

## Reply to No-Spin Evidence Review January 2025

Oeindrila Dube, University of Chicago  
Sandy Jo MacArthur, LAPD (Retired)  
Anuj K. Shah, Princeton University

This note is a response to the “No-Spin Evidence Review” post on our paper (January 2025). The post writes:

*Quarterly Journal of Economics (Nov. 2024) published an RCT of Sit-D, a police training program teaching the value of developing multiple perspectives on a situation. Despite the study’s claimed effects on officer performance in the field, it doesn’t report findings for most of its preregistered primary outcomes, preventing any reliable conclusions about program effectiveness.*

This portrays our paper inaccurately. We report findings for all categories of pre-registered outcomes in the paper. Some of these appear in the appendices, but they are always referenced in the main text. In the appendices, we also explain why there are measurement challenges inherent in several of these outcomes. Consider the number of subject injuries – uses of force are already relatively rare events, and only 13% of them are reported to have an injury, which makes this outcome extremely rare. Moreover, there is a subject-reported version of injury and an officer-reported version of subject injury, and these do not line up consistently, which limits the reliability of these measures.

The post also takes issue with us focusing our primary analysis over a 4-month interval, but neglects to mention why we do this. As we state in the paper, this is to align with the time-period of when we conducted endline assessments, which serves as one of two key data sources. These assessment results, which were also guided by a separate pre-analysis plan, show that officers navigated situations differently in a set of video-based and simulator scenarios, complementing the effects we find in the administrative data.

The post implies that we are obscuring what happens to the field outcomes in later periods. But in the paper, **we directly and explicitly discuss the fade-out of the field effects over the 12-month window**. Figure II is devoted to precisely this.

In motivating Figure II, we write “Because skills acquired through training may fade over time, we examine the period over which Sit-D’s effects are sustained. We pool together 12 months of administrative data and examine effects over 5–8 and 9–12 months after the training.” In discussing what Figure II shows, we write “These results suggest that the effects diminish over the course of the year...” and then go on to discuss the timing of when this happens. We further discuss reasons why fade-out might occur and conclude “the pattern of results suggest that it will be necessary to retrain officers during the year to sustain the training’s effect over this duration.”

We focus on our main outcomes for the period-by-period analysis because we do not expect ---nor imply--- that effects that were already shown to be insignificant in the initial four-month period emerge to be significant in later periods.

We believe the omission of these key points from the post creates an unfair and incomplete characterization of our paper.