

The Leaderboard, Checked Twice

Lesson 21 · Part 6 · first we asked another AI for rival commanders, then a researcher checked its figures against the actual sources · **this version replaces the earlier one, which was mostly wrong**

Start here · the words you need

Exchange ratio (the "K/D") — how many of the enemy you killed for each one of your own you lost. 10× means ten of theirs for one of yours.

Pitched battle — a stand-up fight both sides turned up for on purpose, as opposed to a raid, a siege or a skirmish. **Nobody has ever agreed on where the line is**, which is why "65 battles" is a softer number than it looks.

Primary source — someone writing at or near the time. **Secondary source** — a later historian describing it.

Folklore — a figure everybody repeats that turns out to trace back to nothing, or to one unreliable text.

What happened, in order

Step one. We asked Gemini 2.5 Flash for twelve commanders who might beat Subotai's record, with casualty figures. It gave us a table and **no sources at all**. We published it and said so.

Step two. A researcher checked each figure against the actual ancient and modern literature.

Only three of the eleven pairs survived. Four turned out to be folklore, two are disputed between ancient authors, and two appear in no source anybody can find.

Subotai's own entry — the whole reason we built the table — is one of the folklore ones.

The checked table

Commander	Battle	What we published (own / enemy)	Verdict
Miltiades	Marathon, 490 BCE	192 / 6,400	Confirmed , with a caveat
Napoleon	Austerlitz, 1805	9,000 / 27,000	Confirmed
Frederick the Great	Leuthen, 1757	6,500 / 22,000	Confirmed
Hannibal	Cannae, 216 BCE	8,000 / 60,000	Disputed
Scipio Africanus	Zama, 202 BCE	3,750 / 25,000	Disputed
Julius Caesar	Alesia, 52 BCE	3,000 / 70,000	Folklore
Alexander	Gaugamela, 331 BCE	3,000 / 65,000	Folklore
Khalid ibn al-Walid	Yarmouk, 636	3,500 / 60,000	Folklore
Subotai	Kalka River, 1223	3,000 / 60,000	Folklore
Tamerlane	Ankara, 1402	20,000 / 80,000	Not found
Gustavus Adolphus	Breitenfeld, 1631	4,000 / 9,500	Wrong battle — those are 1642

The four that matter, explained

Marathon — real numbers, doubtful enemy count. Both figures come straight out of **Herodotus**, the Greek historian writing about fifty years later: "about six thousand four hundred men of the foreigners were killed, and one hundred and ninety-two Athenians." So the numbers are genuinely ancient. **But modern historians don't believe the 6,400.** Losing 30% of your army while the other side loses 2% is not how battles go. Some put the Persian dead at 1,000–2,000 instead. Also, the 192 counts **Athenians only** — Herodotus gives no figure for their Plataean allies.

Cannae — two incompatible authors, quietly stitched together. This is the instructive one. **Livy** says Hannibal lost about 8,000. **Quintilian** is the only source for 60,000 Romans dead. **Polybius** — generally the more careful ancient historian — says 70,000 Romans but only **5,700** Carthaginians. Our published pair took the 8,000 from one author and the 60,000 from another and ignored the third.

Why that matters more than the error itself. Nobody lied. Each number has a real source. The ratio was manufactured purely by **choosing which author to believe for each half** — and the two halves were never meant to sit beside each other.

Austerlitz — solid, and the best-documented thing on the page. French losses of 8,569–9,288 come from official returns, broken down as 1,305 killed, 6,940 wounded, 573 captured, and confirmed by four separate modern historians. Napoleon's own propaganda bulletin *did* exaggerate — it claimed 20,000 Allied troops drowned in the Satschan ponds, which did not happen — but the scholarly figures hold.

Yarmouk — the 60,000 is in no source whatsoever. The medieval chroniclers give 120,000 (al-Tabari), 70,000 (Ibn Ishaq), 70,000 (al-Baladhuri). Modern historians put the **entire** Byzantine army at 15,000–50,000, which makes 60,000 dead arithmetically impossible. The "60,000" is popular-history rounding with no ancestor.

Two specific claims we made, both corrected

"Khalid ibn al-Walid won about 100 engagements and was never beaten" — folklore. The ~100 is everywhere in popular writing and **absent from scholarship**. No chronicler keeps a tally. A careful enumeration of his career lists roughly **31 named engagements**. What is defensible: no clear battlefield defeat under his overall command survives in the tradition. The researcher's line on it is the right one — turning that into "100 and 0" *"makes seventh-century history look suspiciously like a boxing record."*

"Subotai conquered 32 nations and won 65 pitched battles, per Muslim chroniclers via Liddell Hart" — the trail stops at Gabriel. That exact wording is the **publisher's description** of Richard Gabriel's 2006 book. The tally could not be found in Liddell Hart's own work, and **no primary chronicler counts Subotai's battles at all** — Juvayni, Rashid al-Din and Ibn al-Athir narrate campaigns without keeping score.

So the number that started this whole investigation has exactly one source, and it is modern.

That does not make it false. It makes it a claim rather than a record — and it is the only candidate figure on the board that beats everyone else, which is precisely the position where you want more than one source.

Our hypothesis survived, and it has a proper name

We guessed that **reported ratios track the quality of the record-keeping, not the skill of the general** — that the spectacular ancient numbers are bookkeeping artefacts.

That turns out to be an established argument, and we were about to over-credit its author.

The name attached to it is **Hans Delbrück**, who built the method for testing reported army sizes against physical constraints — road space, marching order, supply. His demolition of Herodotus's claim that Xerxes invaded with five million men is permanent: he calculated that such a column would stretch roughly 1,960 miles, so "the last men could only have left Susa, beyond the Tigris, when the first arrived before Thermopylae."

Three corrections to how we first described him, all from the same check.

He did not originate number-scepticism — he credits the earlier historian **Niebuhr**. What is his own is the *systematic method* (*Sachkritik*) and a *stronger* claim.

His actual thesis is subtler than ours. Not merely "enemy figures are inflated", but that **the victor usually outnumbered the vanquished** — the "few against innumerable masses" story being a demand of popular taste. That is a **narrative** mechanism, where ours was a **bureaucratic** one. They are compatible but not the same claim.

The field kept his method and rejected his minimalism. Modern estimates of Xerxes' force cluster at 100,000–200,000 — **five to ten times Delbrück's own ceiling of 20–25,000.**

So cite him for the method and the maxim, not for the thesis's origin — that belongs to the Niebuhr–Delbrück tradition. And his strong parity rule is the contested part, not the accepted core.

And the surviving data fits. The one confirmed ancient case is Marathon at 33× (single source, enemy figure now doubted). The two confirmed modern cases are Austerlitz at 3.0× and Leuthen at 3.4×, where own losses come from state paperwork.

Two honest caveats on that, though:

- **The modern "3×" includes prisoners.** Both Austerlitz and Leuthen counted roughly 12,000 captured enemy in that enemy total. Count only killed and wounded and the ratios drop to about **1.7–1.8×** — so the real gap between eras is *wider* than we said.

- **Part of our neat slope was invented.** Four of the eleven figures were folklore, so the tidy 33× → 2.4× gradient was partly an artefact of Gemini's guessing.

The important correction · we had the mechanism backwards

What we published: armies recorded the enemy's dead and not their own, so the missing half of any ratio is your *own* losses.

That is wrong for the modern era, and the right version is more interesting.

Early-modern and modern states recorded their own dead meticulously — because they had to pay for them. Wages, pensions, widows, replacements. France kept roughly three million individual personnel records; there is a published dictionary naming the French dead of Austerlitz *man by man*. Prussia's 6,344 at Leuthen comes from regimental returns.

It is the **enemy's** dead that were always the estimate, for the obvious reason: **a defeated army does not file paperwork for you.**

So why does our conclusion still hold? Because the two eras attest **opposite halves**, and never both at once.

Era	Who counted the enemy well?	Who counted their own well?
Mongols, 13th c.	Yes — ears, one per body, for the next campaign's planning	No — never recorded
Modern states	No — always estimated	Yes — to the individual name, for pay and pensions

The rule underneath both, which is the actual finding:

Each era documented carefully whichever side it owed something to. Modern states owe wages and pensions to their own soldiers, so their own dead are named. The Mongols owed nothing to anyone, but needed an accurate enemy count to decide what to attack next — so the enemy's dead were the measured side.

Nobody ever had a reason to measure both. That is why no commander in any century has a computable career exchange ratio — and the reason is incentive, not ignorance.

What this episode is actually worth

The headline is unchanged and now properly supported: there is no highest K/D in history, because the statistic has never been collectable.

But the better lesson is about the method. We published a table with an explicit warning that nothing in it was verified. The warning was correct and it was not sufficient — **four of eleven rows were fabrications, and a reader skimming the table would have carried them away regardless of the caveat.**

A caveat protects the author. It does not protect the reader. The only thing that protects the reader is checking, and checking took a second pass by someone with a browser.

Credits and limits

The verification was done independently by a browser-based research agent, working from the published table and no access to this site. Its report gave source URLs for every verdict, named the specific ancient authors behind each disputed figure, and flagged where our own reasoning was wrong rather than only where our numbers were.

A second pass resolved the open rows. Defensible replacements, with the ancient author named each time:

Battle	What we published	Defensible figures
Cannae	8,000 / 60,000	Punic 5,700–8,000 dead (Polybius–Livy); Roman 48,200–70,000 dead + ~10–19k captured
Zama	3,750 / 25,000	Roman 1,500–2,500 killed; Carthaginian 20,000–25,000 killed + 8,500–20,000 captured
Ankara	20,000 / 80,000	No defensible figure on either side. The pair appears nowhere
Breitenfeld 1631	4,000 / 9,500	Allied ~5,000–5,550; Imperial ~7,600 + ~6,000 captured. Our pair is the 1642 battle
Kalka River	3,000 / 60,000	Rus' losses catastrophic but unquantifiable; Mongol losses unknown

The sharpest line in the whole check: two of those five pairs contain numbers with **no source at all — not even a bad one.** Zama's "3,750", Ankara's entire pair and Kalka's "3,000" were generated, not misremembered.

Where to read Delbrück himself. *Geschichte der Kriegskunst im Rahmen der politischen Geschichte*, Erster Teil: *Das Altertum* (3rd ed. 1920), Book 1 ch. 1, "Heereszahlen. Vorbereitendes". In English: *History of the Art of War*, Vol. 1, *Warfare in Antiquity*, trans. Walter J. Renfro Jr. (University of Nebraska Press, 1975), **p. 33** for the maxim — "A movement that is made by an organization of 1,000 men

without complications becomes an accomplishment for 10,000 men, a work of art for 50,000, and an impossibility for 100,000." His own English summary, **Numbers in History** (1913), is public domain on archive.org and contains the Xerxes calculation.

One flag carried over honestly: that page reference was verified second-hand through Bucholz's footnotes, not against a physical copy. **Check p. 33 before quoting it anywhere that matters** — which is precisely the discipline this whole lesson exists to teach.