain It* (Cambridge: Deighton, Bell, 1875), 52. Scrivener instead stated, "At the end of each book we notice pretty arabesque ornaments in ink by the first hand" (52–53). Scrivener, unfortunately, is incorrect on this point; the books end with titles and not colophons, and none of the decorative pieces could be considered a volute. For an example of citation error, see: "Present scholars agreed in that case (Metzger, Aland, Hernández, Jongkind).[22]:119–120[18]:101," where footnote 22 links to "Skeat, T. C. 'The Provenance of the Codex Alexandrinus'. *The collected biblical writings of T. C. Skeat*" and footnote 18 links to "Hernández, Juan (2006). *Scribal habits and theological influences in the Apocalypse*. Mohr Siebeck. p. 102." Outdated scholarship appears in the use of free (read: out of copyright) resources that are more accessible to the general population that is editing Wikipedia articles. The latest specialist works are generally prohibitively expensive. This is a regrettable but realistic limitation for these editors. Finally, writing style: Though borrowing Head's term *dreadfulness*, the writing style consistently uses words and concepts imprecisely. The editor who wrote "words are written continuously in a large, round and well-formed uncial hand" likely intended "continuously" to refer to *scriptio continua*.

articles.[24] The researchers note that any frequency of academic words "does not guarantee that they are used correctly, nor that the style of argumentation, the framing of questions or the rhetorical structure of the overall text would meet the standards of what competent members of the academic community would expect from a published text."[25]

In his ethnography of Wikipedia, Dariusz Jemielniak notes that rejection of credential checks as a means of controlling editors and content has a result of allowing "full democratic participation by different people with different backgrounds.… Trust or credential control is substituted with precise behavioral scripts and formalization of discussion rules."[26] He notes that trust in editors, which is "developed locally," has "organizational benefits" despite "its obvious disadvantages."[27] The obvious disadvantage is that a well-cited argument does not guarantee that its citation data have been understood. When the 2014 Codex Alexandrinus article addressed the quire structure of the manuscript, it noted that "most of the folios were originally gathered into quires of eight leaves each. In modern times it was rebound into quires of six leaves each." I reference my own clarification in Head's blog post to explain how the source material was misunderstood:

> The problem most likely results from a misunderstanding of Thompson's comments in the introduction to volume 1 of the full-scale facsimile. When the NT volume of the facsimile was published (1879), he believed that most of the quires were composed of six leaves, but he corrected himself in volume 1 (1909), claiming that "when the MS. was re-bound in the present century, the quire-formation was disregarded, the leaves being separated and re-backed and made up into sets of six" (*Facsimile of the Codex Alexandrinus* [London: Trustees of the British Museum, 1879–1883], 1:8).[28]

24. Turo Hiltunen and Jukka Tyrkkö, "Academic Vocabulary in Wikipedia Articles: Frequency and Dispersion in Uneven Datasets," in *From Data to Evidence in English Language Research*, ed. Carla Suhr, Terttu Nevalainen, and Irma Taavitsainen, Language and Computers 83 (Leiden: Brill, 2019).

25. Hiltunen and Tyrkkö, "Academic Vocabulary in Wikipedia Articles," 302.

26. Dariusz Jemielniak, *Common Knowledge? An Ethnography of Wikipedia* (Stanford, CA: Stanford University Press, 2014), 120.

27. Jemielniak, *Common Knowledge?*

28. See the comment on Head, "How Bad Is Wikipedia?," http://evangelical-textualcriticism.blogspot.com/2014/02/how-bad-is-wikipedia-codex-alexandrinus.html?showComment=1392518201234#c243848788094668817.

Therefore, Head was perfectly justified in commenting that "now this sounds plausible enough, so long as you have *no idea at all what a quire is*." Articles such as this have the trappings of scholarship—frequent use of technical terms, a high number of citations, lists of Greek variants, and so on—without fully understanding the scholarship. As a quick and ready go-to resource on the internet, the mistakes recorded in Wikipedia will spread far and wide.[29]

2.2. Example #2: Forgery and Ill Intentions

The story of the so-called Gospel of Jesus's Wife forgery has made headlines since its debut in September 2012. Numerous peer-reviewed articles, news articles, blog posts, and journalist Ariel Sabar's book-length coverage detail the history of the story as well as address specific problems with the Coptic papyrus fragment and its five other associated Coptic fragments.[30] In summary, the con man and forger Walter Fritz delivered six Coptic papyrus fragments to Harvard professor Karen King, the most sensational of the fragments being named (by King) the Gospel of Jesus's Wife (hereafter GJW) based on the phrase at the middle of the fragment reading "Jesus said to them, 'My wife….'" Suspicion of forgery plagued the papyri from the start, yet the GJW fragment was presented to the press and published in the *Harvard Theological Review* against the recommendation of two peer reviewers. Subsequent analyses of the GJW fragment and another fragment produced by the same hand named the Harvard Lycopolitan John (HLJ) fragment successfully demonstrated that the two were modern creations.

While Fritz may have had rudimentary knowledge of Coptic, he did not have the expertise to generate forgeries without leaning on published work. He also did not have the paleographical background necessary to manage the attendant material issues: making proper lampblack ink, utilizing a pen instead of an anachronistic brush, the muscle memory of writing quality Coptic script, or the ability to reproduce a properly datable hand. Ironically, the online Coptic transcription data that made the forgeries possible also made them detectable.

29. Should an academic scoff at this point and note that there are better resources freely available such as Academia.edu, I suspect that Academia.edu is likely to become the pseudo-scholarly version of Wikipedia.

30. For the larger context, see Ariel Sabar, *Veritas: A Harvard Professor, a Con Man and the Gospel of Jesus's Wife* (New York: Doubleday, 2020).

To produce the GJW fragment, the forger assembled what Francis Watson referred to as a "patchwork" of words and phrases using Michael W. Grondin's 2002 PDF of the single Coptic manuscript of the Gospel of Thomas. Andrew Bernhard detailed the collaborative effort to analyze the fragment in *New Testament Studies*, noting: "On 11 October 2012, Bernhard released an online article calling attention to a number of features in the text that suggested *GJW* was probably prepared by someone relying on Grondin's edition of *GTh*, and Goodacre simultaneously spotlighted the most startling discovery in a blog post: *GJW* seems to reproduce a typographical (and grammatical) error directly from 'Grondin's Interlinear.'"[31]

Additionally, Christian Askeland's analysis of the HLJ fragment, which aggregates the analyses of other scholars with his own, convincingly demonstrated that the fragment was produced from a PDF of Herbert Thompson's 1924 edition of the Lycopolitan Qau codex.[32] Apart from other problems for the authenticity of the fragment (which were not related to data democratization), the forger erred by duplicating seventeen line breaks from Thompson's edition, by misunderstanding the column layout of the edition, and by failing to follow a pattern of skipping lines from the edition (and thus producing too small of a lacuna for the Coptic text) based on a page break in Thompson's edition.[33]

In both instances, the forger had access to enough transcription data to produce a papyrus fragment with Coptic text. In neither case did the forger have the appropriate skillset to use those data with expertise.

3. How Does This Happen?

Historically, institutional web-based and social media–based information sharing has been implemented by academic organizations or individual scholars both externally (to engage the public, students, donors, scholars, etc.) and internally (what has been termed *enterprise social media*).[34] The

31. Andrew Bernhard, "The *Gospel of Jesus' Wife:* Textual Evidence of Modern Forgery," *NTS* 61 (2015): 337.

32. H. Thompson, *The Gospel of St. John according to the Earliest Coptic Manuscript* (London: Quaritch, 1924).

33. Christian Askeland, "A Lycopolitan Forgery of John's Gospel," *NTS* 61 (2015): 314–34.

34. Paul M. Leonardi, Marleen Huysman, and Charles Steinfield, "Enterprise Social Media: Definition, History, and Prospects for the Study of Social Technologies in Organizations," *Journal of Computer-Mediated Communication* 19 (2013): 1–19.

affordances of these externally facing data (broadly, their perceived uses) are very different for expert and nonexpert end users, especially when the data presentation is effectively raw or unmediated.[35] In both previous examples, data that were made openly available and that have served the academic community in a positive way were instead misunderstood and mishandled by those outside academia. What factors are contributing to an environment where data democratization does not simply lead to a wider community of understanding? Despite the malicious nature of the GJW example, let us temporarily set aside willful and agenda-driven data manipulation and give our interlocutors the benefit of the doubt that their mishandling of data is unintentional. Factors that remain and contribute to misunderstood and mishandled data include but are not limited to: (1) the devaluation of authority; (2) the ascension of personal authority; and (3) data handling without domain expertise.

3.1. The Devaluation of Authority

From the aftermath of World War II to the end of the Cold War, the Western world began to devalue obedience to authority and eventually associate obedience with a lack of critical thinking and facility of being manipulated. The very definition of authority was questioned and whether it should be viewed positively or negatively. Abuses of authority—how many coming of age in this era were fed the cautionary tale of Milgram's experiments?—were enough to demonstrate that those who traditionally held authority were no longer worthy of it.[36]

Confusion about how authority might be defined, especially with its relationship to power, was subsequently coupled with an uncertain relationship between authority and legitimacy. Attempting to appeal to some form of legitimacy, science became the impersonal vehicle for appeals to authority. As Frank Furedi observes: "Despite Western culture's disenchantment with rationality, competing parties in the key controversies of

35. Further, "accepting a relational view of affordance" will eliminate the idea of a "generic user" and describing technology as a set of features because "user intent, abilities, social, environment, as well as the specifics of the situation will matter even more." See Samer Faraj and Bijan Azad, "The Materiality of Technology: An Affordance Perspective," in *Materiality and Organizing: Social Interaction in a Technological World*, ed. Paul M. Leonardi et al. (Cambridge: Cambridge University Press, 2012), 255.

36. Frank Furedi, *Authority: A Sociological History* (Cambridge: Cambridge University Press, 2013), 328–98.

our era continually appeal to the authority of science, and the narrative of expertise represents the most influential validation of authority in our times."[37]

3.2. The Ascension of Personal Authority

Even the professional expert and the wielder of science have their authority questioned when controversial or emotional topics arise. Several competing factors are at work to strip external bodies of their authority. When issues of personal importance are involved, persons with *belief superiority*—"the belief that one's views are superior to other viewpoints"—will mistakenly believe they have greater knowledge about the topic than is warranted.[38] In a six-part study of individuals with belief superiority, researchers found that errors in metacognitive ability (e.g., an awareness of one's own limitations and errors) increase when the individual cares deeply about an issue or takes an extremist position. As a result, "belief superiority was associated with a significantly larger gap between perceived and actual knowledge … even when controlling for confidence."[39] The study revealed that those with belief superiority misinterpreted those who shared their belief as sharing greater knowledge and did not seek out all types of belief-related information equally. Additionally, "those high in belief superiority were aware of their biased information-seeking behavior."[40]

Competing ideologies in the current world climate face populations that are more invested in emotionally satisfying beliefs being reinforced by external sources than being challenged by objective facts. This post-truth atmosphere is one in which buzzwords such as "fake news," "echo chambers," "filter bubbles," and "hot cognition" speak to the idea that reasoned, evidence-based decision making is passe in this cultural setting. In his book *Post-truth*, Lee McIntyre comments:

> When a person's beliefs are threatened by an "inconvenient fact," sometimes it is preferable to challenge the fact. This can happen at either a conscious or unconscious level (since sometimes the person we are seeking to convince is ourselves), but the point is that

37. Furedi, *Authority*, 396.

38. Michael P. Hall and Kaitlin T. Raimi, "Is Belief Superiority Justified by Superior Knowledge?," *Journal of Experimental Social Psychology* 76 (2018): 290.

39. Hall and Raimi, "Is Belief Superiority Justified," 296.

40 Hall and Raimi, "Is Belief Superiority Justified," 302.

> this sort of post-truth relationship to facts occurs only when we are seeking to assert something that is more important to us than the truth itself. Thus, post-truth amounts to a form of ideological supremacy, whereby its practitioners are trying to compel someone to believe in something whether there is good evidence for it or not. And this is a recipe for political domination.[41]

Science, with its likelihood of utilizing inconvenient facts, has no foothold where personal authority has taken the reins. This is because humans utilize *motivated reasoning* to safely arrive at conclusions "that are primed by our deeper underlying values, worldviews, vested interests, fears, and self-identities and social identities."[42] It is with little surprise that sociologists Harry Collins and Robert Evans comment that

> In today's world the scales upon which science is weighed sometimes tip to the point where ordinary people are said to have a more profound grasp of technology than do scientists. Our loss of confidence in experts and expertise seems poised to usher in an age of technological populism.[43]

3.3. Data Handling without Domain Expertise

Those of us who regularly work with biblical manuscripts have been excited by the flood of images, transcriptions, and studies in paleography, textual transmission, and codicology that have been made available online. Much of those data remain uninterpreted. For example, the images of a manuscript in an online repository typically come without commentary on their

41. Lee McIntyre, *Post-truth* (Cambridge: MIT Press, 2018), 13.

42. Sara L. Rynes, Amy E. Colbert, and Ernest H. O'Boyle, "When the 'Best Available Evidence' Doesn't Win: How Doubts about Science and Scientists Threaten the Future of Evidence-Based Management," *Journal of Management* 44 (2018): 2998.

43. Harry Collins and Robert Evans, *Rethinking Expertise* (Chicago: University of Chicago Press, 2017), 1–2. From an entirely different perspective, social-justice scholarship is also challenging the scientific method "because empirical research that values knowledge production rooted in evidence and reasoned argument is an unfairly privileged cultural construct of white Westerners" (Helen Pluckrose and James Lindsay, *Cynical Theories* [Durham, NC: Pitchstone, 2020], 62). In this and related models, knowledge and reason are directly tied to social and political power. The topic lies outside the scope of this study, however.

contents. To understand the data in a meaningful way requires experience and expertise.

Although there has been a cultural move to devalue expertise alongside authority, studies indicate that novices and experts in a range of fields operate with fundamental differences when processing data. Additionally, gaining expertise is a two-pronged endeavor involving both theory (subject knowledge) and practice (applied knowledge). Expert theory without practice is labeled *interactional expertise*, "since interactional experts do not contribute to or create within the domain of their expertise but are still immersed in the language of that particular domain."[44] Expertise in theory *and* practice is labeled *contributory expertise.* Virginia Tucker observes that

> a key practice among experts is *reflection*, particularly as it affects the process of learning. Schön described a stark contrast between knowledge acquisition—extreme or otherwise—and learning at a level he called *professional artistry*. He explained, "Artistry is an exercise of intelligence, a kind of knowing, though different in crucial respects from our standard model of professional knowledge."[45]

As a result, the practices used by the expert differ from the novice. Experts, for example, produce "deeper and richer" representations of problems, work as more flexible and opportunistic planners, more skillfully handle ambiguous data as "top down processors," and "develop automaticity in their behavior to allow conscious processing of ongoing information."[46]

While expert-level understanding may be required simply to interpret some data, especially when dealing with specialized knowledge, deeper understanding and popular understanding are particularly at odds when the science is disputed. Collins and Evans conclude that "the last three decades of social studies of science have shown us that, in disputed science, detail, tacit knowledge, and unspoken understanding of who is to be trusted among those who work in the esoteric core of the science are vital components of decision-making at the technical level.

44. Jenny Rice, "Para-expertise, Tacit Knowledge, and Writing Problems," *College English* 78.2 (2015): 123.

45. Virginia M. Tucker, "Learning Experiences and the Liminality of Expertise," in *Threshold Concepts in Practice*, ed. R. Land et al. (Rotterdam: Sense Publishers, 2016), 98.

46. Tucker, "Learning Experiences."

Popular understanding hides detail, has no access to the tacit, and washes over scientists' doubts."[47]

According to Collins and Evans, one of the most significant challenges for the nonexpert seeking to process disputed domain-specific data is not having access to the expert community. They continue:

> Many of the papers in the professional literature are never read, so if one wants to gain something even approximating to a rough version of agreed scientific knowledge from published sources one has first to know what to read and what not to read; this requires social contact with the expert community. Reading the professional literature is a long way from understanding a scientific dispute. The question, then, even for those who read the journals in which primary research findings are published, is whether their knowledge matches the Trivial Pursuit player's, the chess novice's, the experienced chess player's, or the chess master's understanding of the bishop's move. Our claim is that in the case of scientific disputes primary source knowledge is not much better in respect of the science than a chess novice's understanding in respect of the bishop's move.[48]

Even the well-read novice will not know which sources are more reliable, and this is a problem with lacking domain expertise in general: data will be decontextualized from current discussions of the field, sources will be selected without circumspection (and perhaps with too small or too homogenous of a sampling), and the novice will not have enough experience to ask the right questions.

3.4. Demonstration

3.4.1. Example: The Shocking Bible

A freelance writer in Seattle named Jonathan Poletti wrote a 2020 article titled "The Shocking Bible," which was published on Medium.com, a website for digital publishing.[49] The short (1,600-word) article identifies Codex

47. Collins and Evans, *Rethinking Expertise*, 20.
48. Collins and Evans, *Rethinking Expertise*, 22–23.
49. Jonathan Poletti, "The Shocking Bible," Medium.com, 2020 (no longer avail-

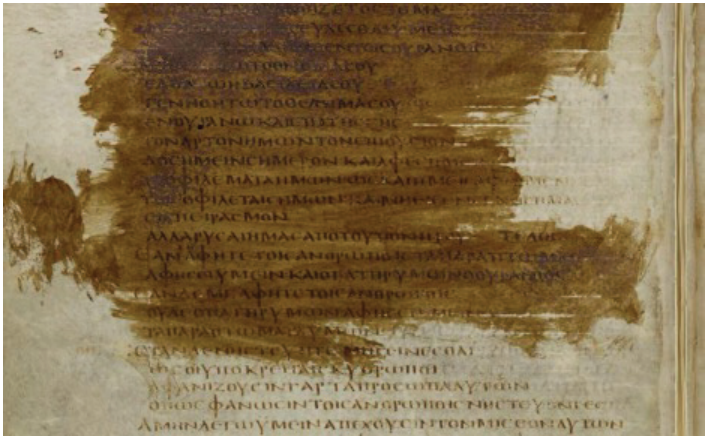

Figure 1.

Bezae as having a "shockingly different" New Testament text. While the piece struggles somewhat to maintain a clear thesis, the overall theme is that this one codex preserves an older, richer, more Jewish, and, overall, more inclusive biblical text that orthodox Christians should have taken seriously but instead suppressed. Though the argumentation is jejune, there are three elements of the article that are pertinent to data democratization.

First, after the opening paragraphs a partial image of a page from Codex Bezae is included without comment (fig. 1). The image is from folio 16v (flesh side) and contains Matt 6:8–20. The immediate impression of the image is that the author has included it as evidence of *suppressed* biblical text, albeit poorly done, since the Greek is still readable. But anyone who has worked with codices immediately recognizes that these brush marks are from chemical agents used to try to *recover* lost or difficult-to-read text. Before ultraviolet technology was available, scholars were limited in how they might recover badly faded ink or palimpsested pages of text. Chemical reagents could be brushed on the necessary pages to make the ink traces more readable. A first scholar to make widespread use of chemical reagents was Angelo Mai (1782–1854) of the Biblioteca Ambrosiana in Milan and

able). Poletti writes the following for the about page: "I wanted *facts* on Christianity and the Bible. I went to get them, and found out a lot that I'd never heard in church. This channel is for highly-researched posts" (https://belover.medium.com/about). Poletti is "Editor of Belover and QueerTheory." He is identified as a freelance writer on Salon's website: https://www.salon.com/writer/jonathan_poletti.

later of the Vatican who experimented with various compounds before settling on distilled gall nut to retrieve erased palimpsest texts.[50]

While the compounds made the ink visible, they would also darken not long after and render the page difficult or impossible to read. Being a folio of the Lord's Prayer, this may have been a popular page to turn to and thus receive greater damage over time. There are a few minor variant readings in the recovered text but nothing interesting. Poletti did not understand the image, could not read the Greek to understand it was unremarkable, and consequently drew the opposite conclusion.

Second, Poletti notes the textual variant in Codex Bezae at Luke 1:28, where the phrase ευλογημενη συ εν γυναιξιν is added to the end of the verse, mimicking the language of verse 45, where Elizabeth says the same thing by the Holy Spirit. What is remarkable is that Poletti is apparently unaware that this same reading appears in many other manuscripts (A C D K Γ Δ Θ f13 33 892 1424 2542 L2211 𝔐 latt sy bo^{mss} Eus). His conclusion, nonetheless, is, "Though the difference is not great, scholars notice the Bezae version has an allusion which was otherwise concealed. Mary is being compared, through the same phrase, to Jael in Judges 5:24, who is 'most blessed of women.'" The phrase, being in the majority text, appears in the ubiquitous King James Bible.[51] In this case, the user of these freely available data had no idea where they fit into the larger data set and again drew an incorrect conclusion that the reading was somehow suppressed outside of this codex.

Third, despite the article's theme of suppression, the author is apparently unaware of the impressive bibliography for Bezae, which has, for example, 344 entries currently published in the *Kurzgefasste Liste*. Poletti cites the first line of David Parker's *Codex Bezae* ("Codex Bezae is a manuscript that has generally managed to provoke strong emotions")[52] with a link to that quote on Google Books. Assuming he read beyond that first page, Parker offers a bibliography in each section of his book that would have been sufficient to demonstrate the codex has been discussed at length.

50. Raymond Clemens and Timothy Graham, *Introduction to Manuscript Studies* (Ithaca, NY: Cornell University Press, 2007), 104–5; A. Németh, "Angelo Mai and the Palimpsests," Vatican Palimpsests, https://spotlight.vatlib.it/palimpsests/feature/angelo-mai-and-the-palimpsests.

51. A commenter labeled "Janet Jones" points this out to Poletti, who responds by commenting that "the phrase disappears in later translations."

52. David C. Parker, *Codex Bezae* (Cambridge: Cambridge University Press, 1992), 1.

Poletti is currently listed as having 111,000 followers on medium.com.

3.4.2. Example: Defending the Textus Receptus

Not only can images of biblical manuscripts published online be misunderstood by the nonexpert viewer, but also metadata regarding the textual history of those manuscripts. Kirk DiVietro is the pastor of Grace Baptist Church in Franklin, Massachusetts, and a prominent member of the Dean Burgon Society, an organization dedicated to defending the "Traditional Masoretic Hebrew Text of the Old Testament," the "Traditional Received Greek Text of the New Testament" (specifically the Greek used for the 1611 King James Bible), and the "Traditional English Translation of the Bible—the King James Version (or Authorized Version)."[53] In July 2017, DiVietro delivered a paper to the Dean Burgon Society titled "Attacking the TR by the Coherence Based Genealogical Method," which is available online.[54]

Near the start of his paper, DiVietro confesses, "I have studied it for the better part of four months pretty seriously and still cannot begin to take it apart completely." His goal, instead, is to make the method understandable to a general audience. After expressing concerns about textual criticism in general, DiVietro introduces the CBGM as follows:

> Now what is the GCBM [*sic*]? Okay, what they have done now is that they have thrown away the arguments that Westcott and Hort have made. They're gone! They don't matter anymore! So, Dean Burgon Society doesn't have to worry about Codex B, Codex Sinaiticus, we don't have to worry about who wrote what on what papyrus because *none of that matters anymore*. We look at each reading. We don't care where we find it—as long as it's not in a TR manuscript. We look at each reading and we try to decide this reading here is related to that reading over there and that reading is related to this one over here. That's the grandfather, this is the son, here's the cousin, here's the second cousin. Okay, now we understand where that reading came from and we give it a com-

53. https://tinyurl.com/SBLPressTC2024h1; https://www.unitedstateschurches.com/massachusetts/grace-baptist-church-franklin/290538; https://www.facebook.com/kirk.divietro. At the time of my writing, the second URL was no longer active.

54. Kirk DiVietro, "Attacking the TR by the Coherence Based Genealogical Method." SermonAudio.com, 26 July 2017, https://www.sermonaudio.com/sermoninfo.asp?SID=717171125246.

> puter value. And then we say, "Okay, now that's those readings. Now here's some other readings in the same verse and they go this way. And then these over here go this way." And the coherence is that we have to make sense out of all these readings, so we look at all these readings, and we give it a computer value. Now where's the computer value come from? Their heads. And I push a button and the computer tells me what the Greek New Testament is. That's the GCBM [*sic*].

The video only occasionally shows the PowerPoint slides that accompanied the paper, but an image of a local stemma is onscreen at one point, and DiVietro is likely referring to textual flow diagrams when he comments, "It gets more complicated. There's just some of the charts. This is to explain it. I'm not going to even try."

DiVietro's frustration at dealing with these data is evident in his comments. "The entire new system of picking the right reading in the GNT is a barrage of baloney. It is purposely engineered to be so complex that they can say well you just don't understand. And most of us don't have a leg to stand on." Later he reiterates: "That's this new method that's in the Nestle-Aland 28th edition of the Greek New Testament and it says there's not a person in the Dean Burgon Society smart enough to unravel this because we don't even understand what we're doing!"

In this instance, data democratization has hindered rather than helped understanding. A user in the public sphere spent four months reviewing the data, found it confusing and frustrating, and in the end believed that using the CBGM involved pushing a button to produce a Greek New Testament. Unfortunately, the result is also misinformation about the CBGM, which is already misunderstood among some scholars.

4. Summary and Conclusion

As biblical manuscript data becomes more openly available to the world at large, current cultural trends indicate that the data will be improperly handled more frequently by inappropriately confident nonexperts. The double-edged sword of accessibility through internet media platforms is that both the good source data and the subsequently generated misinformation will spread quickly and widely. Democratically run and crowd-sourced sites for dissemination of this information will work with dated scholarship that is partially understood and relayed with uninten-

tional errors. This is a transmission history we would not wish upon the Greek New Testament.

What can we do to mitigate this problem? Is this a call for a priesthood, a table at which only the invited can sit? I think not. Unless one is trained in medicine, a broken bone leads one to a doctor. Without experience in repairing engines, one takes a malfunctioning vehicle to a mechanic. Anyone could be trained to work with manuscript data, but the training is necessary. There is significant investment to sit at this table: one must be trained in ancient languages, in papyrology, in paleography, and in text-critical methods, and one must have actual experience examining manuscripts to understand the broader context of any single manuscript.

The Bible—its preservation, its reliability, its texts—is a subject of great personal importance to many, whether perceived as good or bad. Facts and evidence used to explain the Bible—its preservation, its reliability, its texts—will not be sufficient to rise above the din of a post-truth public. Experts and academics will need to evaluate their sources more closely; are data retrieved online coming from trustworthy sources? Experts must improve communication with the community. Hearts must be touched before evidence will matter. Humanizing and contextualizing research findings for a nonacademic audience may go a long way toward injecting truth into the din. Additionally, providing clear and accessible summaries of expert work allows experts to control how information is delivered to the public—do we really want the History Channel reinterpreting what we find? Finally, experts must utilize public-accessible platforms. The average person is far more likely to read a blog post than to purchase a $200 monograph.