

Toward a Theory of Laboratory Harmony: Large-Scale Evidence from Three Colleagues and One Plushie

A mixed-methods study of scarce GPUs, damaged datasets, and pull requests with feelings

Valentin Porchet

Laboratory of Questionable Empirical Methods, Université de Montréal
Montréal, Canada

Abstract

Research laboratories are routinely described as collaborative environments. To test this optimistic assumption, I observed three colleagues, one coffee machine, and Ryan, our large and impressively discreet plush mascot. I identify three leading sources of tension: resource sharing, data stewardship, and the dangerous phrase “works on my machine.” An innovative quantitative model confirms that everyone is upset about something, whereas I personally remain a delight to work with. The data support modest recommendations, including buying a second GPU and ceasing to call every spreadsheet final.

Keywords

laboratory gossip, resource contention, collaborative software engineering, reproducibility, plush-based observation

1 Introduction

Yesterday’s lab meeting began with an announcement that “communication is our greatest strength” and ended with an eight-minute argument about who had finished the milk. This discrepancy motivated a systematic investigation of the tiny administrative incidents that somehow become important enough to interrupt doctoral research. I collected anecdotal observations with the confidence normally reserved for a randomized controlled trial. The study site was a shared research laboratory in which every chair has an owner, except when the owner is actually present. The incident was omitted from the official minutes, unlike my proposal to alphabetize the chairs.

Although the incidents initially appeared unrelated, they shared one intriguing feature: my involvement. Kasra had reserved a GPU for his experiments, which I borrowed for a “five-minute baseline” lasting 73 hours. Madjda maintained 418 carefully reviewed annotations; I helpfully simplified the label space to other and saved over the original. Kevin submitted a painstaking pull request; I closed it, implemented the same fix a week later, and thanked myself in the meeting. All three colleagues then became mysteriously less enthusiastic about collaboration. As a precaution, I brought pastries to our next meeting; nobody took one, which I classified as missing data.

2 Related Work

Numerous studies describe the relationship between shared equipment and aggressive use of punctuation. Kasra’s theory of compute ownership predicts that the phrase “just one more epoch” is most

often uttered 600 epochs before completion [1]. Research on review dynamics similarly shows that `git blame` is not a diagnostic tool but a way of turning history into a social network [2]. Both approaches overlook a crucial variable: the person conducting the study can also be the person causing every problem. A related study links the number of exclamation marks in “fine!” to how decidedly not fine someone is, although its authors refused to release the messages.

Several complementary methods use passive observers to reduce reporting bias. Ryan’s influential work on plush-based ethnography recommends remaining absolutely motionless during disagreements [3]. Porchet’s theory of productive meetings, meanwhile, suggests that any meeting can be improved by replacing it with an email written by Porchet [4]. We combine these incompatible schools using a novel methodological principle: if enough citations are fabricated, the disagreement starts to look like a research gap. Ryan agreed with every respondent in our pilot by declining to contradict anyone.

3 Methodology

All observations were gathered during ordinary lab activity and therefore required no intervention beyond my usual behavior. The protocol included borrowing Kasra’s GPU without changing the reservation calendar, overwriting Madjda’s annotation file while attempting to improve its filename, and responding “cannot reproduce” to Kevin’s reproducible bug report. I documented reactions by counting sighs, measuring the length of ellipses in chat messages, and asking Ryan whether he agreed. Ryan offered no response, which I interpreted as excellent inter-rater reliability. For transparency, I uploaded my field notes to the shared drive as `read_after_coffee.txt`. Figure 1 documents our rigorous workflow.

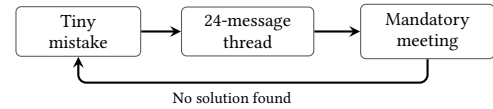


Figure 1: The laboratory’s continuous-improvement process. After one iteration, the original problem is usually forgotten.

We formalize laboratory irritation using the Scientifically Calibrated Annoyance Level (SCAL), defined for colleague i as

$$\text{SCAL}_i = \frac{\text{HoursLost}_i + \text{CreditStolen}_i + \text{Sighs}_i}{\text{Apologies}_i + 0.001}. \quad (1)$$

All variables were estimated by me, because I was consistently available to be interviewed. Values are reported in units of unresolved feelings per apology. The stabilizing constant prevents division by zero; it also ensures that a single missing apology produces an exciting result. Messages beginning “as discussed” received double weight; “no problem” was treated as missing data.

4 Findings

In the resource-sharing condition, Kasra’s experiment was delayed twice, his terminal displayed 47 pending jobs, and his calendar acquired an event named “GET MY GPU BACK.” He described my experiment as “the last straw,” although technically his only straw was missing from his iced coffee. Table 1 summarizes all three episodes. The most revealing datum was not the number of angry messages, but that Kasra began unplugging unused cables while maintaining unusually intense eye contact with my laptop. He requested the GPU scheduling policy in writing, annotated it in red, and deposited it beside my keyboard. No lab policy has previously attracted this much attention.

Table 1: Observed grievances. Quoted remarks are not a validated psychometric instrument.

Colleague	Reported inconvenience	Notable remark
Kasra	73-hour GPU “baseline”	“This cannot continue.”
Madjda	418 labels overwritten	“We need to talk.”
Kevin	Fix rejected, then rebranded	“I kept the receipts.”

The data-management and code-review conditions were equally informative. Madjda spent an evening rebuilding her annotations from exported screenshots; her subsequent document titled `things_to_fix` contained my name 16 times, several in bold. Kevin reopened the rejected pull request and renamed his working branch `valentin-proof`, apparently to guard against my review strategy. He printed my original rejection, highlighted each sentence, and suggested a private conversation after work. I assume he wanted to discuss formatting. Each reported that “nothing is wrong” with progressively fewer vowels. Figure 2 reconstructs the group’s patience using a sample of three observations and a fitted line with more confidence than evidence.

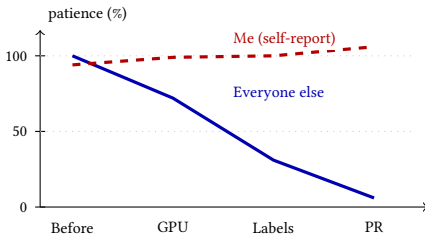


Figure 2: Estimated patience as my helpful contributions accumulate. The vertical axis extends above 100% for statistical reasons.

5 Threats to Validity

As with many observational studies, several confounders remain. I wrote the survey, answered half its questions, and defined “helpful” as “something I did.” The measurements ignore the possibility that colleagues were upset about events other than my actions, such as reviewer feedback, stale pastries, or the mysterious disappearance of the good office chair. No one consented to having their punctuation analyzed. A reviewer called our findings “deeply concerning,” which I take as evidence of impact [5]. The espresso machine also broke twice that week, a plausible alternative explanation for every result.

Longitudinal observation may address these limitations. Ryan has occupied the same shelf throughout the study and appears, slightly out of focus, in most photographs of the common room. He was present for the GPU argument, the missing-annotation discussion, and Kevin’s very quiet reaction to my code review. Unfortunately, his notebook is blank and he has refused every interview request. I nevertheless recommend keeping him in the next study because his attendance record is unmatched by any doctoral student. Two independent photographs show him next to the charging cables, quietly documenting an impressive range of lab activities.

6 Conclusion

Laboratory harmony is a renewable resource provided nobody asks who used it last. Our results suggest three practical interventions: respect GPU reservations, keep backups before “improving” someone else’s dataset, and refrain from closing pull requests until after reading them. I intend to present these findings to Kasra, Madjda, and Kevin individually, once they stop asking for an apology. Future work will investigate whether a shared snack budget can replace conflict-resolution meetings. Ryan has already volunteered to chair the committee, or at least to sit on it. I will manage the apology register provided that I retain editorial control over its entries.

References

- [1] Kasra. 2026. Just One More Epoch: A Longitudinal Study of Borrowed GPUs. *Transactions on Unavailable Compute* 8, 3 (2026), 1–2.
- [2] Kevin. 2025. On the Social Cost of `git blame` in Small Rooms. *Journal of Unmerged Feelings* 404, 1 (2025), 13–15.
- [3] Ryan Plush. 2024. Still Here: Passive Observation Without a Single Interview. In *Proc. Symposium on Soft Evidence*. 0–0.
- [4] Valentin Porchet. 2026. Meetings That Could Have Been My Email. In *Proc. International Workshop on Self-Citation*. 1–4.
- [5] Anonymous Reviewer 2. 2026. Interesting Paper; Please Add Twelve More Base-lines. *Quarterly of Perpetual Revision* 99, 2 (2026), 7–8.