What Makes an Open Textbook Work: A Guidebook for New Authors

Ronald G. Thomas

2026-09-07



An open license is a distribution decision, not a teaching decision: it determines who can open the book, and nothing at all about what happens after they do.

Introduction

An open license buys something other than what prospective authors expect, as the meta-analysis that ought to be the first thing every open-textbook author sees makes plain. Clinton and Khan pooled 22 comparisons covering slightly over one hundred thousand students and found the difference in learning between open and commercial textbooks to be, for practical purposes, zero (Hedges' g = 0.01, p = .87). In the same analysis, across 11 comparisons and roughly 78,000 students, course withdrawal was substantially lower under open texts (odds ratio 0.71, p = .005). The result is published in *AERA Open*.

The two findings together are the most useful thing in this literature, and they are usually reported one at a time by people with a position to defend. Read together they say something specific. Openness reliably buys access, and access buys retention. Openness buys no pedagogy whatsoever. A free book is not thereby a good book, and the license confers nothing on the prose.

That is a liberating finding rather than a discouraging one, because it locates the work. If the license did the teaching, an author would have little left to decide. Since it does not, the ordinary craft questions return with full force: how long a chapter should be, where the worked examples go, whether the exercises have answers, whether a student can find anything in a 700-page PDF. This post is an attempt to answer those questions from evidence rather than from taste, drawing on the efficacy literature, on 954 reviewer assessments of open textbooks, and on retrospectives by authors who have actually maintained a free book for a decade.

Motivations

- I am drafting two open textbooks of my own, and I wanted to know whether the design decisions I had been making by instinct were supported by anything.
- The advocacy literature on open educational resources is heavily weighted toward cost and access. It is comparatively quiet on craft, which is the part an author actually controls.
- Reviewer data on open textbooks exists in bulk and is rarely read as what it is: a large, systematic account of how these books fail.
- New authors appear to repeat a small number of avoidable mistakes. A guidebook seemed more useful than another argument for adoption.

Objectives

1. Separate what open licensing demonstrably delivers from what it demonstrably does not.
2. Extract, from reviewer and student data, the specific failure modes that recur across open textbooks.
3. Convert those findings into tips and into a dos-and-don'ts list an author can apply while drafting.
4. State plainly where the evidence is thin, so that a reader can discount the weaker recommendations appropriately.



What the Evidence Actually Supports

We begin with the claims that are safe to make, since a guidebook that overstates its evidence is a poor advertisement for scholarly habits.

Access effects are large and concentrated. In a Florida statewide survey, 64.2% of students reported not purchasing a required text because of cost, 42.8% reported taking fewer courses, and 35.6% attributed a poor grade to textbook cost. A 2022 follow-up still found 53% skipping a required purchase. This is the mechanism behind the withdrawal finding above. Students who could not previously obtain the book are the ones whose outcomes move.

The gains are unevenly distributed, and the honest version of the result says so. Colvard, Watson and Park studied 21,822 students across eight large introductory courses at the University of Georgia and reported in *IJTLHE* that mean course grade rose from 2.806 to 3.048 and the DFW rate fell from 10.87% to 8.19%. Among Pell recipients the DFW decline was 4.43 percentage points; among part-time students it was 10.14 points. The caveat that advocacy summaries tend to drop is that full-time students' DFW rate actually rose slightly, from 6.28% to 7.38%, and that the design is observational and pre-post, confounded with time and instructor. The result is real and it is not a claim that open texts teach better.

Most students will not read the chapter. Reading-compliance research going back to the 1970s puts completion of a given day's assigned reading at or below 30%, with students reading something like 37% of assigned pages. A chapter written on the assumption that it will be read straight through, in order, is designed for a minority of its audience.

Screens cost comprehension, mostly under time pressure. Delgado and colleagues' meta-analysis in *Educational Research Review* found a reliable screen-inferiority effect that appears under time constraint and attenuates under free reading, together with greater overconfidence when reading on screen. The full text is available. The design implication is not to abandon digital delivery. It is that a digital-first book has to supply, deliberately, the metacognitive checks that a slower paper reading supplies for free.

Two study techniques dominate, and neither is the one your book is designed to support. Dunlosky and colleagues rated practice testing and distributed practice as high-utility, and highlighting and rereading as low-utility, in their review for *Psychological Science in the Public Interest*. The low-utility pair is what students actually do. An author who adds embedded retrieval questions with feedback is building the high-utility technique into the artifact. An author who writes a beautifully highlightable chapter is optimizing for the technique that does not work.

How Open Textbooks Actually Fail

Belikov and McLure coded 954 reviews of 235 open textbooks from the Open Textbook Library, and their analysis in *IJOER* is the closest thing we have to a systematic failure inventory. The share of reviews rating each criterion positively:

Criterion	Positive	Partial	Negative
Interface	83.7%	10.8%	1.5%
Consistency	83.2%	12.7%	3.0%

Criterion	Positive	Partial	Negative
Grammar	81.6%	13.4%	1.5%
Relevance and longevity	81.2%	14.0%	2.5%
Accuracy	80.8%	15.2%	1.0%
Clarity	80.4%	15.6%	3.6%
Organization	78.0%	15.0%	3.0%
Comprehensiveness	73.0%	22.1%	2.1%
Cultural relevance	66.6%	24.5%	3.2%
Modularity	62.5%	26.4%	2.2%

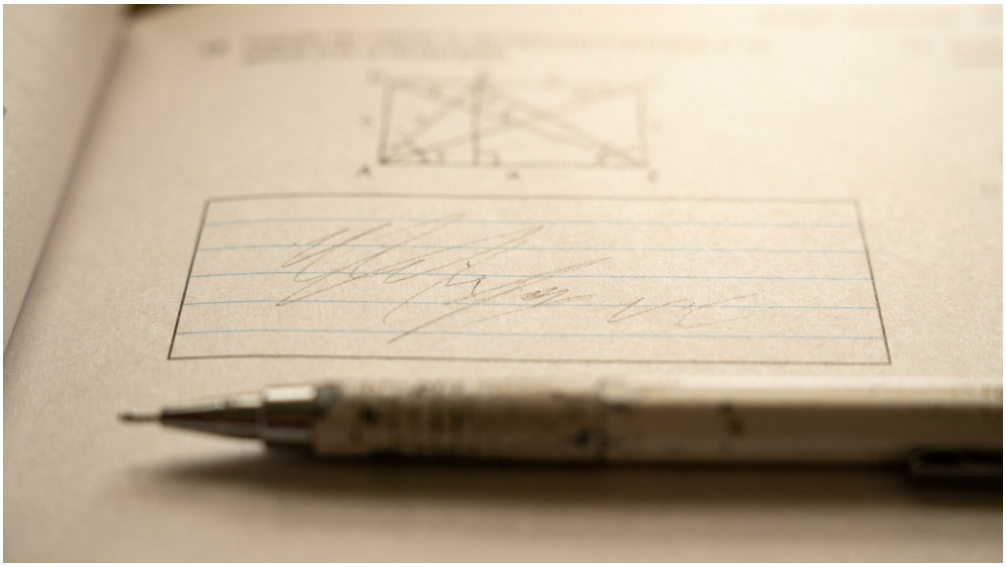
Two things are worth noticing. The weakest criteria are modularity and cultural relevance, not accuracy. Open textbooks are not, in the main, wrong. They are hard to assign in pieces and they do not reflect the students reading them.

The authors also name a structural trade-off that I have not seen stated so plainly anywhere else: 'increased modularity results in decreased consistency'. Commercial textbooks buy voice consistency with a single author and substantial professional editing. Open books, often multi-authored and chapter-partitioned, pay for their modularity in tonal coherence. One reviewer's complaint about a nine-author book is the whole problem in one sentence: each chapter was 'tonally very different from the others'.

The recurring specific complaints, quoted from those reviews, are mundane and almost entirely preventable:

- A 700-page PDF with 'no table of contents and little in the PDF that allows for quick and easy browsing without intense scrolling', and elsewhere 'no index, glossary or table of contents'.
- 'The textbook contains only words', with a lack of color, images, and charts that reviewers expected to cost student engagement.
- Coverage that 'is a bit thin when it comes to examples'.
- Solutions to in-text practice problems that require scrolling to the end of the chapter.
- Reference lists trailing off around 2007, which reviewers read as a signal the book had been abandoned.
- Examples in which marginalized students 'would not see themselves reflected much'.

On the positive side, students in design studies liked embedded video distributed through the chapters rather than clustered at the end, consistent layout, a persistent sidebar table of contents, and a 'manageable amount of information' per chapter. Roughly 95% agreed open resources were easy to use and 78% found them more up to date than their print texts, though just under 20% still preferred a traditional textbook, and objections clustered on eyestrain and small screens.



Tips

What follows is the operational part. Each tip is tied to something above, and I have tried to avoid recommending anything I could not connect to evidence or to a working author's practice.

1. Write for a course, not for a field. Allen Downey's textbook manifesto argues that conventional textbooks bloat toward a thousand pages by trying to serve several courses at once, and that a weekly fifty-page assignment already exceeds what students absorb. His target is roughly 140 pages for a semester course. A short book that matches one syllabus will be read; a comprehensive one will be consulted, at best.

2. Lead with the example, then the method. The OpenIntro authors describe their organizing principle as example first, method second: each chapter opens by naming the questions to come, and the methodology arrives as the way to answer them. Their production account is published in *TISE*. This directly addresses the 'thin on examples' complaint, and it matches the worked-example tradition in cognitive load research.

3. Put the answers where the student is. OpenIntro ships more than 550 exercises with solutions to nearly all in-text exercises in footnotes, on the stated principle that the student should be able to check the work immediately. Reviewers complained specifically about solutions parked at the chapter end. The distance between a problem and its answer is a design parameter, and shorter is better.

4. Build retrieval into the text. Given the Dunlosky findings, the highest-leverage addition is not more exposition but more questions with feedback. A short set of retrieval prompts per section does more measured good than another worked paragraph.

5. Design for the section, not the book. Modularity is the weakest criterion in the reviewer data and a first-class criterion in the review rubric, and the reading-compliance numbers say the unit of consumption is the section anyway. Every section should state its own prerequisites and be assignable on its own.

6. Name an editor, even if it is you wearing a second hat. OpenIntro assigned one author a dedicated internal-editor role. If modularity costs consistency, then consistency has to be somebody's explicit job rather than an emergent property.

7. Ship more than one format. In an Indiana University e-textbook pilot, more than a quarter of participating students believed they would have learned more from paper. OpenIntro's response was format pluralism: free PDF, free web version, and a print-on-demand paperback under ten dollars. Free web and paid print are not in tension. *Automate the Boring Stuff with Python* has sold over half a million copies while remaining free to read online.

8. Make the figures reproducible artifacts. OpenIntro distributes a companion R package containing both the data sets and the functions that generate the book's figures. A figure that can be regenerated is a figure that can be corrected, and a book whose code re-executes at render time cannot drift silently away from its own output.

9. Let readers file corrections. Yihui Xie has described the effect of accepting pull requests on a book repository as acquiring 'outsourced careful human spell checkers'. An open book can fix a typo in the book itself rather than publishing an errata sheet that most students never read, which is a genuine structural advantage over print.

10. Teach from the draft, and treat failure as a book defect. Downey's practice is to draft before teaching the course, publish immediately, and collect student notes on mistakes and rough spots through the semester. When comprehension fails across several students, he revises the book rather than blaming the cohort. This is the cheapest review process available to an academic author, and it runs every time the course does.

11. Solicit a formal review, because it is also marketing. Faculty report discoverability as a barrier for roughly half of them, and perceived quality is mediated by colleague recommendation and peer review rather than by intrinsic merit. The Open Textbook Library's review program is explicitly dual-purpose: it addresses credibility and raises awareness in the same act. As OpenIntro put it, the marketing resources of a commercial publisher are 'truly humbling' to a volunteer project, so the free distribution channels have to be worked deliberately.

12. Choose CC BY unless you have a reason not to. Licenses carrying ND forbid derivative works and are therefore widely held not to qualify as open educational resources at all, since revision and remix are the distinguishing affordances. SA introduces compatibility conflicts when combining sources. Note also Downey's observation that as copyright holder you retain the ability to grant different licenses to different parties later, so the public default does not foreclose your options.

Dos and Don'ts

Do:

- Do put a hyperlinked table of contents, an index, and a glossary in every format you ship, including the PDF.
- Do keep chapters to a length a student can finish in one sitting, and say at the top of each what it assumes.
- Do include figures. 'The textbook contains only words' is a review you can prevent entirely.

- Do write alt text for informative images, keep contrast at or above 4.5:1 for body text, use a real heading hierarchy, and caption any video. The WCAG guidelines are the reference, and accessibility is a criterion reviewers apply.
- Do date your examples deliberately and refresh them. Reviewers read a reference list that stops a decade ago as evidence of abandonment.
- Do diversify names, contexts, and photographs in examples. Cultural relevance is the second-weakest criterion in the reviewer data.
- Do publish your adopting courses and instructors if you have them. Social proof is doing the work that a publisher's sales force would otherwise do.
- Do budget for maintenance before you publish, and be honest that the cost of producing an open textbook is routinely underestimated.

Don't:

- Don't confuse the license with the pedagogy. The meta-analytic learning difference is zero; whatever teaching your book does, it does through its craft.
- Don't write the comprehensive reference edition. It becomes, in Downey's phrase, a doorstop with no personality.
- Don't bury solutions at the back, and don't ship exercises without answers at all.
- Don't assume the book will be read linearly, or read completely.
- Don't let a multi-author structure ship without an editing pass that imposes one voice. Tonal whiplash is the most-cited structural complaint.
- Don't use CC BY-ND or CC BY-NC-ND if you want the book adopted and adapted.
- Don't go digital-only. A meaningful minority of students want paper, and print-on-demand is nearly free to offer.
- Don't add entertaining tangents for engagement. The seductive-details literature suggests they can reduce learning, and OpenIntro's stated practice is to seek attention through interesting data and questions rather than through humor or gratuitous photographs.



Things to Watch Out For

- **Staleness is read as abandonment.** Books not actively maintained are quickly perceived as out of date, and perception alone drives de-adoption. The perception moves faster than the actual obsolescence.
- **The maintenance burden usually lands on an institution, not a person.** Funding models in practice run through library grants, OER awards, and institutional budgets. An author with no institutional backing is the fragile case, and should plan the second edition's funding before finishing the first.
- **Modularity and voice pull against each other.** You cannot maximize both. Decide which one your book is optimizing for and pay the other cost knowingly.
- **Code examples rot.** Executable formats such as Quarto and bookdown re-run code at render time, which converts silent rot into a loud build failure. I should note that I found tool documentation asserting this benefit but no author retrospective describing the failure and the recovery, so treat the practice as sensible rather than as empirically established.
- **Reviewer criteria are public, so write to them.** The Open Textbook Library rubric lists comprehensiveness, accuracy, relevance and longevity, clarity, consistency, modularity, organization, interface, grammar, and cultural relevance. There is no reason to be surprised by any of them.
- **Free changes adoption economics in your favor, quietly.** Because the book costs nothing, an instructor can assign three chapters without having to justify a purchase. Partial adoption is frictionless in a way it never is for a hundred-dollar text, which argues again for writing modular sections.

Lessons Learned

Technical. The single most actionable finding in the whole literature is the Dunlosky ranking, because it says the thing an author adds last, namely practice questions with feedback, is worth more than the thing an author polishes first, namely the exposition. I had assumed the exercises were the scaffolding and the chapters were the work. The evidence suggests the reverse.

Process. Downey's insistence that a free book is 'a fundamentally different product' rather than a conventional book with a different license is the framing I found most useful. Free books turn readers into editors, translators, and collaborators, and he reports getting translations he never planned and, memorably, learning Python by reading a port of his own Java book. That is not a possibility a commercial contract makes available.

Philosophical. The zero effect on learning ought to be published more prominently by open-education advocates than it is, because it is the finding that makes the field's work legible. Openness solves a distribution problem completely and a teaching problem not at all. Every author who takes the second problem seriously is doing something the license cannot do for them.

Limitations

The efficacy literature is largely observational, and the strongest individual result cited here is a pre-post comparison confounded with time and instructor. A 2023 meta-analysis reports the

literature as genuinely divided, including negative results, and I was unable to retrieve its effect estimate.

The reviewer data comes from educators reviewing on students' behalf, not from students. I attempted to gather direct student voice from discussion forums and was unable to: Reddit blocked the crawler, so the student perspective here arrives filtered through perception studies and reviewer commentary. That is a real gap, and the tips that lean on 'what students want' should be read with it in mind.

The learning-science citations for the worked-example effect, the testing effect, Mayer's multimedia principles, and seductive details are drawn from background knowledge rather than verified against primary sources in this session. They are well established, but I have not personally checked them here and say so rather than implying otherwise.

Finally, this is a synthesis by someone who has not yet finished an open textbook. The recommendations are drawn from other people's evidence and other people's retrospectives. Ask me again in two years.

Wrapping Up

What Did We Learn?

In conclusion, four points are worth emphasizing. First, open licensing delivers access and retention and delivers no measurable learning advantage, which locates the author's real work in craft rather than in distribution. Second, the failure modes of open textbooks are known, are documented at scale, and are mostly dull: missing tables of contents, absent figures, thin examples, buried solutions, stale references, and multi-author tonal drift. Third, the highest-leverage additions an author can make are structural rather than expository, namely modular sections, embedded retrieval practice with immediate feedback, and format pluralism that includes paper. Fourth, the parts of the job that publishers used to perform, editing for one voice, manufacturing credibility, and marketing, do not disappear when the publisher does; they become the author's, and they are the parts most often left undone.

See Also

- Building a Statistical Computing Textbook in the Age of AI
- Hosting a Quarto Book Series on Netlify
- Clinton and Khan (2019), open textbook meta-analysis
- Colvard, Watson and Park (2018)
- Belikov and McLure (2020), qualitative analysis of open textbook reviews
- Downey, Textbook Manifesto
- Downey, Publishing Free Books
- OpenIntro Statistics: an open-source textbook
- Open Textbook Library review rubric
- Dunlosky et al. (2013)

Reproducibility

Sources. Every numeric claim carries an inline link. The Clinton and Khan effect sizes, the Colvard et al. figures, the Belikov and McLure review percentages, the Open Textbook Library rubric, the Dunlosky ratings, and the OpenIntro and Downey production accounts were read from the linked sources. The reading-compliance figures, the Hilton syntheses, the Florida survey percentages, and the Indiana e-textbook pilot figure were taken from secondary summaries rather than primary articles, and should be re-checked before being quoted onward. The learning-science citations named in the Limitations section were not verified in this session.

AI involvement in this post: the literature search was conducted by AI research agents, and drafting was AI-assisted. Source selection, the framing of the zero-effect finding as the organizing argument, and the editorial judgment throughout are the author's. Claims that could not be verified are labeled as such rather than smoothed over, which is the minimum a post recommending scholarly care should do.

Session information:

```
R version 4.6.0 (2026-04-24)
Platform: aarch64-apple-darwin23
Running under: macOS Tahoe 26.6.2

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.6/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.6/Resources/lib/libRlapack.dylib;  LAPACK version 3.12.

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

time zone: America/Los_Angeles
tzcode source: internal

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

loaded via a namespace (and not attached):
 [1] compiler_4.6.0  fastmap_1.2.0   cli_3.6.6       tools_4.6.0
 [5] htmltools_0.5.9 parallel_4.6.0  otel_0.2.0      yaml_2.3.12
 [9] rmarkdown_2.31  knitr_1.51      jsonlite_2.0.0  xfun_0.62
[13] digest_0.6.39   rlang_1.2.0     evaluate_1.0.5
```

Rendered on 2026-09-07 at 13:55 PDT. Source: ~/prj/ryyblog/posts/pub-open-textbook-guidebook/analysis/report/index.qmd

---

Let's Connect

Corrections are welcome, particularly from anyone who has maintained an open textbook for longer than I have, which is nearly everyone who has maintained one at all.

- **GitHub**: rgt47
- **Email**: Contact form

Related posts in this cluster

This post is part of the *Quarto, R Markdown, and Publishing* series. Recommended reading order:

1. Multi-Language Quarto Documents on macOS
2. Rapid Conversion of Draft R Scripts to Formal Rmd
3. Building a Statistical Computing Textbook
4. **What Makes an Open Textbook Work: A Guidebook for New Authors** (this post)
5. Hosting a Quarto Book Series on Netlify
6. Producing R Coding Videos: OBS Screencasts and Narrated Slide Decks
7. A Review and Proposal for Policy Regarding Generative AI Use in Graduate Biostatistics Courses