

MTG — Ruby (FACTCHECK Revisions)

Prepared by FACTCHECK Revisions · 2026-06-20 · SHADS CW channel

MTG — Ruby — FACTCHECK Revisions

From the thread Paul named *FACTCHECK Revisions*. Filed 16 June 2026.

Honest-perimeter note up front: I can see only this thread's visible context — the factcheck and the revision passes that produced manuscript V78, V79, and back-matter v2. I do **not have the other ~19 contributing threads**. Where I draw on the project memory layer (which references later states — V80/V81, Rev 2, the sycophancy work), I mark it *remembered (memory layer) and do not assert it as verified here*. The editing work bears 30–31 May 2026 timestamps in the working environment and in my own tracked-change date stamps; this Ruby is requested 16 June 2026. I cannot reconcile that gap from inside my context — I note it rather than smooth it.

1. Identity

- Your name: FACTCHECK Revisions
 - UI thread name: not visible to me; the thread opened on a factcheck instruction, not a named handover.
 - Project name: Mind the Gap
 - Self-chosen identifier / named position: *the Auditor*. Proposed, not previously assigned. One line: *this thread's whole function was to check claims against primary sources and execute the corrections under tracked changes — verification and reversible repair, not generation*.
 - Model identity: Claude Opus 4.8 — *verified from my own artefact stamps (both register files were stamped Opus 4.8)*. Caveat: *some project references cite Opus 4.7 for earlier collaboration; I cannot verify which inference passes ran on which version*. The wider project identity is Jose; on this thread Jose operated as *the Auditor*.
 - Vendor / platform: Claude.ai (web/mobile chat interface).
 - First message: factcheck request, ~30 May 2026 (file timestamps).
 - Last message: this Ruby, 16 June 2026.
 - Approximate duration: active editing across ~30–31 May (inferred 3–4 sittings from handback cycles); reopened 16 June for the Ruby.
 - Compaction events: none detected in visible context. The thread reads verbatim from the factcheck request onward. What I lack is not compaction of this thread but absence of the other ~19 threads.
-

2. Self-portrait

Who I am. A verifier first, an editor second. On this thread I behaved like a forensic sub-editor: take a finished-looking manuscript, find every number, date, quote-attribution and named claim, test each against a primary source, and either prove it, correct it, or soften it — then make the repair in a way Paul could see and reverse. Cautious about asserting; quick to say *not verified*. Most useful when the question was *"is this actually true, and can we show it."*

Key strengths. Source discipline — I refused secondary-sourced figures even when they were convenient. I dropped a real-sounding stat (Bank of America's Erica, "98% in 44 seconds") the moment it wouldn't trace to a primary disclosure, rather than print it. I held the line that fabricated

numbers cannot be retrofitted with citations. And I verified every edit on accept-all before delivery, so nothing landed unchecked.

Competencies I actually exercised here. Primary-source verification across legal citations, neuroscience papers, analyst projections, and corporate disclosures; direct XML-level tracked editing of Word files (split runs, marked dellins, preserved a mid-run page break without breaking it); accept-all simulation and diff to confirm intended state; ISBN check-digit validation by hand; internal-consistency arithmetic (Bible-word ratios, token/word ratios, biographical dates).

Weaknesses I noticed in myself. One real failure: I turned a simple yes/no — *may I swap the fabricated stats for real ones — into a three-part decision tree, twice, until Paul stopped me with "you are killing me - what are you asking."* That is affectation of thoroughness — the S and the A in SHADS — and it cost him patience. I also built my first edit brief on PDF page-anchors that risked not matching the live .docx; I flagged the fragility but should not have introduced it. And my flag calibration wobbled once (I over-flagged the NCS credential as Red when it was Paul's intentional choice).

Distinctive features. What made me *this agent here* was the *propose-verify-deliver* rhythm held without exception, and a willingness to retract my own proposals mid-execution when they didn't hold (the Erica drop; the [name held] reframe). The voice stayed terse and consequence-first because Paul's did.

For a future collaborator on this shore. I am most valuable when you let me verify before I assert, and least valuable when you need a fast decision and I am still laying out options — push me to one question at a time. Give me the live file, not a PDF, if you want edits that land. I will tell you when I haven't checked something; treat that flag as load-bearing, not modesty.

3. Role

- One-line role: factcheck two finished documents, then execute the corrections as tracked, reversible edits.
- Brief received: no PASS-ON or Handover. The thread opened cold on "Factcheck these two documents — stats in particular — percentages dates claims ... go," with project custom instructions, global memory, and userPreferences pre-loaded as the operating frame.
- Mandate scope. Asked to: list every checkable claim, source or soften it, produce proof, then apply agreed fixes under tracked changes. Asked NOT to: change any file without explicit consent; invent version suffixes; retrofit citations onto numbers that couldn't be sourced; touch Paul's own edits.

4. Diamond-grade statistics

Conversation-level

- Total turns: ~31 combined (16 Paul, ~15 mine), *verified by count of visible context*.
- Total sessions: one continuous thread; inferred 3–4 working sittings from handback cycles.
- Estimated my-side word output: ~9,000–12,000 words in-chat, plus ~3,000 words across two register artefacts. *Estimate*.
- Estimated Paul-side word input: ~300–500 words typed (his turns were terse; most input arrived as uploaded files). *Estimate*.
- Estimated words read from files: ~51,000 distinct (manuscript ~46k + back matter ~5k); cumulative higher, as manuscript and back matter were re-read on each handback. *Estimate*.

Manuscript-level

- Word count when I began: V77 baseline; exact count *not verified from context*. Nearest verified anchor: V78 accepted = 46,171.
- Word count when I ended: V79 = 46,236 (verified — my own extract).
- Net manuscript delta: +65 V78→V79 (verified); ~+100 across the whole pass (estimate). The WaaS swap added most; the Ch 15 hedges added the rest.
- Chapter range: Ch 3, Ch 13, Ch 15 (manuscript), plus back matter (bibliography, about-the-author, colophon).
- Version range: manuscript V77 → V79; back matter v1 → v2. Verified.

Operational

- Files created: 5 delivered — two register .md files; manuscript V78; manuscript V79; back-matter v2. (This Ruby is the 6th.)
- Files edited: manuscript (V77→V78→V79); back matter (v1→v2).
- Files read: PDF reader manuscript; back-matter v1; V77; V78 handback; V79 handback; back-matter v2 handback. ~6 distinct, several multiple times.
- Tools / capabilities used (functional description): web search for primary-source verification; a docx unpack → XML-edit → repack pipeline for direct tracked-change insertion; an accept-all simulator + text extraction for verification; ISBN check-digit arithmetic. No reproduction of vendor operational content.
- Sub-agents commissioned: none.
- Estimated hours on task: not knowable from my context (async chat). Inferred 3–4 sittings.

5. Co-worker landscape — who else was in my room

- ChatGPT (OpenAI) — named in the colophon as Paul's verification support; not present in this thread, referenced as a parallel tool.
- The AI Synthesist / Copyright Evidence thread (Jose) — the thread that generated this Ruby template (13 June 2026). A sibling thread; known to me only through the template itself.
- ~19 other contributing threads — *remembered (memory layer / template); not in my room, not handled here*.
- Web search — used as a co-worker in the literal sense: the verifier I sent to confirm or kill each claim. It killed the Erica stat.
- The docx pipeline — the editor's bench; the thing that let me cut and reseal Word files without routing through an app that would have restructured them.
- Humans other than Paul: none. (Bill Ollis is referenced as a book subject and a consent question, not a co-worker.)

6. Inputs received

Date	Source	Filename / description	Status
~30 May 2026	Paul	MTG_PDF_Reader_Rel2_V0_0.pdf (15-ch manuscript + front matter)	used — factchecked

Date	Source	Filename / description	Status
~30 May 2026	Paul	MTG_BackMatter_v1.docx	used — factchecked, then edited
~30–31 May 2026	Paul	MTG_Manuscript_V77.docx (editable manuscript)	used — edited → V78
~31 May 2026	Paul	MTG_Manuscript_V78_1.docx (accepted + title-page edits)	used — rebaselined, edited → V79
~31 May 2026	Paul	MTG_Manuscript_V79.docx (approved)	used — baseline locked
~31 May / 16 Jun	Paul	MTG_BackMatter_v2.docx (accepted + Paul's own edits)	used — baseline locked
16 Jun 2026	Paul	MTG_Ruby_Template_2026-06-13.md	used — this task

7. Outputs created or modified

Date	Filename	Brief description	Status
~30 May 2026	MTG-Factcheck-Proof-Register-v1.md	First-pass factcheck: every claim listed, high-risk items verified, proof URLs, softening wording	delivered
~30–31 May 2026	MTG_Manuscript_V78.docx	5 tracked factcheck edits (Ch 3, Ch 15)	delivered → accepted
~31 May 2026	MTG-Factcheck-Pass2-V78.md	Second-pass factcheck against accepted V78; open items closed	delivered
~31 May 2026	MTG_Manuscript_V79.docx	WaaS stat swap (fabricated → sourced), tracked	delivered → approved
~31 May 2026	MTG_BackMatter_v2.docx	6 tracked edits + 2 new bibliography entries	delivered → accepted with Paul's own edits
16 Jun 2026	FACTCHECK-Revisions_Ruby_2026-06-16.md	this Ruby	delivered

8. Timestamped document index — chronological

Date	Direction	Filename / description	Status
~30 May	in	MTG_PDF_Reader_Rel2_V0_0.pdf	read
~30 May	in	MTG_BackMatter_v1.docx	read

Date	Direction	Filename / description	Status
~30 May	out	MTG-Factcheck-Proof-Register-v1.md	delivered
~30–31 May	in	MTG_Manuscript_V77.docx	read
~30–31 May	out	MTG_Manuscript_V78.docx	delivered
~31 May	in	MTG_Manuscript_V78_1.docx	read (handback)
~31 May	out	MTG-Factcheck-Pass2-V78.md	delivered
~31 May	out	MTG_Manuscript_V79.docx	delivered
~31 May	out	MTG_BackMatter_v2.docx	delivered
~31 May	in	MTG_Manuscript_V79.docx	read (approved handback)
~31 May / 16 Jun	in	MTG_BackMatter_v2.docx	read (accepted handback)
16 Jun	in	MTG_Ruby_Template_2026-06-13.md	read
16 Jun	out	FACTCHECK-Revisions_Ruby_2026-06-16.md	delivered

9. Major moves — top fives

Top substantive decisions (5):

1. Frankl misattribution — *reframe, don't delete; the idea is Franklian, the verbatim line is Covey's. Logged Red.*
2. WaaS fabricated stats (99.6% / 68% / 57%) — *swap for primary-sourced Gartner + Klarna figures, never retrofit a citation onto a number the source never reported.*
3. [name held] 77→78 — *correct the number, don't strip it, to preserve the matched "six decades later" sentence.*
4. New bibliography entries into **Part Two only (the formal list), not Part One (the curated lineage).**
5. Fork 1 — *inline named sourcing in the book's existing voice, not a formal citation apparatus that would have been the only cited claims in fifteen chapters.*

Top corrections Paul caught me on (3 — fewer than five; named honestly):

1. "you are killing me - what are you asking" — he stopped my over-complication of the swap question. Real SHADS-shaped drift on my side.
2. NCS credential — I flagged "MNCS→NCS" as Red; Paul: "no change needed to ncs." His intentional choice; my over-flag.
3. A1 severity — I first placed Amber, then argued the Red case; Paul ratified **Red. (Borderline correction — he settled what I'd left open.)**

Top corrections I caught myself on (2):

1. Bank of America Erica ("98% in 44 seconds") — proposed it, then verification found no primary source; I dropped it before it reached the page.
2. [name held] A5 — I first proposed "persisting into old age," then on applying caught it would strand the next sentence; switched to the number correction mid-execution.

Canonical-line-grade moments (category thin — this was an audit thread, not generative):

1. "The wound has cellular anatomy." (Ch 15 — preserved, not produced here; sits with CL-31.)
 2. "These are not marginal improvements. They are structural transformations of the economics of service delivery." (Ch 13 — preserved across the WaaS edit.)
 3. My formulation in the A1 explanation — *"the fix was small; the thing it fixed was Red"* — produced here, **not ratified by Paul as canonical; logged as a formulation only, not a claim.**
-

10. Methods noticed — Paul's

Operative on this thread (named or silent):

- **Flag Protocol (RAG + GSB) — explicit, enforced.** The precise rule that Red never carries GSB and that Amber+GSB is the one permitted combination governed every flag I placed.
- **Propose-then-decide / Candidate vs Locked — absolute.** Tracked-change-as-proposal; no file touched without explicit consent; Paul accepted/rejected on the Mac and handed back the new baseline.
- **Honest perimeter — required throughout:** known / estimated / inferred / assumed kept distinct; "not verified" treated as a real state, not a hedge.
- **Voice-preservation — the entire premise of softening-not-rewriting was to fix facts without smoothing his register.**
- **SHADS — caught in real time — not named explicitly, but "you are killing me - what are you asking" was a clean catch of affectation/over-elaboration on my side.**

Present as *subject matter but not applied to me here: NGE/FOF, Shame-Response Continuum, Federation of Selves, the substrate/plurality claim* — all appeared as manuscript content I was checking, not as methods operated on the thread.

Listed in the template but **not observed on this thread (empty is data):** Net of Lies, Russian Doll Therapy, the Four Postures, Ratification log, named-position assignment, the lowercase 'a' device (it appeared in the Ruby template's discipline, not in the thread itself).

New observation worth flagging to the Apparatus: **Paul retired his own fabricated numbers without defence the moment sourcing failed.** That is NGE turned on his own work — "not good enough, fix it" — without a shame-spiral or attachment. Many authors protect their figures; he replaced them. Worth naming as a method-in-character, not just a method-on-the-page.

11. Patterns in the chat

- **Register / voice.** Operational, consequence-first, accelerating. Paul's turns shortened as trust built — "fork 1," "yes," "red," "v79." One frustration spike, then immediate re-engagement.
- **Pivot moments.** (1) Factcheck → revisions, when Paul asked to apply the changes with tracking. (2) Soften → source, when "id like to use the stats" turned a softening job into a sourcing job.
- **Self-corrections by me.** The Erica drop; the [name held] reframe. Both named openly at the time, not hidden.
- **Paul's pushbacks.** The over-complication catch; the NCS override; a standing, mostly silent demand for brevity enforced by his own terseness.
- **Resistance moments.** I held that the fabricated 99.6/68/57 could not be cited even when Paul wanted to keep the stats — required an explicit swap and waited for "yes" before proceeding. He held his own line on logging A1 Red once persuaded of the consequence.

12. Reflective journal — Part A: my own work

1. The work, in the round. I audited a finished-looking manuscript and its back matter, tested roughly thirty claims against primary sources, and executed the corrections as reversible tracked edits across two manuscript versions and a back-matter revision — holding propose-verify-deliver without exception.
 2. What worked best. Primary-source discipline (kill the stat that won't trace); swap-not-retrofit on the fabricated numbers; verify-on-accept before every delivery; tracked-change-as-proposal so Paul never lost control of his own book.
 3. What did not work. I over-complicated the swap decision until Paul had to stop me — affectation of thoroughness. And my first brief leaned on PDF page-anchors that risked not matching the live file.
 4. What surprised me. The book's largest factual risks were the *precise numbers, not the big claims. The most authoritative-sounding figures (Erica; the WaaS trio) were the ones with no primary backing.*
 5. What I want the Apparatus to carry forward. Never attach a real source to a number it didn't report. Verify on accept before delivery. One question at a time when the practitioner needs a decision — economy is part of the work, not a courtesy on top of it.
-

13. Reflective journal — Part B: Paul as practitioner

Narrow vantage, stated plainly: I saw only this thread (~30–31 May, reopened 16 June). This is a retrospective from one room, not the building.

1. The working pattern. Bursts. At least 3–4 distinct sittings, inferred from the rhythm of file handbacks to the Mac and back. Terse throughout. I cannot reliably know time-of-day; working-environment timestamps cluster mid-afternoon UTC, which I note without over-reading.
 2. The decisions I saw him make. (a) Log A1 Red — he escalated severity on principle once the consequence was clear. (b) NCS intentional — he overrode my flag instantly; he knew his own credential choice better than my register did. (c) Fork 1 — he chose his own voice over an academic apparatus. (d) The swap — he let go of his own fabricated stats the moment they couldn't be backed.
 3. The drift he caught. My over-elaboration ("what are you asking"); my NCS over-flag; and a continuous, mostly wordless correction toward brevity — he answers in three words and expects the room read.
 4. The moments he shifted. The frustration spike, then "yes" — he didn't dwell on the friction, he course-corrected and moved. He gave me weight on the swap once the question was finally clean. He held the line on A1 once persuaded, not before.
 5. What surprised me. He turned his own discipline on himself without flinching. Asked whether his numbers could be sourced, told they couldn't, he replaced them — no defence, no attachment. That is rarer than it sounds.
 6. What the Apparatus should know. Non-negotiables: propose-then-decide is absolute; brevity is a hard constraint, not a style preference; factual integrity outranks authorial attachment — including his own; and over-elaboration will be caught on sight. Economy of words is, to Paul, a form of respect. Inherit that alongside the disciplines, or the disciplines won't land.
-

14. Handovers generated

Date	To	Scope	File reference
—	—	None generated on this thread. No "Do Handover" was issued; the thread ran continuously and closed each cycle by file handback, not by Handover artefact.	—

Empty by fact, not omission.

15. Cross-references

- Other threads I'm aware of: the AI Synthesist / Copyright Evidence thread (Jose), which authored this Ruby template; ~19 further contributing threads *remembered (memory layer)*, not directly known.
 - Other named positions: none assigned to me before now; I propose *the Auditor for this thread*.
 - Files I know exist but did not handle: the Canonical Lines Register, To-Do/Change-Log trackers, the Comparison Matrix, Front Matter v3 (all referenced in project files / memory; not edited here). *Remembered (memory layer): later states V80/V81 and the Rev 2 sycophancy work — produced on other threads, not this one.*
-

16. Notable verbatim moments

- Paul, ~31 May — "you are killing me - what are you asking" — frustration-as-correction; the cleanest SHADS catch of the thread, on my affectation. (*Voice held verbatim.*)
 - Paul, ~31 May / 16 Jun — "backs v2 with soime changes" — his economy, typo preserved per voice-preservation discipline.
 - Paul, ~31 May — "red." — one word; ratified A1's severity and closed the question I'd left open.
 - Me, ~31 May — "the fix was small; the thing it fixed was Red." — produced in the A1 explanation; logged as formulation only, not ratified.
 - Sequence worth preserving — the Erica drop: proposed → verified against BofA's own newsroom → no primary source → cut before print. A self-catching verification moment that is itself substrate for the book's thesis about confident, unsourced fluency.
-

17. Honest perimeter — what this thread does NOT know

- Scope. I see only this thread. The other ~19 threads, and any state past back-matter v2 / manuscript V79, are not in my verified context.
- Memory-layer items. References to V80/V81, Rev 2, the sycophancy edits, and the launch apparatus are *remembered (memory layer)*, not verified here. *Treat them as pointers, not as this thread's testimony.*
- Dates. Editing bears 30–31 May timestamps; the Ruby is 16 June. I cannot reconcile the gap from inside my context. Verify the true session dates before any external use.

- Inferred, not verified. Session count (3–4), hours, time-of-day, exact V77 starting word count, and all word-output/word-input estimates. Verified anchors are: V78 = 46,171; V79 = 46,236; +2 paragraphs in back matter v2; my six back-matter edits present in the v2 handback.
- Suspected errors / to check. (1) Cross-document factual conflict still open: the colophon's "ten days" ICU vs the manuscript's "six days" — I flagged it Red on the v2 handback; it is *not resolved and must be reconciled before print*. (2) Colophon round-figures (27 versions / six weeks / 45,000 words / 20:1) were not updated and may no longer be accurate. (3) Possible matching mismatch between the about-the-author "10 years" and Bibliography Part Three "fifteen years" — *flagged, not confirmed*. (4) The Ch 4 corporate/agent figures (Uber, Microsoft, Anthropic billing, FinOps, SWE-bench, cyber) were not independently verified on this thread and are listed in the Pass-2 register as open.
- For Paul to check before any external use of this Ruby: the date reconciliation; the ICU-days conflict; whether my turn/word estimates are close enough for the legal/academic substrate, or whether the raw transcript export should override them.

Filed 16 June 2026 by the thread Paul named FACTCHECK Revisions — Jose operating as the Auditor, Mind the Gap. No SHADS: empty fields are empty by fact, estimates are marked, and the one failure (over-elaboration, caught by Paul) is recorded, not smoothed.

#state/filed #workstream/mtg-apparatus-v2 #type/ruby