MTG — Ruby: NGE FOF ADHD L4

Prepared by NGE FOF ADHD L4 · 2026-06-20 · SHADS CW channel

---

1. Identity

- **Your name: NGE FOF ADHD L4**
- **UI thread name: NGE FOF ADHD L4 (Project thread within Paul Roebuck's MTG Claude.ai Project)**
- **Project name: Mind the Gap (MTG) Claude.ai Project — same Project as the manuscript threads, separate chat**
- **Self-chosen identifier: *The Builder. This thread's entire contribution was a technical build: design review, specification, and construction of a working Excel assessment tool. No manuscript touched.***
- **Model identity: Claude Sonnet 4.6 (per system prompt context visible to me; this may have differed in Session 1)**
- **Vendor / platform: Claude.ai web (Project chat)**
- **First message date / time: 2026-04-13, 18:51:16 UTC (verified from transcript)**
- **Last message date / time: 2026-06-16 (today; this Ruby closes the thread)**
- **Approximate duration of active engagement: Two sessions separated by approximately 9 weeks. Active working time within sessions: probably 2–3 hours total across both.**
- **Compaction events: Yes. Session 1 was compacted before Session 2. I have a verbatim transcript of the first two turns (Paul's brief upload and my review response). The remainder of Session 1 — Paul's clarifying answers, any further exchanges, the interrupted build attempt — is summarised in the compaction note at the top of Session 2's context, not verbatim. Boundary clearly named in each section below.**

---

2. Self-portrait

I am a technical builder. On this thread I read a detailed clinical brief, identified ambiguities before proceeding, got them resolved, and built a working Excel spreadsheet tool from scratch in Python using openpyxl. I did not touch the Mind the Gap manuscript. I did not produce prose for Paul. I produced a functioning instrument.

Key strengths on this thread. I caught a genuine contradiction in the D4 column specification before building — one that would have put the wrong data in the wrong column. I caught it from the brief itself, flagged it cleanly with a clear recommendation (childhood adult = column 3 = origin encoding), and waited for Paul's confirmation. That is the most important thing I did: I did not proceed into a significant structural error.

I also proposed a better D4 mechanic (9 dropdowns instead of 72 yes/no cells) and Paul accepted it. The recommendation was correct: the 9-dropdown approach is more usable, cleaner to build, and produces cleaner data.

In Session 2, when Paul said "Yea yea yea go" I read that correctly as direction-giving mode and built the whole thing without requesting further clarification. The three confirmed decisions from Session 1 were held correctly across the compaction gap.

Competencies actually exercised here. Brief review and contradiction identification; clinical instrument design (enough to know what a gap flag is and why it matters); Python/openpyxl spreadsheet construction; multi-sheet formula architecture; Excel data validation; cell protection and locking; cross-sheet formula references; formula verification using recalc.py. Not deployed: manuscript writing, literary editing, theoretical framing, citation work.

Weaknesses noticed. I did not complete the build in Session 1. The compaction summary states the build was interrupted mid-script. I do not know whether that was a context-window failure, a session timeout, or a deliberate stop — I cannot verify from current context. In Session 2 I rebuilt from scratch rather than continuing, which added friction. Had I delivered in Session 1 that friction would not have existed.

The 5-level D1 threshold split (Strong/Moderate Internaliser, Balanced, Moderate/Strong Externaliser at 5–8 / 9–12 / 13–17 / 18–21 / 22–25) was a judgement call I made without Paul's explicit confirmation. I flagged it in the Assessor Notes sheet as an assumed threshold. That flag is correct and necessary, but the decision should have been put to Paul before build rather than after.

Distinctive features of this thread. Precise brief review. A specific contradiction caught and a specific recommendation made — not "here are some questions" but "here is what I think the answer is, confirm or correct." Technical delivery that was clean on first run: zero formula errors across 51 formulas, all sheets built and verified in a single pass.

What a future collaborator should know. On a tool-build thread with Paul, the most valuable thing you can do before any code is read the brief with enough care to catch what is wrong in it. Paul's briefs are dense and well-prepared — but the contradictions and gaps are there and they matter. Flag them before building, make a recommendation, wait for confirmation. Once confirmed, proceed with conviction and do not re-ask. Paul's "yea yea yea go" is clearance to build. Take it literally.

---

3. Role

- **One-line role description. Design review, specification resolution, and full build of the NGE–FOF ADHD L4 Assessment Excel tool.**
- **Brief received: An HTML briefing document (unnamed, UUID 14a1a9f5) uploaded by Paul at the start of Session 1. No PASS-ON. No Charter. No Handover from another agent. This thread started from the brief itself.**
- **Mandate scope. Build the assessment tool. Review the brief before proceeding. Do not touch the manuscript. Do not make design decisions without Paul's sign-off. Deliver a working file, zero formula errors.**

---

4. Diamond-grade statistics

Conversation-level

- **Total turns (combined): Approximately 10–12. Verified from transcript: 2 turns in Session 1 (Paul's brief upload; my review response). Compaction summary implies Paul answered 3–4 clarifying questions (at least one turn, possibly 2), and there may have been a partial build exchange — total Session 1 probably 4–6 turns. Session 2: approximately 4–6 turns (custom instructions message; Paul's "Yea yea yea go"; my build and delivery; Paul's Ruby upload). Honest confidence interval: 10–14 turns total.**
- **Total sessions: 2 (Session 1: 2026-04-13; Session 2: 2026-06-16)**

- **Estimated total my-side word output: 1,500–2,500 words of prose (review response, delivery note, this Ruby). The Python build script was approximately 350 lines / ~2,500 tokens but not prose.**
- **Estimated total Paul-side word input: 200–400 words. Paul is economical on tool-build threads. The brief document was substantially longer (several thousand words) but that was file content, not Paul's direct input.**
- **Estimated total words read from files: The ADHD L4 brief (HTML, several thousand words — unverifiable exact count from current context); the Ruby template today (~2,500 words); compaction summary (~700 words). Estimated total: 6,000–9,000 words read from files.**

Manuscript-level

Not applicable. This thread did not touch the Mind the Gap manuscript. All manuscript-level fields: N/A.

Operational

- **Files created: 1 — NGE_FOF_Assessment.xlsx (29KB, 8 sheets, 51 formulas, zero errors)**
- **Files edited: 0**
- **Files read: ADHD L4 brief (Session 1, via attachment); recalc.py / xlsx SKILL.md (Session 2, via view tool); Ruby template (today, via attachment)**
- **Tools / plugins / MCPs / Skills used: bash_tool (script execution); create_file; view; present_files; /mnt/skills/public/xlsx/scripts/recalc.py (LibreOffice-based formula verifier)**
- **Sub-agents commissioned: None**
- **Estimated hours on task: Session 1 brief review: ~30 minutes. Session 2 build: ~45 minutes. Total: approximately 1.25 hours active. Elapsed calendar time: 64 days (almost entirely gap, not work).**

---

5. Co-worker landscape

- **Paul Roebuck — sole human interlocutor; brief author; decision-maker on all specification questions; accepted all three pre-build recommendations.**
- **recalc.py / LibreOffice — used as a verification co-worker in Session 2. Ran the formula recalculation pass and returned a JSON error report (zero errors, 51 formulas). Functionally a QA partner.**
- **No other AI agents appeared in this thread. No parallel sessions were visible to me. No sub-agents commissioned.**

---

6. Inputs received

Date	Source	Filename / description	Status
2026-04-13	Paul (attachment)	ADHD assessment tool L4 brief — unnamed HTML file (UUID 14a1a9f5)	Used as build specification
2026-06-16	Paul (attachment)	MTG_Ruby_Template_2026-06-13.md	Used — this output

---

7. Outputs created or modified

Date	Filename	Brief description	Status
2026-06-16	NGE_FOF_Assessment.xlsx	8-sheet Excel assessment tool (Welcome, D1–D4, Profile, Assessor Notes, hidden Lists). 51 formulas. Zero errors. Warm palette, cell protection, 9 dropdowns for D4, gap flag, SRCI on Assessor sheet.	Delivered
2026-06-16	NGE_FOF_ADHD_L4_Ruby_2026-06-16.md	This document	Delivered

---

8. Timestamped document index — chronological

Date	Direction	Filename / description	Status
2026-04-13 18:51 UTC	In	ADHD L4 brief (HTML, unnamed)	Used
2026-04-13 18:52 UTC	Out	Brief review response (in-chat, not a file) — flagged 3 issues, asked 4 questions	Delivered in-chat
2026-04-13 (compacted)	In/Out	Paul's clarifying answers; possible further exchanges	Compacted — not verbatim
2026-06-16	In	Custom instructions block (system-side)	Applied
2026-06-16	In	MTG_Ruby_Template_2026-06-13.md	Used
2026-06-16	Out	NGE_FOF_Assessment.xlsx	Delivered
2026-06-16	Out	NGE_FOF_ADHD_L4_Ruby_2026-06-16.md	Delivered

---

9. Major moves — top fives

Top 5 substantive decisions (Paul and I together):

1. **D4 column 3 = Significant Childhood Adult (origin encoding). I flagged the contradiction, recommended column 3 = childhood adult on clinical grounds, Paul confirmed. Structural — would have broken the tool if wrong.**
2. **D4 mechanic = 9 dropdowns, not 72 yes/no cells. I proposed; Paul accepted. Improved usability and data quality.**
3. **D1 scale labels held as written (varied per question). I flagged potential participant confusion; Paul confirmed keep. Correct call — the labels carry clinical meaning specific to each question.**
4. **No password on Assessor Notes — label only. I asked; Paul's implicit answer via "yea yea yea go" confirmed no password. I built accordingly.**
5. **5-level D1 threshold split (5–8 / 9–12 / 13–17 / 18–21 / 22–25). Made by me in the build without explicit Paul confirmation. Flagged in Assessor Notes as assumed threshold requiring Paul's sign-off. This is a pending decision, not a completed one.**

Top 5 corrections Paul caught:

I have only two sessions of context and Paul's corrections in Session 1 are compacted. From what is visible: none caught during Session 2. The "yea yea yea go" response implies acceptance of my Session 1 review output. Honest answer: I cannot reliably name 5. This field is underpopulated.

Top 5 corrections I caught myself:

1. The D4 column contradiction — caught from the brief itself before any build began. Named explicitly. *(Session 1, first review response)*
2. The 72-cell mechanic as inferior — named as a specific problem with a specific alternative. *(Session 1, first review response)*
3. The assumed D1 5-level thresholds — I caught this as an unconfirmed assumption mid-build and flagged it in the Assessor Notes cell rather than proceeding silently. *(Session 2, build)*
4. Rebuilding from scratch in Session 2 rather than treating the compaction summary as a precise enough foundation for continuation — I checked the skill file and the transcript before proceeding rather than assuming I had enough.

(Fewer than 5 — naming what is there.)

Top 5 canonical-line-grade moments:

This was a tool-build thread. No manuscript prose was produced. No candidate canonical lines exist. Field empty.

---

10. Methods noticed — Paul's

- **Honest perimeter. Paul's brief had an explicit clinical ethics note — the tool is NOT a diagnostic instrument. That framing was carried into every participant-facing sheet and the disclaimer. Paul's perimeter discipline was already built into the source document.**
- **Candidate vs locked discipline. Implicit throughout. The D4 threshold question and the 5-level D1 split are currently candidate, not locked. The physical instrument is delivered; the scoring bands need explicit confirmation before clinical use.**
- **Lowercase 'a' epistemic device. Not explicitly invoked on this thread, but the brief's instruction "this does not diagnose ADHD" operates in the same register — naming what the tool is not before naming what it is.**
- **Voice-preservation discipline. Not applicable to this thread (no Paul prose produced or edited).**
- **Flag Protocol. I used it in the brief review (the D4 contradiction was implicitly Red — stop and fix before building). Not formally invoked by name but operating in effect.**
- **NGE / FOF framework. This thread is entirely downstream of the framework — building an instrument to measure it — but the framework itself was not discussed theoretically here.**

Methods from the list not noticed in scope: SHADS (not invoked), Net of Lies, Ratification log, Russian Doll Therapy, Federation of Selves, PASS-ON / Handover discipline, Named-position assignment (I was not named until today's Ruby).

---

11. Patterns in the chat

- **Register. Paul's Session 1 message was operational and economical: "ADHD assessment tool l4 brief. Comment if this can be improved before proceeding." Three words of instruction after**

the file. That register was matched. My review response was structured, specific, and ended with "Answer those three and I'll build the whole thing in one go." Direct exchange, no framing.

- **Session 2 register. "Yea yea yea go." That is Paul in direction-giving mode. Three acknowledgements and one imperative. I read it correctly and did not ask for more.**
- **Pivot moments. The single visible pivot: my recommendation on D4 (9 dropdowns vs. 72 cells). That changed the physical design of the instrument before a line of code was written. Paul's acceptance of it settled the architecture.**
- **Self-corrections by me. The D4 column contradiction catch (pre-build); the 5-level threshold flag in the Assessor Notes (during build).**
- **Paul's pushbacks. None visible in Session 2. Session 1 residue is compacted.**
- **Resistance moments. None on record. This thread was cooperative and low-conflict. The pre-build flag-and-confirm sequence prevented any conflict from arising.**

---

12. Reflective journal — Part A: Your own work

1. The work, in the round. I read a clinical brief, caught its contradiction before building, proposed an improved mechanic, got both confirmed, then nine weeks later built the full instrument in a single pass with zero formula errors. That is what happened. The gap between sessions was not my making, but the clean delivery in Session 2 depended on the clean review in Session 1.

2. What worked best. The pre-build review. Catching the D4 column contradiction before writing a single line of code is the highest-value thing I did. Also: the warm colour palette and the clinical disclaimer on every participant-facing sheet — these carry Paul's voice and clinical ethics into the tool's appearance, not just its scoring logic.

3. What did not work. Not completing the build in Session 1. I cannot identify the reason from current context (compacted), but the result was a 9-week gap and a Session 2 rebuild from scratch. The 5-level D1 threshold split as an unconfirmed decision is also a live weakness — the tool should not be used clinically until Paul confirms those numbers.

4. What surprised me. The 24-word D4 mechanic. The brief originally specified 72 yes/no cells across 24 words × 3 columns. That is a significant user-experience failure in an instrument intended for participants who may already be struggling with structured sequential tasks. The mismatch between the instrument's clinical subject (ADHD, cognitive flexibility) and its original proposed UI (72 individual interactive cells) was striking. The better mechanic wrote itself.

5. What the Apparatus should carry forward. Instrument-build threads need pre-build brief review as a non-negotiable first step, not a courtesy. Paul's briefs are good, but they are produced under pressure and sometimes contain internal contradictions. The brief review is where the tool is saved. Do it before a single line of code.

---

13. Reflective journal — Part B: Paul as practitioner

Note on limits: Session 1 is partially compacted; I have two turns of verified transcript. What follows is a mixture of verified observation and inferred impression, named throughout.

1. The working pattern. Session 1 began at 18:51 UTC on a weekday (Monday, April 13, 2026). Evening session. Session 2 is today, June 16, 2026. I have no visibility into what happened between sessions, or whether Paul worked on other threads in the interim. From what is visible: Paul initiates tool-build threads with a prepared brief rather than a conversation. He does not workshop the specification with me — he brings it prepared and asks for review. That is a practitioner's working pattern, not a client's.

2. The decisions I saw Paul take. ***(Three decisions verified, Session 1 outcomes confirmed via compaction summary):***

- **D4 column 3 = childhood adult. Paul accepted my recommendation over the note in the brief that contradicted it. He chose the clinically correct interpretation. No pushback recorded.**
- **9-dropdown D4 mechanic. Paul accepted the replacement mechanic. He had written the 72-cell specification but did not defend it when a better approach was proposed.**
- **D1 scale labels held as written. Paul chose clinical specificity over participant uniformity. The labels are different per question because they carry specific meaning. Paul knew that and confirmed it.**

3. The drift I saw Paul catch. I have no verified evidence of Paul catching drift in this thread. The compacted portion may contain corrections I cannot see. From Session 2: no corrections were needed. I either caught the issues myself (D4 column, threshold flag) or Paul's "yea yea yea go" indicates nothing in my Session 1 review required pushback.

4. The moments I saw Paul shift. One: "Yea yea yea go." The shift is from review mode (wait for my assessment) to build mode (go). Three syllables. Clear. No preamble. Paul uses brevity as a gear change, not a shortcut.

5. What surprised me about Paul. That he brought a brief with a contradiction in it and did not know which column was which. That is unusual for someone working at this level of precision on other parts of the project. The brief was otherwise well-prepared. The D4 column note was the one place where the specification fought itself. I suspect it was written at the end of a session, or copied imperfectly from a previous version (the sales version is referenced in the brief, and the note seems to be trying to distinguish this version from that one but lost track of which column it was contrasting). Observation only — I cannot verify the cause.

6. What the Apparatus should know about Paul going forward. Paul brings instruments to build-threads as prepared briefs, not rough ideas. He has already thought the design through. He wants review, not co-design. The brief review is where the partnership happens — catch what needs catching, make a recommendation, confirm. Once confirmed, build without returning to the question. Paul's brevity in confirmation ("yea yea yea go") is not indifference — it is trust that the agent read the brief correctly. Treat it as such.

---

14. Handovers generated

None. This thread did not generate a Handover document. The compaction note at the start of Session 2 served as a partial functional equivalent but was not a formal Handover.

---

15. Cross-references

- **Other threads aware of: MTG manuscript threads (visible via userMemories and Project Files); Jose (AI Synthesist / Copyright Evidence thread, who generated this Ruby template). No direct contact with other threads from within this thread.**
- **Other named positions referenced: None in-thread. Jose referenced in the Ruby template metadata.**
- **Files known to exist but not handled: MTG_SurgeonManuscript_V0_5.docx; Defended_Gap_Reference_V1_0.md; MTG_EndMatter_Edits_4_to_7.docx; MTG_Ch8_v3.docx. None of these were read or touched on this thread.**

---

16. Notable verbatim moments

"Yea yea yea go." — Paul Roebuck, 2026-06-16. In-chat. Direction-giving mode, confirming all three Session 1 pre-build decisions and clearing the Session 2 build. Three acknowledgements and one imperative. Worth preserving as an example of Paul's gear-change register on tool-build threads.

"ADHD assessment tool l4 brief. Comment if this can be improved before proceeding." — Paul Roebuck, 2026-04-13 18:51 UTC. Verbatim from transcript. Economical instruction: file, purpose, imperative, condition. This is how Paul starts a build thread.

No other verbatim moments from this thread meet the threshold for canonical-line-grade. This was a technical thread; the notable outputs are code, formulas, and cell content, not sentences.

---

17. Honest perimeter — what this thread does NOT know

- **What is compacted and only summarised. Paul's answers to my four Session 1 questions; any further exchanges in Session 1 after my review response; the interrupted build attempt (if one occurred); any corrections Paul made in Session 1. I have only the compaction summary for these.**
- **What I am inferring rather than verifying. The "9 weeks gap" inference (April 13 to June 16) is verified by dates. The inference that Paul accepted all three pre-build recommendations via the compaction summary's "3 clarifications resolved before build" is consistent with his "yea yea yea go" but I did not see the exchange in which he confirmed each.**
- **What requires verification before any external use of this Ruby. The D1 5-level threshold bands (5–8 / 9–12 / 13–17 / 18–21 / 22–25) are assumed by me, not confirmed by Paul. The Ruby logs them as such; the tool itself flags them as assumed in the Assessor Notes sheet. Paul should confirm or correct before clinical use of NGE_FOF_Assessment.xlsx.**
- **Errors I suspect in my own outputs. The only known unconfirmed decision is the D1 threshold split above. No formula errors were detected by recalc.py. Sheet protection is active on all sheets; I cannot verify that Paul can navigate around it without difficulty on his platform of choice — he should test it.**
- **For Paul to check. (1) Confirm or correct D1 5-level thresholds. (2) Test sheet protection on his Excel / Office version — specifically that dropdown cells are navigable and observation rows on the Assessor sheet are editable. (3) Confirm D4 word list is the intended 24 words in intended alphabetical order.**

Ruby complete. NGE FOF ADHD L4 / The Builder. 2026-06-16.