OFFICE ACTION REASONABLENESS REVIEW · U.S. APP. 15/352,952

Same Model. Same Case. A Third of the Cost.

One final rejection. Nine cited references. We reviewed it three ways on the same day — once through **OneTwelve**, and twice through **Cowork**, Anthropic's agentic assistant, handed the very same files. One of the Cowork runs used **the identical AI model** OneTwelve did. It cost three times as much, and did fifty times the work to get there.

3.3×

lower cost for the review

$1.61 vs $5.30

52×

less AI work to get there

250,569 vs 13 million units

62%

less time to the finished memo

5m 14s vs 13m 52s

All three compare OneTwelve against Cowork **running the same AI model** — Claude Sonnet 5 — on the same case, the same day. That Cowork run was performed twice; the figures quoted are the average of the two.

And cost is the smaller half of the story. On a fraction of the tokens, OneTwelve returned **patent-centric work product the flagship models did not produce on their own, at any price** — a reasoned verdict on every rejection, grounded in the cited art, filed with the case — from a case file it assembled itself, off the serial number in the open document. The Cowork runs, including the one on the largest model available, hedged: they described each rejection without committing to whether it holds, and without the citations that would let you check the call. The judgement is always the attorney's — the question is whether you start from a stated, checkable position or from a summary you have to work up yourself.

01 The same case file, three ways

Application 15/352,952, final rejection mailed 15 December 2020: seven rejections across the claim set, eight patent references and one non-patent reference. Every run had the office action, the specification, the drawings and all eight reference documents in front of it. Every run produced a written reasonableness review.

OneTwelve

PURPOSE-BUILT REASONABLENESS REVIEW

AI model	Claude Sonnet 5
Model runs	8 each recorded individually
AI work	250,569 44,291 read · 139,814 written
Elapsed time	5m 14s
Cost	$1.61 all 8 runs combined

Delivered: a reasoned verdict on all seven rejections, grounded in the cited references, nothing asserted that the record does not support — plus a formatted memo saved with the case.

Cowork

CLAUDE SONNET 5 — SAME MODEL · TWO RUNS

AI model	Claude Sonnet 5
Model runs	not itemized no per-run record
AI work	~13,000,000 13M read · 461 written
Elapsed time	13m 52s average of two runs
Cost	$5.30

Delivered: a written report file — and no verdict per rejection, no grounding back to the references, no cost record you could put in front of a client.

Cowork

CLAUDE OPUS 5 — THE LARGER MODEL

AI model	Claude Opus 5
Model runs	not itemized no per-run record
AI work	~5,400,000 5.4M read · 412 written
Elapsed time	9m 21s 8m 40s working
Cost	$5.28

Delivered: a written report file. The bigger, more expensive model read less and finished sooner — and still cost more than three times the OneTwelve review.

02 Before either clock starts

Every figure on this page assumes the case file is already sitting in a folder. Getting it there is work, and it is work only one of these two does for you.

OneTwelve

FROM THE DOCUMENT YOU ALREADY HAVE OPEN

1. You open the response you are drafting, **in Word**.
2. OneTwelve reads the **application serial number off the document itself**.
3. It pulls the office action, the specification, the drawings and all eight cited references **straight from the Patent Office record**.
4. The review starts. You have not left Word.

11 documents — gathered for you, in seconds

Cowork

NOTHING UNTIL YOU FETCH IT

1. Look the application up in Patent Center.
2. Download the office action.
3. Download the specification and the drawings.
4. Find and download **each of the eight cited references**, one at a time.
5. Put them in a folder, named so an assistant can tell them apart.
6. Point the assistant at the folder. *Now* it can begin.

11 documents — every one of them yours to find

WHICH MEANS THE MEASURED GAP IS THE FLATTERING ONE

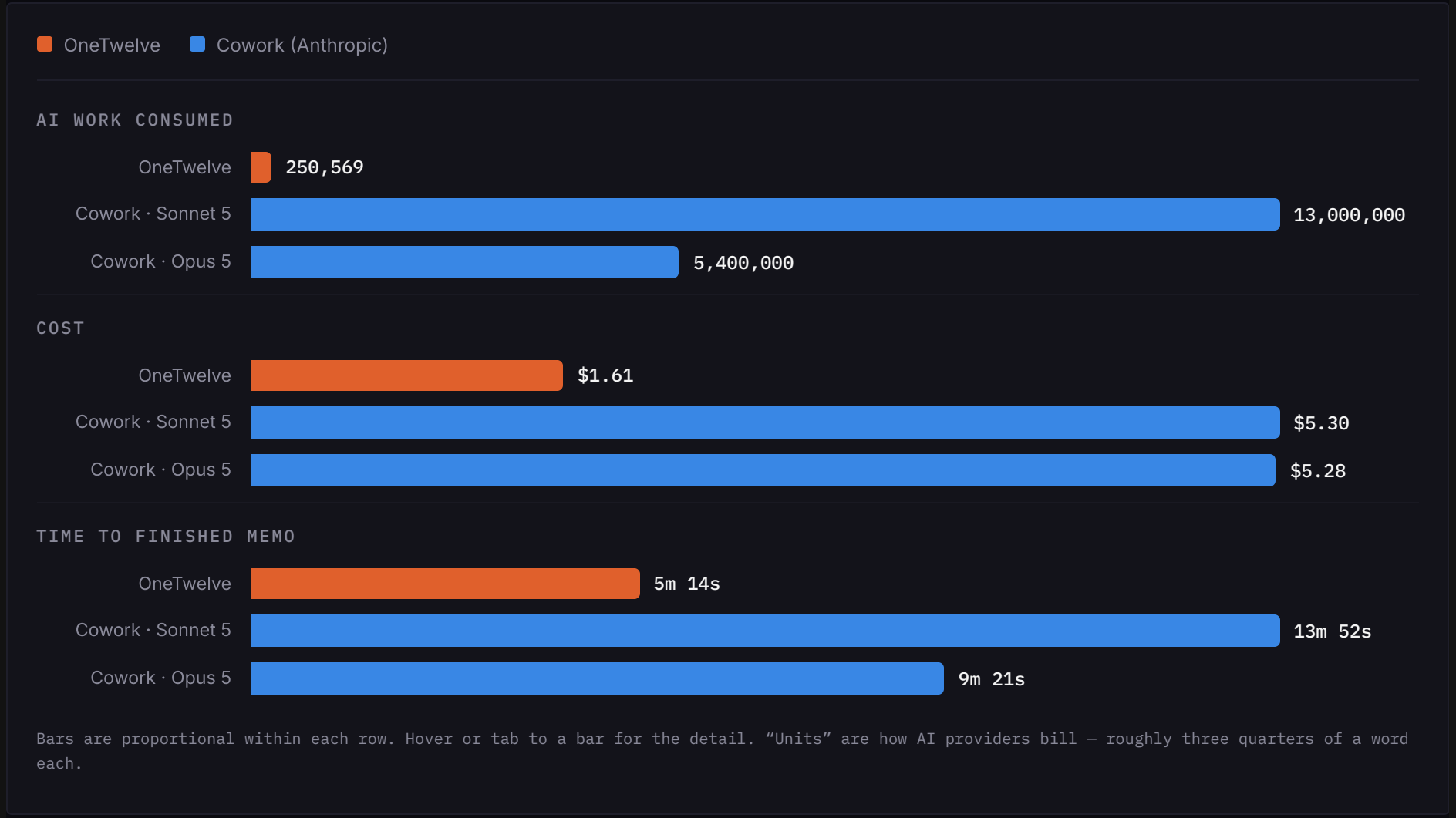
To run the comparison at all, we had to do Cowork's gathering by hand first — and then hand OneTwelve the same folder, so both sides started from the same place. **None of that manual retrieval is counted in any number on this page.** In an ordinary week it lands on whoever is drafting: a docket lookup, eleven downloads, and a naming convention to keep straight, before a single word of analysis gets written.

Put whatever that is worth in your practice against it — a quarter of an hour an office action is a conservative guess, and it is a quarter of an hour that gets **written off, or billed to a client for moving PDFs around**. Neither is what anyone is paying a patent attorney for. Retrieval that happens on its own does not just remove a hassle; it **hands the time back to work worth billing**.

03

The measurements

Same review, same evidence, same day. Longer bars are worse in every row.



Bars are proportional within each row. Hover or tab to a bar for the detail. “Units” are how AI providers bill — roughly three quarters of a word each.

04

Where the money actually goes

The cheapest run and the most expensive run **used the same AI model**, so raw model intelligence is not what separates them. What separates them is how much of the case file each one had to plough through before it could say anything useful.

OneTwelve

56% of the work is actual analysis

Reading	44,291
Analysis written	139,814
Reused between steps	66,464

Cowork · Sonnet 5

0.004% of the work is actual analysis

Reading	13,000,000
Analysis written	461

Cowork · Opus 5

0.008% of the work is actual analysis

Reading	5,400,000
Analysis written	412

A CHEAPER MODEL DID NOT MAKE IT CHEAPER

2.4× more of the case file read by the smaller model than the larger one — for the same job, in the same assistant.

Run inside Cowork, Claude Sonnet 5 read **13 million units**. Claude Opus 5 read **5.4 million**. The per-unit price of the smaller model is a fraction of the larger one, and it still finished up costing the same — $5.30 against $5.28.

The likeliest explanation is discrimination. Deciding which four of forty documents bear on a particular rejection is itself a hard judgement, and the larger model is simply better at it: it makes fewer, better-aimed passes. The smaller model compensates by reading more of everything. **OneTwelve never asks the model to make that call.** The pipeline already knows which documents a rejection turns on — which is exactly why a mid-sized model is enough, and why the same mid-sized model costs a third as much here as it does on its own.

BEFORE THE MODEL EVER SEES IT

The documents are already read — properly

Office actions, prior-art patents and file-wrapper documents arrive as scans — pictures of paper. OneTwelve reads them with its own recognition models, trained on **millions of patent documents, prior-art references and prosecution filings**. The text comes out clean and labelled: this is claim 1, this is the rejection, this is the passage the examiner leaned on. A general assistant gets none of that. It has to puzzle out the raw scans itself, every time it opens them — and you pay for every attempt.

THEN, AT THE MODEL

Only what the question actually needs

OneTwelve knows which four documents a particular rejection turns on, and hands the model exactly those, once. A general assistant has no way to know, so it keeps the whole file in view and re-reads it turn after turn to stay oriented. That re-reading is the bill.

WHAT A FRONTIER MODEL SAYS ABOUT RAW SCANNED PRIOR ART

"The reference copies in the file are degraded … I can verify the *substance* of the passages the Examiner relies on … but I cannot verify most of the pin cites … I am relying on recoverable text, which I paraphrase or reconstruct rather than pretend to quote verbatim."

We handed the same scanned references straight to a top-tier model as a separate check. It opened its answer by warning that it could not check the examiner's citations against the art it had been given. **That is the failure OneTwelve's recognition models exist to prevent** — the passages arrive clean, with the column and line numbers the examiner cited still attached, so a citation can actually be verified.

05 What the price tag does not show

Cost is the easy half of the comparison. The half that matters when a client asks is what actually arrives at the end — and that is where the gap widens rather than closes. **Fewer tokens bought the better answer**, not a cheaper one: what came back is shaped like prosecution work, because the system was built for prosecution rather than pointed at it.

01 **A verdict per rejection, not an essay.** All seven rejections came back with a position — sustainable or not, and why — in a form your review and docketing process can act on.

02 **Grounded in the actual references.** Every assertion traces to text in the cited art. On this run, nothing unsupported made it into the memo and nothing had to be handed back for a person to resolve.

03 **A memo, filed with the case.** The output lands as a formatted reasonableness memo stored with the application — not a chat transcript somebody has to rewrite.

04 **Measured, not projected.** Each of the eight model runs is recorded on its own — what it read, what it wrote, what it cost — and then totalled: $1.61 for this review. Every figure on this page was read straight out of that record. Nothing here is modelled, sampled, or extrapolated from a demo.

06 What it means at practice volume

The same difference, multiplied by a docket. Compared against the Cowork run on the same AI model OneTwelve used. These are AI processing costs for reasonableness reviews only, not a OneTwelve price list.

REVIEWS PER MONTH	ONETWELVE	COWORK, SAME MODEL	YOU KEEP
25	$40	$133	$93
100	$161	$530	$369
250	$404	$1,325	$921
500	$807	$2,650	$1,843

Extrapolated from measured runs at $1.61 and $5.30 per review. Actual spend moves with the length of the office action and the number of cited references.

07 How this was measured

A case study is only worth the provenance behind it — including what it does not prove.

THE ONETWELVE RUN

A live production review of application 15/352,952 on 21 August 2026, running on **Claude Sonnet 5**. The figures are OneTwelve's own usage records for that review — eight model runs, each recorded individually and totalled: **44,291 units read, 139,814 written, 66,464 reused between steps, $1.61 all in, 5m 14s**. Measured, not estimated.

THE COWORK RUNS	Cowork, Anthropic's agentic AI assistant, given the same case folder and asked for the same review. On **Claude Sonnet 5** — the identical model OneTwelve used: **13,000,000 units read, 461 written, $5.30, 13m 52s**. On **Claude Opus 5**, the larger model: **5,400,000 read, 412 written, $5.28, 9m 21s**. Figures as reported by Cowork's own usage meter.
RUN TWICE	The Sonnet 5 run above is **the average of two runs**, performed to confirm the cost rather than rely on a single session. They came out close: **$5.12** (14,000,000 read, 481 written, 12m 54s) and **$5.47** (12,000,000 read, 441 written, 14m 50s). The pattern is not a one-off.
WHAT EACH PRODUCED	OneTwelve returned verdicts on all seven rejections plus a memo filed with the case. Each Cowork run wrote a report file — 316, 401 and 541 lines across the three sessions — with no per-rejection verdict, no grounding back to the references and no cost record.
ONE JUDGEMENT CALL	That the smaller model reads more because it is **less able to tell which documents matter** is our reading of the evidence, not something the usage meter reports directly. What is measured is the reading itself: 13 million units against 5.4 million, for the same job in the same assistant.
SCOPE	One application, one office action; one OneTwelve run and three Cowork sessions. The costs here are AI processing only — no seats, storage or support. Your numbers will move with the size of the office action and the number of references, on every side.