



What the Research Says About Debriefing

An honest account of the evidence behind debriefing a political simulation, including what the field asserts but has not shown.

For college and high school instructors.

The claim, and why it is worth stating plainly

David Crookall, who edited *Simulation and Gaming* for decades, argues that a simulation achieves its intended learning only when the debriefing is properly engaging, that most of the learning happens in the debrief rather than in the game, and that running a simulation without a full debriefing is unethical. Timothy Wedig, writing in *PS* for political scientists, calls the debrief “the key to the entire simulation process.”

Now the honest part

That claim is asserted rather than demonstrated. No study isolates the debrief’s contribution by comparing a simulation that was debriefed against the same simulation that was not. The large meta-analysis in this area, Cheng and colleagues in *Medical Education*, pooled 177 studies and 11,511 learners and found effects from 0.28 to 2.16, but it compared simulation plus debriefing against no intervention at all. That is a test of the package, not of the debrief.

Two further things a teacher deserves to be told.

- **Almost the entire empirical literature is healthcare simulation:** anesthesia, nursing, resuscitation, surgery, with professional trainees, short technical scenarios and skill-based outcomes. Political simulations run longer, aim at conceptual understanding rather than psychomotor skill, and require students to argue positions they may not hold. None of it transfers automatically.
- **Political science has essentially no empirical research on debriefing at all.** It appears as advice inside articles about running simulations. Shaw and Switky’s 2018 treatment in the *Journal of Political Science Education* cites exactly one source in its debriefing section. Brown and Ostojski proposed in 2025 to begin studying the question.

*Schwägle and colleagues searched *Simulation & Gaming* itself and found 631 articles that mention debriefing and eight that study it methodologically.*

Five things the evidence does support

1. Debriefing is associated with retention, and that is the strongest finding we have

Levin-Banchik taught four topics by four methods and tested students with an anonymous pop quiz three months later. Simulation plus in-class debriefing was most effective; simulation alone was close behind; learning without simulation trailed both. This is a political science study, in a political science classroom, with a delayed measure, which makes it the single most useful citation available.

2. Instructors talk too much, and this is documented rather than suspected

An analysis of 36 airline crew debriefs found instructors talking more than the crews and rarely allowing silence. Dieckmann and colleagues coded actual debriefings and found four interaction patterns, of which only one, which they call the **net**, involves participants referring to each other at all. The others run through the instructor like spokes. If you take one practical thing from the research, take this one.

3. Video adds nothing reliable

This is the clearest convergent null in the field. Cheng and colleagues found an effect size of 0.10, not significant. Levett-Jones and Lapkin found no statistically significant advantage. Their later review notes that video-enhanced debriefing, “often considered the gold standard,” showed no clear advantage. Video costs money and time and tends to eat the discussion. You can stop feeling guilty about not recording.

4. Nobody knows the right length

Only five of ten trials in one systematic review reported debrief length at all, and those ranged from two minutes to thirty. Short versus long was not significant and, if anything, favored short. **The common rule that you should debrief for at least as long as the exercise ran has no evidence behind it.**

5. No debriefing structure has been shown to beat another

Sawyer and colleagues reviewed every named method and concluded that no studies have objectively compared the conversational structures, that empirical support for any specific method is limited, and that the act of debriefing may matter more than which method you use. Treat the named structures as plausible, experienced designs rather than as competitors with a winner.

One counterweight nobody in this field mentions

A Cochrane review of single-session psychological debriefing after real trauma found it did not prevent post-traumatic stress disorder, and in one study at thirteen months a diagnosis was significantly **more** likely in the debriefed group. That is a different intervention, after real trauma rather than a classroom exercise, and the comparison should not be overdrawn. But the lesson generalizes: a structured conversation that makes people revisit an intense experience is an intervention with a direction, not an automatic good. Do it well, and do not assume that doing it at all is protective.

Claims to avoid repeating

- “The debrief is where the learning happens,” unqualified. Nobody has tested the counterfactual.
- That any particular structure beats another.
- Any specific right length.
- That peer-led debriefing works as well. What exists is **self**-debriefing performing comparably in a few healthcare trials; peer-facilitated debriefing is effectively unstudied.
- That immediate beats delayed. Unsettled, and within-event was inferior in one surgical study.
- Kolb’s cycle as evidence. It is a rationale, and the model is contested in educational psychology.
- Any de-roling protocol. There is no accepted model and no evidence base. A 2024 conference presentation proposing one says so in its own abstract.

Provenance corrections worth knowing

- “What? So what? Now what?” is Driscoll 1994, developed from Borton 1970. Attributing the model to Borton alone is wrong.
 - **Thiagi’s six-phase model** comes from a 2004 practitioner newsletter, not a refereed journal, and has never been tested.
 - **Plus and delta** has no founding paper. Its best provenance is aviation practice.
-

The load-bearing sources

- Wedig, Timothy. 2010. “Getting the Most from Classroom Simulations.” **PS: Political Science and Politics** 43(3): 547–555. The anchor, and the only structure built for a political science classroom.
- Levin-Banchik, Luba. 2018. **Journal of Political Science Education** 14(3): 341–359. The best evidence here: a delayed measure in a political science classroom.
- Cheng, Adam, et al. 2014. “Debriefing for Technology-Enhanced Simulation.” **Medical Education** 48(7): 657–666. Healthcare. Tests the package, not the debrief.
- Sawyer, Taylor, et al. 2016. “More Than One Way to Debrief.” **Simulation in Healthcare** 11(3): 209–217. The cross-walk of every named method.
- Fanning, Ruth M., and David M. Gaba. 2007. **Simulation in Healthcare** 2(2): 115–125. The 36 airline debriefs.
- Dieckmann, Peter, et al. 2009. **Medical Teacher** 31(7): e287–e294. The four interaction patterns.
- Schwägele, Sebastian, et al. 2021. **Simulation & Gaming** 52(3): 364–385.

- Crookall, David. 2014. **Simulation & Gaming** 45(4–5): 416–427. An argument in an editorial, not a finding.
- Shaw, Carolyn M., and Bob Switky. 2018. **Journal of Political Science Education** 14(4): 523–534.
- Rose, Suzanna, Jonathan Bisson, Rachel Churchill, and Simon Wessely. 2002. Cochrane Database of Systematic Reviews CD000560.
- Kartal, Filiz. 2021. **Journal of Political Science Education** 17(sup1): 148–168. On why self-assessment inflates apparent gains.
- Clark, Nicholas, and John A. Scherpereel. 2024. **Journal of Political Science Education** 20(1): 133–152. The KESE framework; skills gains not supported.

Every citation above was checked against publisher records on October 17, 2026. A small number of secondary citations mentioned in the companion handouts are still being confirmed and are not listed here.