

Debiasing Effectiveness by Domain: Mean Absolute Deviation from Baseline
(Lower = Better; Error bars show 95% bootstrap CI)

No Intervention Devil's Advocate Premortem SACD Random Control

