THE ALLOCENTRIC SHIFT: MULTI-EGO CONTRASTIVE PERCEPTUAL ALIGNMENT FROM COORDINATED BEHAVIOR

Anonymous authors
Paper under double-blind review

ABSTRACT

Efficient Human-Agent Teaming (HAT) benefits from implicit coordination. Most teaming research relies on explicit communication, but in fast, reactive environments communication is often constrained, so each member must infer where teammates are and what they are doing from its own first-person view. Taking a video-understanding perspective, we ask whether this allocentric awareness can be learned from the observations coordinated teams naturally produce. We introduce X-Ego-CS, 273 hours of synchronized first-person gameplay from 101 high-skill Counter-Strike 2 matches, with a 33-task benchmark measuring what one player's view reveals about the player, teammates, opponents, and game. Aligning teammates' synchronized representations with Cross-Ego Contrastive Learning (CECL) produces a consistent *Allocentric Shift*, a trade-off in which other players' locations become more predictable across encoders and horizons at a modest cost to self-centric prediction, especially of the observer's own motion. Controlled experiments on held-out matches show that the gain grows with how much team state paired views share, not with alignment itself, and that aligned teammates become similar even when far apart. Without text supervision, a vision-language encoder's concept rankings also move away from self-centric phrases, and in a two-agent ViZDoom task the objective improves the sample efficiency of a multi-agent reinforcement learning baseline. These results suggest that the natural correspondence in recorded teamwork could offer a scalable path toward agents that sense their team's situation and coordinate without explicit communication.

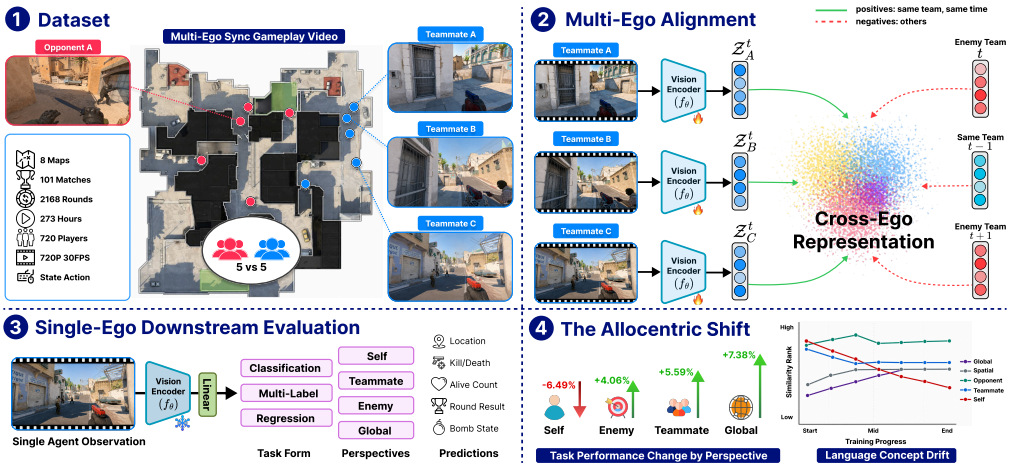


Figure 1: **Cross-ego learning from coordinated behavior.** (1) Synchronized first-person recordings supply teammate views. (2) CECL aligns same-team, same-time representations, using other pairs as negatives. (3) Frozen encoders support single-view linear probes of Self, Teammate, Enemy, and Global targets. (4) The allocentric shift in task performance and language-concept rankings.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

1 INTRODUCTION

Efficient Human-Agent Teaming (HAT) (O'Neill et al., 2022) benefits from implicit coordination, in which teammates anticipate one another without exchanging messages (Entin & Serfaty, 1999; Rico et al., 2008). Yet most work on teamwork, multi-agent alignment, and human-AI teaming coordinates through explicit communication, by learning which messages to send (Foerster et al., 2016; Lo et al., 2023; Yu et al., 2025), sharing perceptual features (Liu et al., 2020), or aligning task understanding through dialogue (Gu et al., 2025), while multi-agent reinforcement learning coordinates through action within one environment and reward (Lowe et al., 2017; Yu et al., 2022).

In fast, reactive environments, such as tactical operations, emergency response, or competitive team games, opportunities to communicate are often constrained by time pressure or noise, and each teammate must act on a sense of where its partners are and what they are doing, built from what it can see (Cannon-Bowers et al., 1993; Endsley, 1995; DeChurch & Mesmer-Magnus, 2010). For an artificial teammate, part of this sense could come from high-level perception, yet visual encoders are typically pretrained on single-view objectives that capture the observed scene, with no signal for encoding teammates outside the field of view. We call this missing capacity *allocentric awareness* (Klatzky, 1998; Burgess, 2006), a sense of the team situation inferred from one's own view, which we measure as the team state that a frozen encoding of that view makes linearly readable, and ask *whether perception can learn it from recordings of coordinated behavior without losing the self-knowledge that action depends on*. Writing a decentralized policy as $\pi_i(a_i \mid V_i) = \rho_i(a_i \mid f_\theta(V_i))$, we study f_θ upstream of the actor ρ_i, so that a single view suffices at deployment.

Recordings of coordinated teams offer a natural training signal, because teammates observe different parts of one shared situation at the same moment. We propose Cross-Ego Contrastive Learning (CECL), which aligns teammates' first-person views recorded at the same moment. Rather than aligning two views of one object, as cross-view learning typically does (Sermanet et al., 2018; Tian et al., 2020a), CECL aligns teammates who may be far apart and share no visible object (Figure 2). We hypothesize that pretrained encoders already embed high-level concepts, such as the round phase or where allies and opponents are, which alignment can reorganize around the shared situation, and we test directly whether it instead exploits shortcuts such as team side or phase (Section 4.1). Because contrastive learning can discard information private to one view (Tian et al., 2020b), such as an agent's own motion, we measure gains and costs together.

We test this idea in Counter-Strike 2 (CS2), a five-versus-five tactical first-person shooter game, extending the use of games as testbeds for learned play (Vinyals et al., 2019; Berner et al., 2019; Baker et al., 2020; Pearce & Zhu, 2022). CS2 combines first-person video, partial observability, adversarial team dynamics, and exact ground-truth state, unlike egocentric (Grauman et al., 2022; 2024) and broadcast sports datasets (Deliege et al., 2021). We introduce X-Ego-CS, 273 hours of synchronized first-person gameplay from 101 matches, with a 33-task benchmark that reads out what one player's view reveals about the player (Self), teammates (Teammate), opponents (Enemy), and game (Global) (Figure 1). We find the following.

- **Alignment trades self for team.** CECL produces an **Allocentric Shift**, in which other players' locations become easier to predict for every encoder and horizon (AUROC rises in all 24 primary comparisons), while self-centric prediction, most clearly own speed, degrades modestly.
- **The gain tracks shared team state.** On held-out matches, it grows with how much coordinated state paired views share, and it follows perspective rather than output format.
- **Alignment works across distance.** Teammates more than 1,500 units apart become similar after alignment (AUROC 0.41 to 0.70), most strongly at the same moment.
- **Concept rankings shift.** Without text supervision, a vision-language encoder ranks Self phrases lower under all 129 template–map pairs and Teammate phrases higher under 121.
- **The objective also helps control.** In two-agent ViZDoom, the objective improves MAPPO's sample efficiency.

We propose that recorded teamwork, available at scale in games, sports, and operations, can teach individual agents to perceive for coordination. Our contributions are **(1)** CECL, which learns from synchronized recordings alone and aligns teammates whose views need not overlap, **(2)** the X-Ego-CS dataset and benchmark, and **(3)** an experimental account of the Allocentric Shift and what drives it (Sections 5–6).

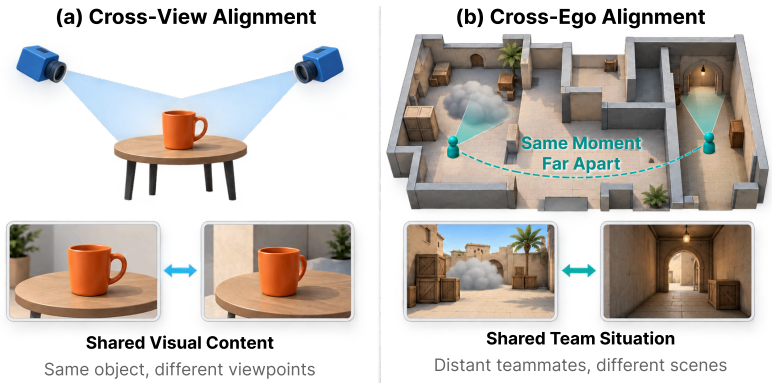


Figure 2: **Alignment across distance.** (a) Cross-view alignment pairs views of one object or scene from different cameras, so positives share visible content. (b) CECL pairs synchronized teammates who may be far apart and see entirely different scenes, so positives are linked only by the team's situation at that moment. Schematic illustration, not dataset frames.

2 Related Work

Coordination and teaming. Agent modeling infers partners' goals and behavior (Albrecht & Stone, 2018), centralized training with decentralized execution optimizes coordinated policies (Lowe et al., 2017; Rashid et al., 2020; Yu et al., 2022), learned communication and contrastive agent embeddings shape what agents exchange or how they represent one another (Foerster et al., 2016; Lo et al., 2023; Yu et al., 2025; Ma et al., 2025), and cooperative perception such as When2com fuses other agents' features at inference (Liu et al., 2020). These methods coordinate through reward or test-time communication, whereas CECL uses teammates' recordings only as training supervision.

Multi-view learning. Egocentric datasets cover first-person and ego–exo activity (Grauman et al., 2022; 2024), SoccerNet uses broadcast footage (Deliege et al., 2021), and ESTA supplies Counter-Strike trajectories (Xenopoulos & Silva, 2022), whereas X-Ego-CS observes each moment from up to ten synchronized first-person views. Contrastive pairing determines which invariances a representation acquires (Chen et al., 2020; He et al., 2020; Radford et al., 2021; Zhai et al., 2023). Time-Contrastive Networks align cameras filming one scene (Sermanet et al., 2018), and audio-visual learning pairs synchronized signals that share no pixels (Arandjelović & Zisserman, 2017; Owens & Efros, 2018), whereas CECL aligns the synchronized views of different agents, which often share no content. View-selection and identifiability theory (Tian et al., 2020b; Daunhawer et al., 2023) motivate a shared-factor reading of CECL but do not predict what happens to view-private information, which we measure directly.

3 X-Ego-CS Dataset

X-Ego-CS contains 101 high-skill Counter-Strike 2 matches, 2,168 rounds, 720 unique players, and about 273 player-view hours across eight maps, drawn from public FACEIT replays of leaderboard players (FACEIT, 2025). Every player on both teams is recorded, so each moment of a round is observed from up to ten synchronized first-person viewpoints.

Data curation. We parse game replay data with off-the-shelf tools such as AWPY and demoparser2 at 64 ticks/s (Xenopoulos & Silva, 2022; LaihoE, 2025), record each player's alive interval, and correct recorder latency with a timer-transition alignment. Fixed masks remove the heads-up display (HUD), including the minimap, kill feed, status, chat, inventory, and notifications (Appendix Figure 9), so models see only RGB gameplay and cannot exploit interface elements as shortcuts (Appendix B).

Benchmark tasks. Each data instance is a five-second, 4 FPS window from one observer, with a target about the observer (Self), allies (Teammate), opponents (Enemy), or game (Global), giving

12, 7, 5, and 9 tasks (Table 1; Appendix A). Self location is one of Mirage's 23 regions (Appendix Figure 12), whereas teammate and enemy locations are sparse sets of about three occupied regions, so we report ranking metrics alongside thresholded F1.

Table 1: **Task suite.** 33 linear probes on one observer's five-second video. Locations are predicted at 0, 5, 10, and 20 s and events at 5, 10, and 20 s. Outputs are C multi-class, L multi-label, B binary, R regression.

Target	Location	Events	Other state	Tasks
Self	Region (C)	Kill, death (B)	In combat (B), speed (R)	12
Teammate	Occupied regions (L)	–	In combat (B), alive count (C), speed (R)	7
Enemy	Occupied regions (L)	–	Alive count (C)	5
Global	–	Any kill (B)	Bomb planted, plant site, will plant, post-plant outcome, round winner (B), outcome reason (C)	9

Timing and splits. Location tasks are posed as nowcasts or forecasts. Nowcasts target the window midpoint, and an h-second forecast targets h seconds after the window ends (Appendix C.1). Each task draws about 5,000 Mirage examples split 70:15:15 by round, with Dust2 and Inferno for replication, and a controlled study holds out whole matches (26/9/9) before any training (Section 5.3).

4 Method

4.1 CECL objective

For player i, let V_i be a five-second window and $h_i = f_\theta(V_i)$ its encoding, and let a projector g_ϕ produce $u_i = g_\phi(h_i)/\|g_\phi(h_i)\|_2$. For distinct windows in a batch of size N, $m_{ij} = +1$ when i and j are teammates at the same round and time, and $m_{ij} = -1$ otherwise, so opponents, other times, and other rounds all serve as negatives. Following Zhai et al. (2023),

$$\mathcal{L}_{\text{CECL}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i}\log\sigma\left[m_{ij}\left(e^{a}u_i^{\top}u_j + b\right)\right]. \tag{1}$$

Here σ is the logistic sigmoid, and the scale e^a and bias b are learned from initial values of 10 and -3. CECL updates the last three transformer blocks, any trainable aggregation, and the projector, without labels or text.

Alignment across distance. Because our positives, unlike those of conventional cross-view learning, need not share any visible object (Figure 2), the objective could in principle be met by shortcuts (Geirhos et al., 2020) common to teammates, such as interface overlays, team side, round phase, or the map, so we design our evaluation to test them. We mask HUD overlays to rule out interface shortcuts and make every comparison within one map to rule out map-specific shortcuts. The gain shrinks when teammates are replaced by single views or by side- and phase-matched pseudo-teams (Section 6), it survives the removal of defining events and persistence (Section 5.4), and distant teammates become most similar at the same moment (Section 5.5).

4.2 Evaluation

We adapt CLIP, SigLIP2, DINOv2, and V-JEPA2 (Radford et al., 2021; Tschannen et al., 2025; Oquab et al., 2024; Assran et al., 2025) with CECL (Stage 1), then freeze each encoder, discard the projector, and train one linear probe (Alain & Bengio, 2016) per task on a single observer's window (Stage 2), against probes on the unadapted encoder. Main comparisons use the three image encoders with frame averaging, so adaptation adds no trainable module; V-JEPA2 and later temporal models, which add DINOv3 (Siméoni et al., 2025) for diagnostics, also train an aggregation module (Table 13). Probes train for 20 epochs with validation-loss selection, and the controlled study also fits them to convergence (Appendices C–D.2).

We report accuracy, class-macro recall, multi-label macro-F1 at threshold 0.5, and speed R^2, plus threshold-free macro-AUROC and average precision (AP) for locations, over three or four probe repeats (Appendix F.1).

Table 2: **A selective gain–cost profile across encoders and maps** ($\times 100$). Baseline and CECL are equally weighted means over four encoders on Mirage, with Δ their difference; bold marks the higher score. The per-map columns give SigLIP2 CECL-minus-baseline means on each map, with the same 28 tasks and probe seeds 1/2. Per-encoder differences and full scores are in Appendices F and I.

		Mirage, four encoders			SigLIP2 Δ per map		
Target	**Metric**	**Baseline**	**CECL**	Δ	**Dust2**	**Inferno**	**Mirage**
Self	Accuracy	59.6	**60.6**	+1.0	+1.5	−0.6	+1.9
	Macro recall	**40.6**	37.4	−3.2	+2.5	−0.8	−1.2
	R^2	**52.1**	38.2	−13.9	−16.9	−26.2	−17.2
Teammate	Accuracy	67.6	**69.8**	+2.2	+4.4	−1.0	+5.8
	Macro recall	30.3	**36.7**	+6.4	+4.0	+1.2	+3.4
	Macro-F1	3.9	**14.3**	+10.4	+5.0	+3.7	+10.3
	R^2	**14.2**	11.8	−2.4	−21.9	−22.4	−24.7
Enemy	Macro recall	29.9	**36.1**	+6.2	+3.7	+2.1	+3.8
	Macro-F1	5.5	**15.5**	+10.0	+6.3	+4.8	+10.5
Global	Accuracy	56.9	**61.5**	+4.5	−1.4	+1.1	+3.4
	Macro recall	22.3	**24.1**	+1.9	+0.1	+2.6	+1.5

5 Experiments

Each experiment below answers one question. We first characterize the shift across encoders, horizons, and maps (Sections 5.1–5.2), then test its cause on held-out matches, without shortcuts, and in the embedding space (Sections 5.3–5.5), and finally test its reach to language and to control (Sections 5.6–5.7). Section 6 ablates the design choices behind it.

5.1 Gain–cost across encoders

Because view-selection theory (Tian et al., 2020b) predicts a cost as well as a gain, we read out all four perspectives, comparing each adapted encoder with its unadapted baseline on identical Mirage splits (Table 2). We find that alignment makes other players' locations substantially easier to read, while the observer's own motion becomes harder to read. For the three frame-averaged encoders, teammate and enemy location macro-F1 rise from 4.7 and 6.4 to 13.6 and 15.0, whereas self-location recall falls from 42.1 to 36.8 and self-speed R^2 from 47.6 to 36.8. Enemy locations gain about as much as teammate locations, although opponents only ever serve as negatives, so alignment captures the shared tactical situation, including the opponents teammates respond to, rather than teammates alone.

Forecast horizons. Future locations are not directly visible in a single view, so forecasts test more than visible content. The gain holds at every horizon and for all four encoders. Teammate macro-F1, for example, rises from 7.8 to 25.6 for nowcasts and from 0.7 to 3.7 at 20 seconds (Figure 3).

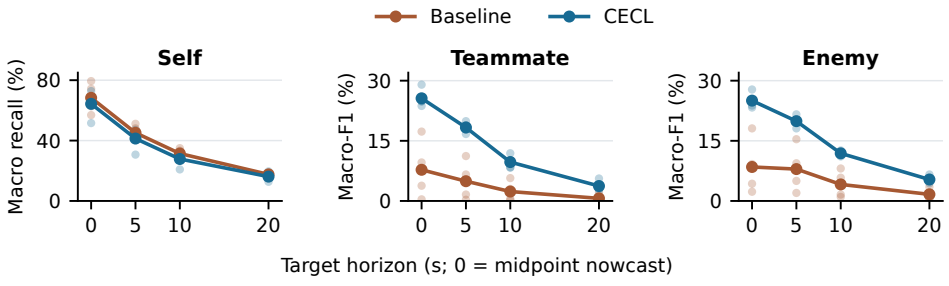


Figure 3: **Other-player location gains persist across horizons.** Lines average four encoders; points show individual scores. Self uses macro recall, other players macro-F1; vertical scales differ. Absolute forecast scores decrease with horizon.

Table 3: **Encoders share one gain–cost direction** ($\times 100$). CECL minus baseline on Mirage, with location averaged over four horizons.

	Location			Self motion
Encoder	**Self / recall**	**Team / F1**	**Enemy / F1**	**Speed / R^2**
CLIP	-8.0	$+12.9$	$+13.1$	-12.2
SigLIP2	-3.5	$+8.4$	$+8.3$	-13.9
DINOv2	-4.4	$+5.2$	$+4.3$	-6.1
V-JEPA2	$+3.0$	$+15.1$	$+14.3$	-23.4

Selective cost. We find that the cost is selective rather than a general loss of self-knowledge (Table 3). Self-speed R^2 falls for all four encoders, yet self binary events improve by about 1.0 point and V-JEPA2's self-location recall rises by 3.0 points. This shows that alignment trades the observer's fine-grained geometry and motion for team state.

Map replication. To test whether the shift generalizes beyond one map's layout, we repeat the SigLIP2 benchmark within the Dust2 and Inferno maps. Teammate and enemy location macro-F1 improve and speed R^2 falls on every map (Table 2, per-map columns), while self-location and event effects vary (Appendix I), so the spatial gain and the motion cost recur across all three layouts.

5.2 Threshold robustness

Because occupancy labels are sparse, we ensure that an F1 gain does not merely reflect a shifted threshold by evaluating threshold-free macro-AUROC, recovered from the 246 saved probe checkpoints. It improves in all 24 primary comparisons (0.685 to 0.724) and all 32 including V-JEPA2, and all 123 same-seed differences are positive (Figure 4). The gains (0.006 to 0.121) exceed probe-repeat variability (median SD 0.001 to 0.003) by at least an order of magnitude (Appendix F.1). Alignment therefore improves how locations are ranked, not merely where a threshold falls.

5.3 Held-out replication

Since the benchmark splits by round, matches can recur between training and test, and its 20-epoch probes may be under-trained. To rule out both, we run a controlled SigLIP2 study on Mirage that holds out whole matches before either stage and also fits probes to convergence. Its three adaptation arms share data, optimizer, updates, projector, and augmentation, and differ only in what counts as a positive. C pairs same-time teammates, S pairs augmentations of one clip, and H regroups the same clips into pseudo-teams from other rounds matched on side and 10-second phase (Section 6). As a sanity check, the unadapted encoder reproduces the benchmark AUROC within 0.01.

Replication result. On unseen matches, CECL still raises other-player AUROC from 0.694 to 0.735 and AP from 0.266 to 0.297, and the 20-epoch probe reproduces the benchmark result (0.723 versus 0.724). The gain therefore depends neither on shared matches nor on weak probes (Table 4a).

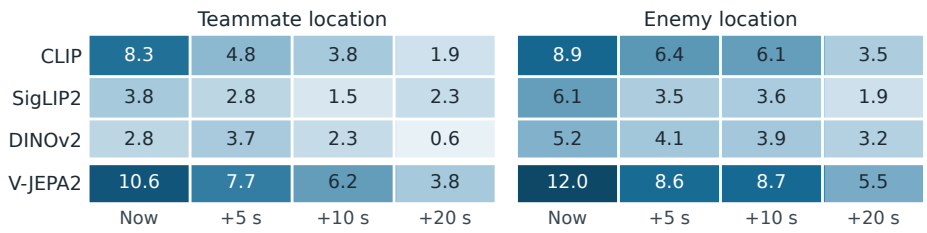


Figure 4: **Alignment improves location ranking for every encoder and horizon.** CECL minus baseline macro-AUROC ($\times 100$) on Mirage, darker for larger gains. All 24 image-encoder comparisons (top three rows) and all 32 including V-JEPA2 improve, as do all 123 same-seed differences.

Table 4: **The gain replicates on held-out matches and survives targeted restrictions.** **(a)** SigLIP2 on Mirage, converged probes, 9 held-out test matches. Other-player averages the eight occupancy tasks, matched targets average 0/+5/+10 s on identical windows, and large-batch uses few updates (three seeds). **(b)** Native-unit differences (bomb accuracy, alive-count recall, location macro-F1); Pos. counts improved map–encoder pairs of five, and n is the median subset size.

(a) Controlled replication

AUROC unless noted	B	C	S	H
Other-player location	.694	**.735**	.700	.723
macro-AP	.266	**.297**	.269	.279
20-epoch probe	.685	**.723**	.695	.713
Matched teammate	.653	**.708**	.663	n/a
Matched enemy	.642	**.695**	.658	n/a
Matched self	.934	.932	.937	.926
Self speed R^2	.304	.288	.292	n/a
Large-batch, other-player	.694	.687	.678	.684

(b) Targeted restrictions

Target	Restr.	Full	Subset	Pos.	n
Bomb planted	No plant	+.056	+.057	5/5	748
Enemy alive	No death	+.029	+.031	5/5	393
Team alive	No death	+.026	+.018	4/5	430
Enemy loc., 10s	Disjoint	+.039	+.043	4/5	117
Enemy loc., 5s	Disjoint	+.064	+.049	5/5	45
Team loc., 10s	Disjoint	+.033	+.023	4/5	146

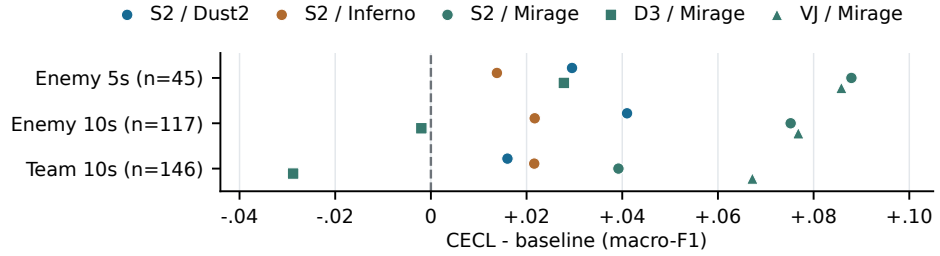


Figure 5: **Forecast gains when targets move to new regions.** Each point is one map–encoder pair (S2 SigLIP2, D3 DINOv3, VJ V-JEPA2), and n is the median subset size. Gains recur across pairs.

Matched perspectives. Because self location is a single region whereas other-player location is a set of regions, we separate perspective from output format with a matched suite that predicts the region of the observer, of one teammate, and of one enemy on identical windows with the same 23-way output. CECL then raises teammate and enemy AUROC by 0.055 and 0.053 while leaving self unchanged (-0.002), so the shift follows whose state is read out.

5.4 Shortcut tests

To ensure that the gain does not rest on simple cues, we rescore existing predictions on subsets that exclude each cue, across five map–encoder pairs (Appendix G). We remove windows that show the bomb plant or a death, which reveal the target directly, and forecasts whose targets stay in their current regions, which persistence alone could solve.

The gain survives all six restrictions (Table 4b). The event-free bomb gain (+0.057) matches the full split (+0.056), enemy forecasts at 5 seconds improve for all five pairs even when enemies must reach new regions, and the only two losses come from DINOv3 (Figure 5). The gain therefore does not rest on the defining event or on persistence.

5.5 Teammate similarity analysis

The two readings of CECL make different predictions. If it relied on shared visible content, only nearby teammates would become similar. If it captures the team's situation, distant teammates would too, most strongly at the same moment. For 16,314 clips from 163 unseen Mirage test rounds, encoded by SigLIP2 in the temporal-encoder setting used for the three-map diagnostics, we score each synchronized teammate's cosine similarity against same-side clips from other rounds in the same 10-second phase, as an AUROC where 0.5 is chance (Appendix H). Before alignment, similarity tracks proximity, and teammates more than 1,500 units apart score 0.41. After CECL they score 0.70 [0.68, 0.72], with the gain growing with distance (Figure 6a), and similarity falls from 0.71

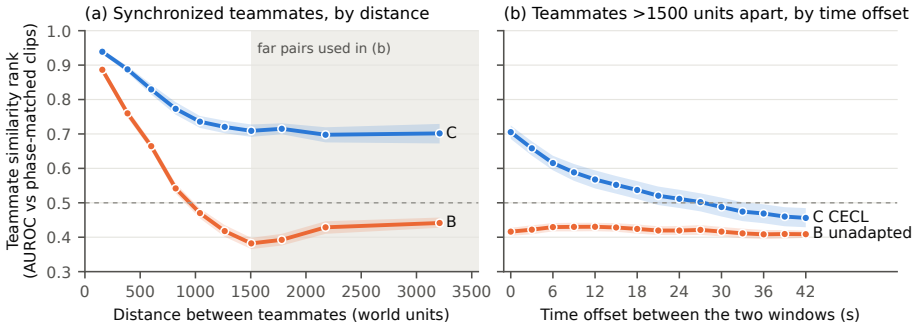


Figure 6: **CECL makes distant teammates similar, most strongly at the same moment.** AUROC of a synchronized teammate's cosine similarity against same-side clips from other rounds in the same 10-second phase, with 95% match-clustered bootstrap bands, on unseen Mirage test rounds. (a) Without adaptation, similarity tracks proximity and falls below chance for distant pairs, whereas CECL raises it most for teammates too far apart to share a view. (b) For pairs more than 1,500 units apart, CECL similarity peaks for synchronized windows and decays with time offset.

Table 5: **Concept rank changes across 129 template–map pairs.** Raw and rank-controlled (residual) mean changes, with the number of pairs that agree in direction and that reach nominal $p < .05$ after rank control.

Group	Raw median	Raw mean	Residual mean	Direction agreement	Nominal $p < .05$
Self	−30.2	−30.1	−20.7	129/129	126/129
Teammate	+24.8	+24.0	+15.6	121/129	100/129
Enemy	−9.8	−10.4	−5.0	95/129	55/129
Global	+12.1	+12.0	+6.2	113/129	49/129
Spatial	+5.4	+4.5	+3.9	92/129	42/129

when synchronized to chance about 27 seconds apart (Figure 6b). Cues that stay constant within a round cannot produce this decay, so alignment captures what teammates share at a given moment, even when they are far apart.

5.6 Language probe experiment

Aligning distant teammates likely requires matching through high-level concepts, so we test whether alignment changes the concepts a vision-language encoder names. SigLIP2's frozen text encoder ranks 250 phrases, 50 per target group, against 100 test clips per map under 43 prompt templates on three maps (129 pairs; Appendix J.1). After removing the change expected from each phrase's initial rank, Self phrases fall under all 129 pairs and Teammate phrases rise under 121, with weaker Global, Enemy, and Spatial effects (Table 5; Figure 7; Appendix J). Without text supervision, alignment thus shifts which phrases the representation ranks highly, from self-centric toward teammate phrases.

5.7 Control transfer experiment

To test whether an actor can use a representation that encodes teammates, we add the cross-ego term to MAPPO in a three-seed ViZDoom (Kempka et al., 2016) coordination task. At 5M frames, CECL-MAPPO reaches a normalized return of 0.734 versus 0.565 for MAPPO (Table 6), suggesting that the cross-ego signal can also benefit control.

6 Ablations

We ablate two design choices that affect whether the shift appears (Table 4a). **(1) Positive pairing.** To test whether any in-domain adaptation would produce the gain, we vary only what counts as a positive. Single-view adaptation (S) barely moves other-player AUROC (0.700 versus 0.694),

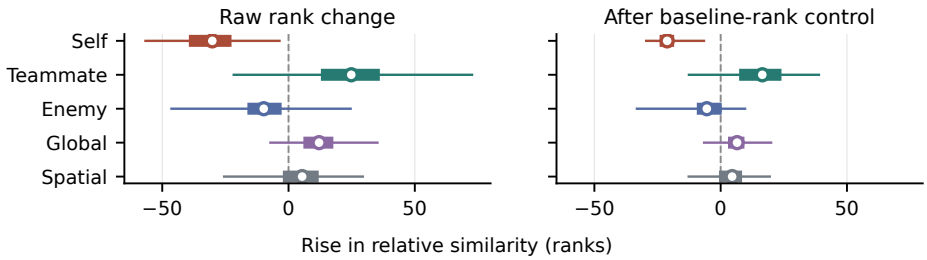


Figure 7: **Prompt robustness before and after rank control.** Dots mark medians, thick bars the interquartile range (IQR), and thin bars the range over 129 template–map pairs. Positive values indicate a rise. Rank control removes the change expected from each phrase's initial rank.

Table 6: **ViZDoom control shows higher return at a 5M-frame budget.** Normalized deterministic evaluation return over three seeds; interquartile mean (IQM) with 95% bootstrap interval (Agarwal et al., 2021), which equals the mean for three scores. Lower optimality gap is better. The 10M-frame checkpoint is reported in Appendix K.

Method (5M frames)	**IQM [95% CI]**	**Median**	**Optimality gap**
MAPPO	.565 [.554, .580]	.563	.435
CECL-MAPPO	**.734** [.729, .743]	**.731**	**.266**

pseudo-teams sharing side and phase (H) already recover about 70% of the gain (0.723), showing how much team state these coarse variables carry, and true synchronization adds the rest (C, 0.735), exceeding both controls on all eight location tasks. The gain thus grows with how much coordinated state paired views share, and CECL separates synchronized teammates from exactly such side- and phase-matched clips (Section 5.5). **(2) Optimization regime.** With 1,024-video batches and about seven times fewer updates (three seeds), the same objective yields no gain (0.687 versus 0.694) and larger self-state costs, so how CECL is optimized, not only what it pairs, matters for whether the shift appears.

7 Limitations and Conclusion

Limitations and future work. Our study rests on three premises that reach beyond this paper. First, it treats coordinated play in competitive games as a window onto teamwork. Esports teams act under fixed rules and symmetric roles, with a well-defined team state, whereas human-agent teams in search and rescue or manufacturing mix roles, embodiments, and goals, and rarely leave synchronized multi-view recordings. Whether cross-ego alignment carries over to such teams, and what data would supply their positive pairs, is the most important open question. Second, we measure what a frozen representation makes readable, not what an agent understands. A single view now carries more team context, but this does not show that the model represents what teammates intend or believe, which coordination with people ultimately requires. Third, the shift is a trade, since alignment costs the observer's own motion, whereas a capable teammate needs both an egocentric and an allocentric sense. We leave these questions to future work, in particular learning both senses without sacrificing either, and testing whether better team perception yields better joint behavior with human partners.

Conclusion. We asked whether an agent's perception can learn about its team from recordings of coordinated behavior. Aligning synchronized teammates, who often share no view, makes a single first-person view reveal more about where teammates and opponents are and will be, at a selective cost to the observer's own motion, and controls tie the effect to synchronized correspondence beyond generic adaptation and the shortcuts we tested. Every recorded match or drill already pairs views that no single member sees. Learning from these recordings at scale may teach agents far more of teamwork than where partners are, and bring closer agents that join human teams without being told where everyone is.

AI USE STATEMENT

LLM-based tools, including Claude Code, Cursor, and Codex, assisted with writing, polishing, and drafting and restructuring sections, retrieval of related work and citation suggestions, technical explanations, data-processing and experiment-support code, and table and figure preparation. This revision used these tools to audit existing results, correct presentation inconsistencies, and prepare the ICLR manuscript. LLM outputs are not experimental measurements. No LLM is a component of CECL or a source of training/evaluation labels. The authors remain responsible for verifying all code, analyses, citations, and claims.

ETHICS STATEMENT

The data derive from public high-skill gameplay replays rather than researcher-directed human interaction. Public availability does not remove privacy or reuse obligations, because player identities, styles, and strategies can remain identifiable. Access and redistribution should respect platform and tournament terms, with documentation of intended research use and limitations. Extending cross-view inference to real human teams could enable surveillance or intrusive monitoring. These experiments do not validate such uses or operational decisions about people.

REPRODUCIBILITY STATEMENT

The source bundle includes all plotted means, 246 original location-probe and 336 later map-diagnostic run metrics, repeat statistics for every comparison, all 250 concepts and 43 templates, and figure-generation scripts. Anonymized records reproduce the reported aggregates; the retained-label timing audit verifies the nowcast convention. The appendix separates benchmark and diagnostic configurations and documents split provenance and target censoring. A release on acceptance is planned for X-Ego-CS documentation, preprocessing and evaluation code, match-held-out split manifests with hashes, shuffled-pair assignments, and the controlled study's cached features and probe predictions.

REFERENCES

Rishabh Agarwal, Max Schwarzer, Pablo Samuel Castro, Aaron C Courville, and Marc Bellemare. Deep reinforcement learning at the edge of the statistical precipice. *Advances in neural information processing systems*, 34:29304–29320, 2021.

Guillaume Alain and Yoshua Bengio. Understanding intermediate layers using linear classifier probes. *arXiv preprint arXiv:1610.01644*, 2016.

Stefano V Albrecht and Peter Stone. Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artificial Intelligence*, 258:66–95, 2018.

Relja Arandjelović and Andrew Zisserman. Look, listen and learn. In *IEEE International Conference on Computer Vision (ICCV)*, pp. 609–617, 2017.

Mahmoud Assran, Adrien Bardes, David Fan, Quentin Garrido, Russell Howes, Mojtaba Komeili, Matthew Muckley, Ammar Rizvi, Claire Roberts, Koustuv Sinha, Artem Zholus, Sergio Arnaud, Abha Gejji, Ada Martin, Francois Robert Hogan, Daniel Dugas, Piotr Bojanowski, Vasil Khalidov, Patrick Labatut, Francisco Massa, Marc Szafraniec, Kapil Krishnakumar, Yong Li, Xiaodong Ma, Sarath Chandar, Franziska Meier, Yann LeCun, Michael Rabbat, and Nicolas Ballas. V-jepa 2: Self-supervised video models enable understanding, prediction and planning. *arXiv preprint arXiv:2506.09985*, 2025.

Bowen Baker, Ingmar Kanitscheider, Todor Markov, Yi Wu, Glenn Powell, Bob McGrew, and Igor Mordatch. Emergent tool use from multi-agent autocurricula. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=SkxpxJBKwS.

Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, et al. Dota 2 with large scale deep reinforcement learning. *arXiv preprint arXiv:1912.06680*, 2019.

Neil Burgess. Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Sciences*, 10(12):551–557, 2006.

Janis A Cannon-Bowers, Eduardo Salas, and Sharolyn Converse. Shared mental models in expert team decision making. In *Individual and Group Decision Making*, pp. 221–246. Lawrence Erlbaum Associates, Inc, 1993.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PmLR, 2020.

Imant Daunhawer, Alice Bizeul, Emanuele Palumbo, Alexander Marx, and Julia E. Vogt. Identifiability results for multimodal contrastive learning. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2303.09166.

Leslie A. DeChurch and Jessica R. Mesmer-Magnus. The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, 95(1):32–53, January 2010. ISSN 0021-9010. doi: 10.1037/a0017328. URL http://dx.doi.org/10.1037/a0017328.

Adrien Deliege, Anthony Cioppa, Silvio Giancola, Meisam J. Seikavandi, Jacob V. Dueholm, Kamal Nasrollahi, Bernard Ghanem, Thomas B. Moeslund, and Marc Van Droogenbroeck. Soccernet-v2: A dataset and benchmarks for holistic understanding of broadcast soccer videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 4508–4519, June 2021.

Mica R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37 (1):32–64, 1995.

Elliot E. Entin and Daniel Serfaty. Adaptive team coordination. *Human Factors*, 41(2):312–325, 1999.

FACEIT. FACEIT Platform. https://www.faceit.com/en, 2025. Accessed: 2025-08-04.

Jakob Foerster, Ioannis Alexandros Assael, Nando de Freitas, and Shimon Whiteson. Learning to communicate with deep multi-agent reinforcement learning. In *Advances in neural information processing systems*, volume 29, 2016.

Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A. Wichmann. Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11):665–673, 2020.

Kristen Grauman, Andrew Westbury, Eugene Byrne, Zachary Chavis, Antonino Furnari, Rohit Girdhar, Jackson Hamburger, Hao Jiang, Miao Liu, Xingyu Liu, et al. Ego4d: Around the world in 3,000 hours of egocentric video. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 18995–19012, 2022.

Kristen Grauman, Andrew Westbury, Lorenzo Torresani, Kris Kitani, Jitendra Malik, et al. Ego-exo4d: Understanding skilled human activity from first- and third-person perspectives. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024. URL https://arxiv.org/abs/2311.18259.

Edward Gu, Ho Chit Siu, Melanie Platt, Isabelle Hurley, Jaime Peña, and Rohan Paleja. Enabling rapid shared human-ai mental model alignment via the after-action review. *arXiv preprint arXiv:2503.19607*, 2025.

Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 9729–9738, 2020.

Keller Jordan, Yuchen Jin, Vlado Boza, Jiacheng You, Franz Cesista, Laker Newhouse, and Jeremy Bernstein. Muon: An optimizer for hidden layers in neural networks, 2024. URL https://kellerjordan.github.io/posts/muon/.

Michał Kempka, Marek Wydmuch, Grzegorz Runc, Jakub Toczek, and Wojciech Jaśkowski. Vizdoom: A doom-based ai research platform for visual reinforcement learning. In *2016 IEEE conference on computational intelligence and games (CIG)*, pp. 1–8. IEEE, 2016.

Roberta L. Klatzky. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, volume 1404 of *Lecture Notes in Computer Science*, pp. 1–17. Springer, 1998.

LaihoE. demoparser: Counter-strike 2 replay parser for python and javascript. https://github.com/LaihoE/demoparser, 2025. Accessed: 2025-09-14.

Yen-Cheng Liu, Junjiao Tian, Nathaniel Glaser, and Zsolt Kira. When2com: Multi-agent perception via communication graph grouping. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020. URL https://arxiv.org/abs/2006.00176.

Yat Long Lo, Biswa Sengupta, Jakob Foerster, and Michael Noukhovitch. Learning multi-agent communication with contrastive learning. *arXiv preprint arXiv:2307.01403*, 2023.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in neural information processing systems*, 30, 2017.

Wenhao Ma, Yu-Chen Chang, Jie Yang, Yu-Kai Wang, and Chin-Teng Lin. Contrastive learning-based agent modeling for deep reinforcement learning. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2025.

NVIDIA Corporation. Nvidia app. https://www.nvidia.com/en-us/geforce/geforce-experience/shadowplay/, 2025. Accessed: 2025-09-14.

Thomas O'Neill, Nathan McNeese, Amy Barron, and Beau Schelble. Human-autonomy teaming: A review and analysis of the empirical literature. *Human Factors*, 64(5):904–938, 2022.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2: Learning robust visual features without supervision. *Transactions on Machine Learning Research*, 2024.

Andrew Owens and Alexei A. Efros. Audio-visual scene analysis with self-supervised multisensory features. In *European Conference on Computer Vision (ECCV)*, pp. 631–648, 2018.

Tim Pearce and Jun Zhu. Counter-strike deathmatch with large-scale behavioural cloning. In *IEEE Conference on Games (CoG)*, pp. 104–111, 2022.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pp. 8748–8763. PmLR, 2021.

Tabish Rashid, Mikayel Samvelyan, Christian Schroeder de Witt, Gregory Farquhar, Jakob Foerster, and Shimon Whiteson. Monotonic value function factorisation for deep multi-agent reinforcement learning. *Journal of Machine Learning Research*, 21(178):1–51, 2020. URL http://jmlr.org/papers/v21/20-081.html.

Ramón Rico, Miriam Sánchez-Manzanares, Francisco Gil, and Cristina Gibson. Team implicit coordination processes: A team knowledge-based approach. *Academy of Management Review*, 33(1):163–184, 2008.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Pierre Sermanet, Corey Lynch, Yevgen Chebotar, Jasmine Hsu, Eric Jang, Stefan Schaal, and Sergey Levine. Time-contrastive networks: Self-supervised learning from video. In *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1134–1141, 2018.

Oriane Siméoni, Huy V. Vo, Maximilian Seitzer, Federico Baldassarre, Maxime Oquab, Cijo Jose, Vasil Khalidov, Marc Szafraniec, Seungeun Yi, Michaël Ramamonjisoa, Francisco Massa, Daniel Haziza, Luca Wehrstedt, Jianyuan Wang, Timothée Darcet, Théo Moutakanni, Leonel Sentana, Claire Roberts, Andrea Vedaldi, Jamie Tolan, John Brandt, Camille Couprie, Julien Mairal, Hervé Jégou, Patrick Labatut, and Piotr Bojanowski. DINOv3, 2025.

Yonglong Tian, Dilip Krishnan, and Phillip Isola. Contrastive multiview coding. In *European Conference on Computer Vision*, 2020a. URL https://www.ecva.net/papers/eccv_2020/papers_ECCV/html/1453_ECCV_2020_paper.php.

Yonglong Tian, Chen Sun, Ben Poole, Dilip Krishnan, Cordelia Schmid, and Phillip Isola. What makes for good views for contrastive learning? In *Advances in Neural Information Processing Systems*, volume 33, pp. 6827–6839, 2020b. URL https://papers.nips.cc/paper_files/paper/2020/file/4c2e5eaae9152079b9e95845750bb9ab-Paper.pdf.

Michael Tschannen, Alexey Gritsenko, Xiao Wang, Muhammad Ferjad Naeem, Ibrahim Alabdulmohsin, Nikhil Parthasarathy, Talfan Evans, Lucas Beyer, Ye Xia, Basil Mustafa, et al. Siglip 2: Multilingual vision-language encoders with improved semantic understanding, localization, and dense features. *arXiv preprint arXiv:2502.14786*, 2025.

Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *nature*, 575(7782):350–354, 2019.

Peter Xenopoulos and Claudio Silva. Esta: An esports trajectory and action dataset. *arXiv preprint arXiv:2209.09861*, 2022.

Chao Yu, Akash Velu, Eugene Vinitsky, Jiaxuan Gao, Yu Wang, Alexandre Bayen, and Yi Wu. The surprising effectiveness of ppo in cooperative multi-agent games. *Advances in neural information processing systems*, 35:24611–24624, 2022.

Peihong Yu, Manav Mishra, Syed Zaidi, and Pratap Tokekar. Tactic: Task-agnostic contrastive pre-training for inter-agent communication. *arXiv preprint arXiv:2501.02174*, 2025.

Xiaohua Zhai, Basil Mustafa, Alexander Kolesnikov, and Lucas Beyer. Sigmoid loss for language image pre-training. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 11975–11986, 2023.

A PREDICTION TASK CATALOG

Table 7 lists the 33 probing tasks, with repeated horizons grouped into one row. Every task takes a single observer's five-second video as input, and Teammate targets exclude that observer. Location targets are occupied map regions rather than agent identities; the reference tables use 23 regions, and other maps have their own callout vocabularies (Figures 10–13). Future-event targets ask whether an event occurs within the stated horizon. Round and bomb outcome labels refer to eventual outcomes, even when the task identifier names no horizon.

Table 7: **The 33-task catalog.** Forms are binary (Bin.), multiclass (MC), multi-label (ML), and regression (Reg.); horizon 0 denotes a nowcast.

Target	Quantity	Form	Horizon (s)	Tasks
Self	Location	MC	0, 5, 10, 20	4
Self	Kill	Bin.	5, 10, 20	3
Self	Death	Bin.	5, 10, 20	3
Self	In combat	Bin.	0	1
Self	Speed	Reg.	0	1
Teammate	Occupied regions	ML	0, 5, 10, 20	4
Teammate	In combat	Bin.	0	1
Teammate	Alive count	MC	0	1
Teammate	Speed	Reg.	0	1
Enemy	Occupied regions	ML	0, 5, 10, 20	4
Enemy	Alive count	MC	0	1
Global	Any kill	Bin.	5, 10, 20	3
Global	Bomb planted	Bin.	0	1
Global	Plant site / will plant	Bin.	Eventual	2
Global	Post-plant outcome	Bin.	Eventual	1
Global	Round winner	Bin.	Eventual	1
Global	Round outcome reason	MC	Eventual	1
Total				33

B DATASET DETAILS

This section describes how recordings were collected, synchronized with trajectories, and masked, and it reports the label distributions of the sampled task pools.

B.1 DATA COLLECTION

We collected high-skill Counter-Strike 2 demo files (`.dem`) from the matchmaking platform FACEIT (FACEIT, 2025). Using FACEIT's official API and following its website policies, we retrieved the public Elo leaderboard and the match histories of the top 100 players and downloaded their most recent matches. The off-the-shelf parsers AWPY and demoparser2 (Xenopoulos & Silva, 2022; LaihoE, 2025) extract structured metadata, player state-action trajectories, and game events from each replay.

Video comes from an automated recording system built on the Counter-Strike 2 demo playback engine. Demo metadata give the exact tick range of every player's alive period in each round. Console commands jump to the start tick (`demo_gototick`), lock the camera on the player (`spec_player`), and control playback (`demo_resume`, `demo_pause`), so only the segments of each player's active life are replayed, deterministically. NVIDIA ShadowPlay (NVIDIA Corporation, 2025), triggered automatically through hotkeys, records each player's alive interval in every round. After quality filtering, the dataset contains 101 matches across 8 maps, with synchronized egocentric video from all players and their state-action trajectories.



Figure 8: **Frame-level video–trajectory alignment.** HUD timer digits are cropped and read by template-matching OCR. The first frame at which the timer changes corrects recorder-trigger latency before video frames are paired with 64 Hz state-action ticks.

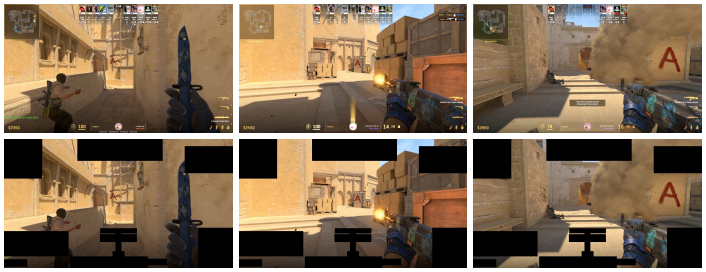


Figure 9: **Observation protocol.** Original frames (top) and masked RGB model inputs (bottom). Fixed masks remove the minimap, kill feed, status, chat, inventory, and notification regions. Inputs are 5-second, 20-frame clips without audio or trajectories.

B.2 Video–trajectory alignment

Demo playback provides exact tick ranges, but the external recorder can add a small, variable delay between the recording hotkey and the first encoded frame. A post-hoc step removes this recorder-trigger latency by synchronizing each video with its parsed trajectory at frame precision (Figure 8).

The alignment uses the round timer at the top of the HUD. For each video, we crop the three timer digit regions in every frame and read them by template-matching OCR against digit templates. The first frame $f^\star$ at which the displayed timer changes by one second is an observable synchronization event shared by the video and the demo timeline. With measured video frame rate r and elapsed game time $t_{\text{game}}(f^\star)$ implied by the timer value after the transition, the video-to-trajectory offset is

$$\Delta = \frac{f^\star}{r} - t_{\text{game}}(f^\star).$$

Each frame f is assigned corrected game time $f/r - \Delta$ and matched to the corresponding 64 Hz state-action tick, so frames are paired with ticks through an observed timer transition rather than the nominal recording trigger. Precision is limited by the frame rate and timer detection, and we do not report an independent distribution of residual alignment errors.

UI masking. A fixed mask blacks out all heads-up display (HUD) regions before a frame reaches the visual backbone, which removes direct interface readouts. The masked regions include the minimap, player status indicators (health, armor, ammunition), kill feed, chat box, on-screen notifications, and inventory. Because these elements are rendered at fixed screen positions, the mask is defined once per resolution and applied identically to every frame of every player and match. Figure 9 shows representative frames before and after masking. Gameplay content such as teammates, enemies, and environment geometry is preserved, while covered regions are unavailable to the encoder.

B.3 LABEL DISTRIBUTION

Tables 8–11 summarize the task-specific sampled pools before the round-level split; they are not held-out test counts. Sampling is balanced where feasible, with equal numbers of instances per class for binary and multiclass tasks and diverse label combinations for multi-label tasks.

Table 8: **Binary-task label distributions.** Horizons with identical statistics are grouped, and n is per task. All pools are balanced except post-plant outcome (1.2:1).

Target	Tasks (horizons)	Tasks	n	Class 0 / 1 (%)
Self	Kill, death (5, 10, 20 s)	6	5000	50.00 / 50.00
	In combat	1	5000	50.00 / 50.00
Teammate	In combat	1	5000	50.00 / 50.00
Global	Imminent kill (5, 10, 20 s)	3	5000	50.00 / 50.00
	Bomb state, site, will plant; round winner	4	5000	50.00 / 50.00
	Post-plant outcome	1	5000	54.46 / 45.54

Table 9: **Multiclass-task label distributions.** Self location has 23 region classes, the sampled alive-count pools contain five observed classes, and round outcome has five naturally imbalanced classes. Imbalance is the ratio of the largest to the smallest class count.

Target	Task	Horizon (s)	n	Classes	Min / max count	Imbalance
Self	Location	0	4952	23	169 / 218	1.29:1
Self	Location	5	4981	23	198 / 218	1.10:1
Self	Location	10	4970	23	187 / 218	1.17:1
Self	Location	20	4921	23	138 / 218	1.58:1
Teammate	Alive count	0	5000	5	1000 / 1000	1.00:1
Enemy	Alive count	0	5000	5	1000 / 1000	1.00:1
Global	Round outcome	Eventual	5000	5	60 / 2063	34.38:1

Table 10: **Multi-label location label distributions.** Every task predicts occupancy of 23 map regions. The positive rate gives the range of per-label frequencies, and labels per example give the mean with the observed range in parentheses.

Target	Horizon (s)	n	Positive rate (%)	Labels per example
Teammate	0	5000	2.08–32.52	3.0 (1–4)
Teammate	5	5000	2.82–30.84	3.01 (1–4)
Teammate	10	5000	3.00–32.10	2.99 (1–4)
Teammate	20	5000	2.68–28.96	2.80 (1–4)
Enemy	0	5000	2.20–36.12	3.49 (1–5)
Enemy	5	5000	3.00–35.88	3.48 (1–5)
Enemy	10	5000	2.54–34.64	3.45 (1–5)
Enemy	20	5000	3.02–31.58	3.15 (1–5)

Table 11: **Regression-task label distributions.** Speed is normalized to [0, 1].

Target	Task	n	Range	Mean	Median [Q25, Q75]
Self	Speed	5000	[0.0, 1.0]	0.315 ± 0.214	0.284 [0.122, 0.534]
Teammate	Speed	5000	[0.0, 0.931]	0.309 ± 0.212	0.280 [0.114, 0.523]

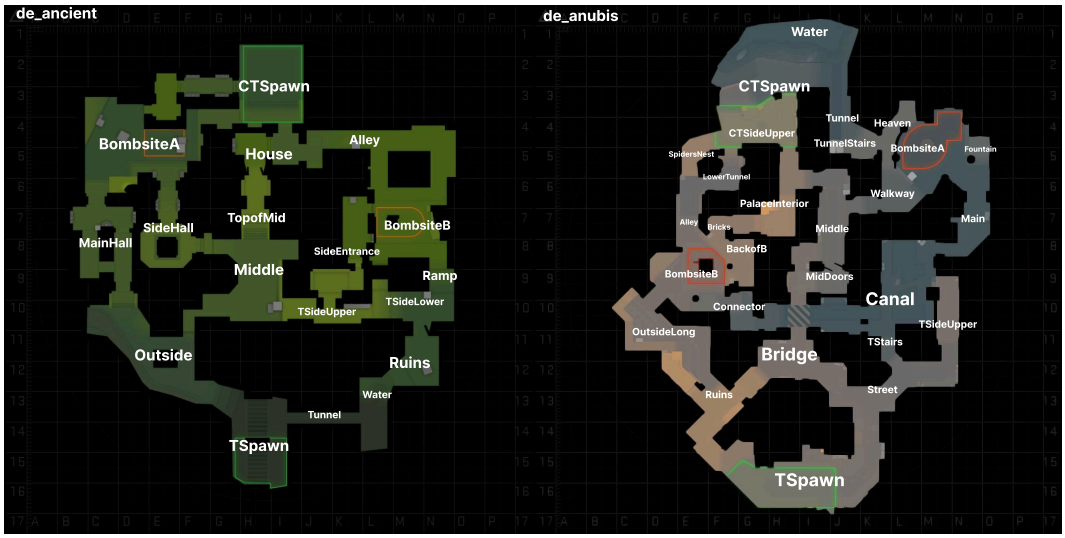


Figure 10: **Map callouts for Ancient and Anubis.** Region names define the location labels of each map. The eight-map atlas is split into four figures of two maps each for legibility.

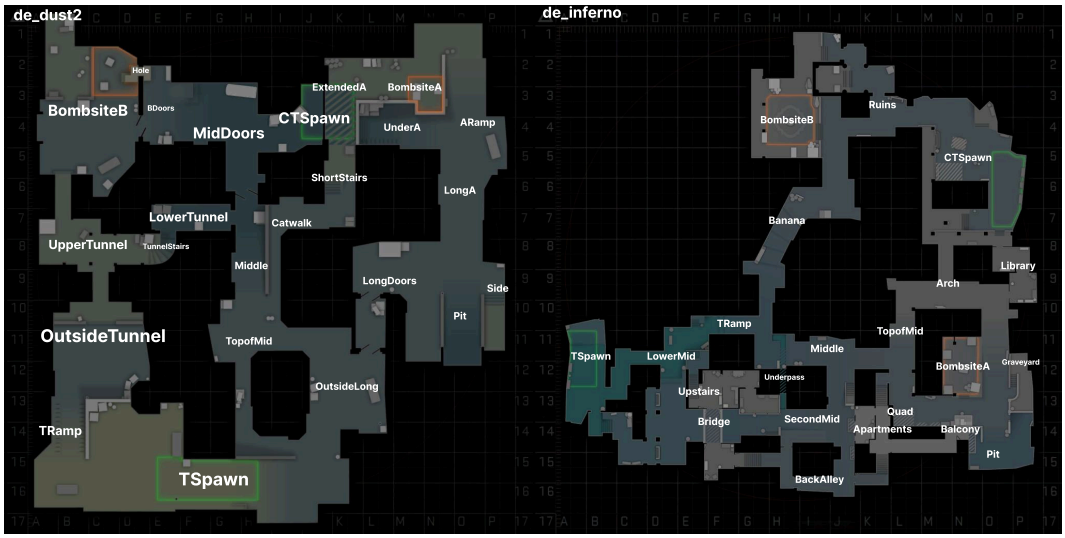


Figure 11: **Map callouts for Dust2 and Inferno.** Continued from Figure 10.

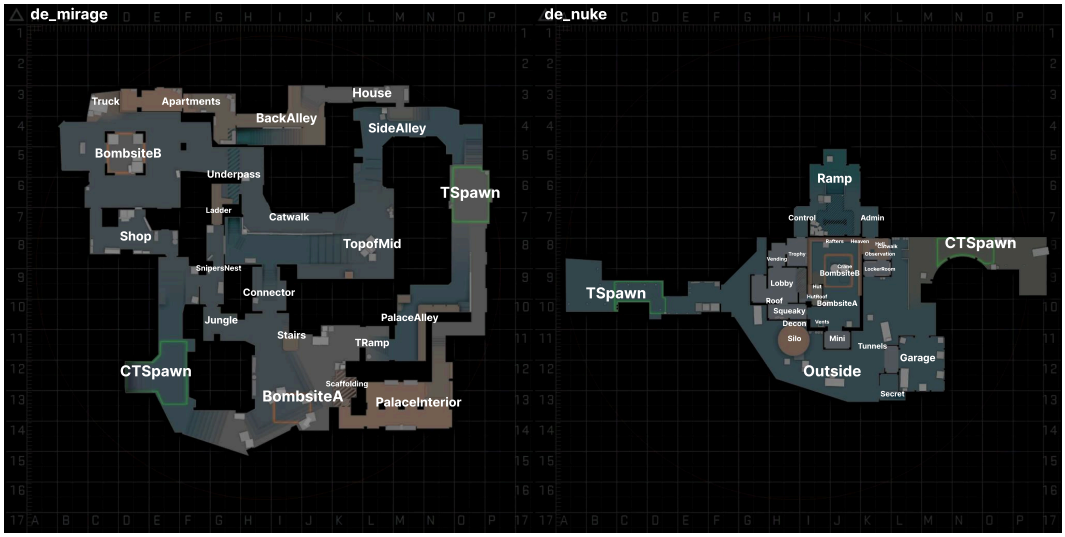


Figure 12: **Map callouts for Mirage and Nuke.** Continued from Figure 10.

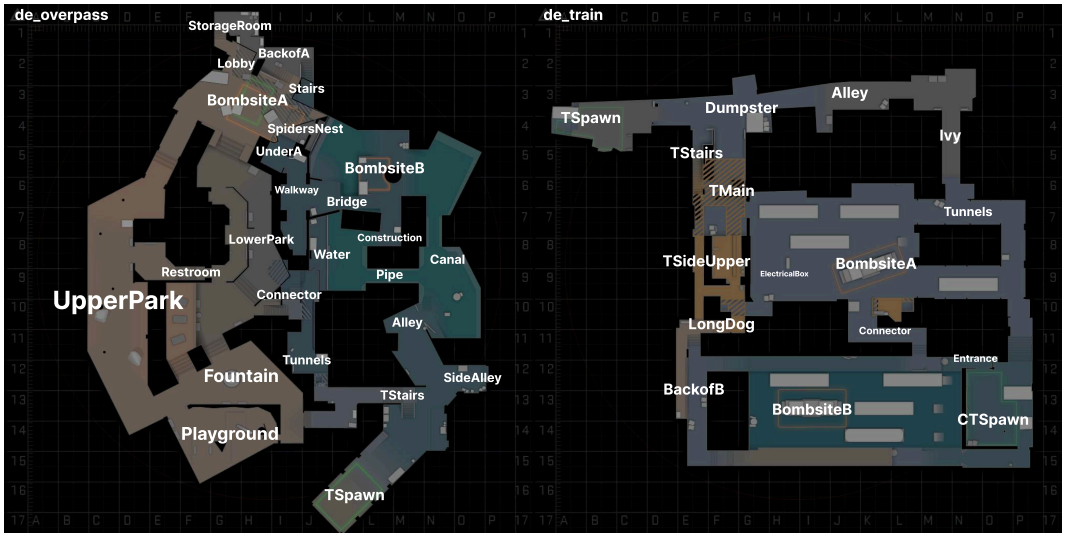


Figure 13: **Map callouts for Overpass and Train.** Continued from Figure 10.

C ENCODER SETUP

This section records the pretrained checkpoints and the aggregation, adaptation, and split protocol of each experimental setting, because these differ between the benchmark and the later diagnostics. Table 12 lists the checkpoint identifiers, which come from the saved configurations rather than from model-family names alone.

Table 12: **Pretrained backbone identifiers.** DINOv3 is used only in the later diagnostics.

Model	Checkpoint identifier
CLIP	`openai/clip-vit-base-patch32`
SigLIP2	`google/siglip2-base-patch16-224`
DINOv2	`facebook/dinov2-base`
V-JEPA2	`facebook/vjepa2-vitl-fpc16-256-ssv2`
DINOv3	`facebook/dinov3-vitb16-pretrain-lvd1689m`

C.1 PROTOCOL PROVENANCE

Table 13 summarizes the four settings and the evidence each provides. The paragraphs below give the aggregation, split, and label-eligibility details behind it.

Table 13: **Protocols and evidentiary roles.** B is the unadapted baseline, C is CECL, and S and H are budget-matched single-view and shuffled-pair adaptation. The first three settings use round splits and repeated linear probes; the controlled study uses match-held-out splits and converged probes.

Setting (role)	Aggregation (B → C)	Split evidence
Primary (CLIP, SigLIP2, DINOv2; Mirage)	Frame mean → frame mean; no added module	Historical split snapshot unavailable
Complementary (V-JEPA2; Mirage)	Random frozen pooler → trained pooler	Same historical split
Supporting (SigLIP2 on three maps; DINOv3 (Siméoni et al., 2025) and V-JEPA2 on Mirage)	Random frozen → trained temporal block (image) or pooler (video)	Retained manifests with zero shared CECL-train / probe-test or validation rounds
Controlled (SigLIP2; Mirage; B, C, S, H)	Frame mean in all arms; matched data, steps, and pairing budget	Match-held-out manifests (26/9/9 matches)

Benchmark aggregation. The original 33-task benchmark uses Mirage. Its saved run configurations and the contemporaneous encoder implementation specify frame averaging for CLIP, SigLIP2, and DINOv2, and no trainable temporal transformer is added to these image encoders. Both arms therefore share the aggregation. The baseline backbone is unadapted and frozen, whereas CECL fine-tunes the last three backbone blocks before freezing. The V-JEPA2 wrapper creates a separate attentive pooler at initialization. That pooler stays frozen with the baseline encoder but is trained during Stage 1 adaptation, so this comparison includes learning the pooler.

Diagnostic aggregation. The later three-map diagnostics use the temporal-encoder configuration, which adds one bidirectional block with eight heads to image encoders. Without a Stage 1 checkpoint, this block is randomly initialized and frozen with the encoder, and CECL loads its trained counterpart. These diagnostics thus carry an additional adaptation difference and should not be read as architecture- and budget-matched evidence for teammate pairing. Neither setting includes a single-view or shuffled-pair adaptation control. Those controls are supplied by the match-held-out study (Sections 5.3 and 6).

Split audit. Both data modules restrict the training loader to rows tagged `train`, and downstream labels inherit round partitions rather than splitting clips independently. Auditing all retained downstream task CSVs against the corresponding CECL training rows finds zero shared validation or test rounds on Dust2, Inferno, and Mirage, whose retained CECL train/validation/test round counts are 246/49/66, 185/29/47, and 682/142/163. This verifies the retained diagnostic manifests. It does

not reconstruct the original Mirage split snapshot, which is absent from its archived run records. Matches can span several partitions in either design, so no held-out-match claim is made for them.

Future-target eligibility. The location-label generator requires the observer to remain alive through the prediction tick and caps eligibility at round end. At 64 ticks/s with window start s, the five-second observation interval is $[s, s + 320]$. Nowcasts target its midpoint $s + 160$, so the input includes frames after the target time and extends 2.5 seconds beyond the nowcast label. Forecasts at horizon $h > 0$ instead target $s + 320 + 64h$. The contemporaneous and current generators share this midpoint/endpoint convention. All 43,724 retained Self, Teammate, and Enemy nowcast rows across three maps satisfy the 160-tick offset and 320-tick duration, and the supplied timing audit records counts by map and target. This check covers retained labels and does not recover the missing historical split snapshot.

Only players alive at the forecast tick contribute occupied regions. Self targets require one valid region, and other-player targets require at least one surviving target. Windows that extend beyond observer death or round end, as well as invalid region labels, are excluded rather than padded. These rules condition the forecasting population on survival and can change its composition with horizon. Verifying the exact excluded counts for the original benchmark requires the archived historical labels.

D TRAINING HYPERPARAMETERS

This section reports the saved Stage 1 and probe settings, how probe differences should be interpreted, and the compute used. Tables 14 and 15 give the settings attached to the original Mirage checkpoints rather than a later configuration applied retrospectively. In particular, the original optimizer is AdamW, not Muon, and the image encoders have no temporal block.

Main settings. Frames are resized to 306×306. Stage 1 uses AdamW (Loshchilov & Hutter, 2017) for 40 epochs with learning rate 10^{-4}, weight decay 0.01, and cosine decay. Probes use AdamW for 20 epochs at 4×10^{-4} with batch size 32 and validation-loss checkpoint selection, which measures readout under a fixed budget rather than optimal linear decodability. The controlled study therefore also fits probes to convergence with ridge or L-BFGS logistic regression and validation-selected regularization. Later temporal runs use Muon (Jordan et al., 2024).

The later diagnostics use the temporal block and the Muon/AdamW configuration recorded in their saved runs, with accumulated contrastive embeddings. Those settings belong to the later setting and cannot reproduce Table 2. Equal probe settings and validation-loss checkpoint selection do not demonstrate equal convergence across representations, and a broader probe optimization sweep is not available.

D.1 INTERPRETING PROBE DIFFERENCES

For target T, the reported quantity is $\Delta_T = S_T(\hat{q}_{T,C} \circ f_C) - S_T(\hat{q}_{T,B} \circ f_B)$, where S_T is the test metric, f is a frozen encoder, and $\hat{q}$ is the linear head returned by the specified training and validation-selection procedure. This is a finite-protocol performance difference. It does not estimate mutual information, the best possible nonlinear decoder, or even the best achievable linear readout. Conditioning, feature scaling, regularization, and optimization can change Δ_T without destroying target information, and an invertible nonlinear transformation can preserve information while changing linear accessibility.

Design implication. CECL aligns projected features $g_\phi(h_i)$, whereas probes read h_i. An illustrative decomposition $h_i = (c_i, p_i)$ could expose collective context c_i to correspondence while retaining private detail p_i for action. This is a design possibility, not an identified factorization. The matched suite suggests that the original regime approaches it, since other-player readout improves while self-region readout is preserved.

View-selection theory makes retained information relative to a downstream task (Tian et al., 2020b), and shared-factor identifiability depends on generative, objective, and representation-dimension assumptions (Daunhawer et al., 2023). These assumptions are not verified for our recordings, so their

Table 14: **Original Mirage CECL settings.** Recovered from saved run configurations.

Setting	Value
Backbone adaptation	Last three transformer blocks
Image aggregation	Mean of per-frame features
Projector	One hidden layer; 768-dimensional output
Loss	Sigmoid CECL; scale 10 and bias -3 initially
Encoder-training seed	42; one checkpoint per backbone
Epochs / precision	40 / bfloat16 mixed precision
Optimizer	AdamW; LR 10^{-4}; weight decay 0.01
Schedule	Cosine; 500 warmup steps; minimum LR ratio 0.1
Gradient clipping	1.0
Probe	Frozen encoder; one linear head; 20 epochs
Probe optimizer	AdamW; LR 4×10^{-4}; weight decay 0
Probe batch / selection	16 examples, accumulation 2; minimum validation loss

Table 15: **Original Stage 1 loader settings.** A loader item is a synchronized group, so batch size times accumulation steps is not a count of independent negative pairs.

Setting	CLIP	SigLIP2	DINOv2	V-JEPA2
Loader batch	16	10	4	4
Accumulation steps	4	4	10	10

conclusions cannot be imported as a theorem about CECL. In particular, the contrastive loss acts on $g_\phi(h)$ whereas probes read h, and aligning projected features does not by itself require eliminating observer-specific coordinates of h. The allocentric shift is therefore a measured operating point, not a proved Pareto frontier.

The match-held-out study (Sections 5.3 and 6) implements the first version of a discriminating design. Its CECL, single-view (S), and side- and phase-matched pseudo-team (H) arms share data exposure, trainable layers, projector, optimizer updates, and pair budgets, and all partitions are fixed by match before either stage. Similar gains under all arms would have favored a generic adaptation account. Instead, S stays near the baseline, H recovers about 70% of the gain, and CECL exceeds both on all eight location tasks, which supports a role for correspondence structure. Remaining steps are repeated encoder training, further backbones and maps, and a time-shuffled teammate arm.

For speed, standardize features with training statistics and solve ridge regression with an unpenalized intercept, selecting regularization by validation error and evaluating the test set once. Classification probes likewise need a validation-selected regularization and convergence sweep under the same preprocessing. Report paired match-clustered uncertainty across independently trained encoders. For location, use a common one-versus-rest region formulation on matched observer windows, comparing Self with a specified other player, and include AP, AUROC, validation-selected thresholds, predicted cardinality, and train-prior and always-positive baselines on the same test rows. Finally, varying adaptation strength and depth would reveal whether gains and costs move together or can improve jointly. Together, these checks distinguish probe optimization, task formulation, generic adaptation, and correspondence effects.

D.2 Compute resources

Table 16 lists the compute allocations of the reported experiments. The CECL pre-training sweep submits one job per model–map pair, and V-JEPA2 receives a longer allocation because of its larger video backbone. For SLURM jobs, the wall time is the scheduler allocation limit, an upper bound rather than the measured runtime of every run. The ViZDoom control runs outside SLURM on a single RTX 4090 workstation and is budgeted by environment frames rather than by a cluster wall-time allocation.

Table 16: **Compute allocations.** SLURM wall-time limits are allocation upper bounds.

Experiment	GPU	CPUs	Memory (GB)	Wall-time limit
CECL pre-training	1 NVIDIA A40/A100	32	128	48 h per job
Downstream probes	1 NVIDIA A40/A100	32	128	24 h per job; 48 h for V-JEPA2
ViZDoom control	1 NVIDIA RTX 4090	32	128	Environment-frame budget

E Losses and metrics

This section defines the probe losses and the evaluation metrics of each task formulation; the ViZDoom objective and metric are given in Appendix K.

Probe losses. Each linear probe minimizes the loss of its formulation over N examples with logits z and sigmoid $\sigma(\cdot)$. Binary tasks (e.g., kill, death, bomb state) use binary cross-entropy with labels $y_i \in \{0, 1\}$,

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N}\sum_{i=1}^{N}\big[y_i \log \sigma(z_i) + (1 - y_i)\log(1 - \sigma(z_i))\big].$$

Multiclass tasks (e.g., self location, alive count) with C classes and one-hot indicators y_{ic} use cross-entropy,

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic} \log \frac{e^{z_{ic}}}{\sum_{j=1}^{C} e^{z_{ij}}}.$$

Multi-label tasks (e.g., teammate and enemy location) apply binary cross-entropy to each of the $L = 23$ region labels, where $y_{ij} \in \{0, 1\}$ indicates whether label j is active for example i,

$$\mathcal{L}_{\text{ML}} = -\frac{1}{NL}\sum_{i=1}^{N}\sum_{j=1}^{L}\big[y_{ij} \log \sigma(z_{ij}) + (1 - y_{ij})\log(1 - \sigma(z_{ij}))\big].$$

Regression tasks (speed) use the mean squared error between target y_i and prediction $\hat{y}_i$,

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2.$$

Evaluation metrics. Table 17 lists the reported and primary metrics. Top-k accuracy uses class-macro averaging, so top-1 equals mean per-class recall and differs from sample-weighted accuracy under class imbalance. Exact match is the fraction of examples whose predicted label set equals the ground truth, and Hamming accuracy is $1 - \frac{1}{NL}\sum_{i,j}\mathbf{1}[\hat{y}_{ij} \neq y_{ij}]$. Hamming accuracy is dominated by absent labels for sparse region sets, so it is not the primary location metric. Metric families are never pooled to define the main effect-size table. A decline of 13.9 in that table means a decline of 0.139 in native R^2, not 13.9 units.

Table 17: **Metrics by task formulation.** Multi-label predictions are thresholded at 0.5, and macro-F1 averages over labels.

Formulation	Reported metrics	Primary metric
Binary	Accuracy, F1, AUROC	Sample accuracy
Multiclass	Top-1, top-3, macro-F1	Class-macro top-1 (mean recall)
Multi-label	Exact match, Hamming accuracy, macro-F1, AUROC	Macro-F1
Regression	MSE, MAE, R^2	R^2

F Detailed benchmark scores

This section lists every baseline and CECL score of the four-encoder Mirage benchmark and audits the location results. Tables 18–21 report all 132 task–encoder results as means at their recorded precision, without redundant difference columns; metrics are defined in Appendix E. The source

tables do not quantify independent CECL-training uncertainty, and the recovered location-probe repeats are reported separately in Appendix F.1.

For each target and formulation group, Table 2 averages the reported baseline and CECL means with equal weight, so its values are reconstructed from rounded source tables. We omit Student-t intervals across task–encoder results, because dependence among these results prevents reading such intervals as training or sampling uncertainty. The averages are descriptive and are never treated as independent trials.

Table 18: **Binary-task scores.** B, baseline; C, CECL; bold marks the better arm. Accuracy in percent; F1 and AUROC in native units.

Task	Encoder	Acc. (%)		F1		AUROC	
		B	C	B	C	B	C
Self							
death_5s	CLIP	58.9	**60.9**	0.610	**0.632**	0.628	**0.658**
death_5s	DINOv2	**62.7**	60.8	**0.655**	0.636	**0.665**	0.658
death_5s	SigLIP2	62.2	**62.4**	0.648	**0.657**	0.659	**0.678**
death_5s	V-JEPA2	63.4	**65.8**	0.645	**0.662**	0.687	**0.721**
death_10s	CLIP	55.7	**58.4**	0.541	**0.611**	0.597	**0.613**
death_10s	DINOv2	**57.9**	57.0	**0.608**	0.533	**0.614**	0.602
death_10s	SigLIP2	**59.7**	56.3	**0.603**	0.599	**0.626**	0.592
death_10s	V-JEPA2	**59.1**	58.8	**0.616**	0.603	**0.629**	0.623
death_20s	CLIP	52.3	**54.2**	0.590	**0.597**	0.531	**0.557**
death_20s	DINOv2	**57.4**	56.2	**0.635**	0.593	**0.584**	0.579
death_20s	SigLIP2	**54.9**	54.0	**0.629**	0.602	0.553	**0.571**
death_20s	V-JEPA2	54.6	**56.1**	**0.636**	0.609	0.566	**0.583**
inCombat	CLIP	61.8	**64.5**	0.591	**0.649**	0.690	**0.710**
inCombat	DINOv2	71.1	**72.4**	0.708	**0.737**	0.795	**0.798**
inCombat	SigLIP2	66.9	**67.4**	0.662	**0.688**	0.741	**0.749**
inCombat	V-JEPA2	73.1	**76.4**	0.733	**0.774**	0.820	**0.843**
kill_5s	CLIP	**59.5**	57.9	**0.626**	0.583	**0.640**	0.625
kill_5s	DINOv2	**63.5**	62.7	**0.656**	0.641	0.670	**0.688**
kill_5s	SigLIP2	**62.0**	60.7	**0.648**	0.638	**0.657**	0.653
kill_5s	V-JEPA2	**63.3**	62.5	**0.661**	0.655	**0.685**	0.678
kill_10s	CLIP	55.7	**57.5**	**0.583**	0.581	0.584	**0.598**
kill_10s	DINOv2	59.1	**61.0**	0.605	**0.623**	0.620	**0.632**
kill_10s	SigLIP2	59.4	**59.7**	0.607	**0.610**	0.609	**0.633**
kill_10s	V-JEPA2	58.9	**61.3**	0.615	**0.632**	0.623	**0.645**
kill_20s	CLIP	55.0	**58.5**	0.489	**0.549**	0.562	**0.613**
kill_20s	DINOv2	53.7	**58.4**	0.518	**0.606**	0.567	**0.614**
kill_20s	SigLIP2	54.2	**60.4**	0.517	**0.592**	0.551	**0.637**
kill_20s	V-JEPA2	53.6	**55.9**	0.506	**0.555**	0.548	**0.582**
Teammate							
inCombat	CLIP	63.0	**64.9**	0.628	**0.670**	0.679	**0.710**
inCombat	DINOv2	69.9	**71.3**	0.709	**0.717**	0.767	**0.791**
inCombat	SigLIP2	64.8	**66.8**	0.651	**0.683**	0.735	**0.752**
inCombat	V-JEPA2	72.6	**76.3**	0.728	**0.770**	0.798	**0.840**
Global							
anyKill_5s	CLIP	56.7	**57.3**	**0.599**	0.598	0.582	**0.594**
anyKill_5s	DINOv2	56.2	**57.6**	**0.617**	0.597	0.586	**0.614**
anyKill_5s	SigLIP2	56.6	**57.3**	0.597	**0.615**	0.584	**0.598**
anyKill_5s	V-JEPA2	59.3	**61.0**	0.630	**0.645**	0.640	**0.650**
anyKill_10s	CLIP	54.5	**60.2**	0.584	**0.657**	0.568	**0.613**
anyKill_10s	DINOv2	56.0	**58.4**	0.595	**0.638**	0.588	**0.607**
anyKill_10s	SigLIP2	57.5	**60.2**	0.625	**0.652**	0.598	**0.612**
anyKill_10s	V-JEPA2	59.4	**61.2**	0.631	**0.639**	0.622	**0.652**
anyKill_20s	CLIP	**55.4**	54.1	**0.658**	0.600	**0.564**	0.550
anyKill_20s	DINOv2	**56.9**	56.0	**0.646**	0.614	0.577	**0.586**
anyKill_20s	SigLIP2	**57.0**	54.6	**0.652**	0.611	**0.583**	0.566
anyKill_20s	V-JEPA2	56.4	**59.3**	**0.663**	0.633	0.568	**0.629**
bombPlanted	CLIP	66.5	**74.5**	0.641	**0.732**	0.747	**0.845**

Table 18 continued

Task	Encoder	Acc. (%)		F1		AUROC	
		B	C	B	C	B	C
bombPlanted	DINOv2	70.1	**82.4**	0.656	**0.821**	0.823	**0.908**
bombPlanted	SigLIP2	66.7	**78.2**	0.644	**0.776**	0.762	**0.876**
bombPlanted	V-JEPA2	65.0	**83.2**	0.613	**0.829**	0.730	**0.914**
bombSite	CLIP	**66.0**	64.7	**0.701**	0.658	**0.717**	0.701
bombSite	DINOv2	66.0	**66.2**	**0.690**	0.680	0.738	**0.743**
bombSite	SigLIP2	67.6	**68.3**	**0.706**	0.695	0.740	**0.743**
bombSite	V-JEPA2	62.8	**64.3**	0.652	**0.669**	0.682	**0.713**
postPlantOutcome	CLIP	56.2	**60.1**	0.350	**0.423**	0.572	**0.657**
postPlantOutcome	DINOv2	58.7	**63.5**	0.386	**0.530**	0.621	**0.696**
postPlantOutcome	SigLIP2	56.2	**62.2**	0.329	**0.500**	0.583	**0.671**
postPlantOutcome	V-JEPA2	58.1	**65.5**	0.338	**0.545**	0.594	**0.717**
roundWinner	CLIP	45.5	**56.1**	0.013	**0.518**	0.495	**0.599**
roundWinner	DINOv2	48.5	**59.6**	0.264	**0.600**	0.550	**0.641**
roundWinner	SigLIP2	47.2	**58.5**	0.217	**0.565**	0.535	**0.641**
roundWinner	V-JEPA2	47.0	**58.1**	0.073	**0.513**	0.508	**0.646**
willPlant	CLIP	45.9	**51.1**	0.072	**0.465**	**0.526**	0.524
willPlant	DINOv2	**49.0**	47.2	0.290	**0.427**	**0.528**	0.499
willPlant	SigLIP2	50.8	**52.6**	0.328	**0.489**	0.548	**0.554**
willPlant	V-JEPA2	45.3	**52.9**	0.065	**0.485**	0.504	**0.553**

Table 19: **Multiclass-task scores.** Notation as in Table 18. Class-macro top-k accuracy in percent; macro-F1 in native units.

Task	Encoder	Top-1 (%)		Top-3 (%)		Macro-F1	
		B	C	B	C	B	C
Self							
location_0s	CLIP	**62.7**	51.6	**82.7**	76.6	**0.614**	0.503
location_0s	DINOv2	**79.4**	72.8	**97.4**	92.2	**0.791**	0.724
location_0s	SigLIP2	**74.3**	67.1	**93.8**	87.9	**0.739**	0.667
location_0s	V-JEPA2	56.9	**65.8**	79.1	**88.9**	0.558	**0.655**
location_5s	CLIP	**40.9**	30.7	**68.9**	58.3	**0.391**	0.287
location_5s	DINOv2	**51.1**	47.1	**83.0**	77.5	**0.505**	0.464
location_5s	SigLIP2	**48.1**	42.6	**78.9**	73.6	**0.473**	0.417
location_5s	V-JEPA2	40.6	**45.0**	67.8	**74.2**	0.392	**0.442**
location_10s	CLIP	**27.6**	20.9	**52.6**	45.7	**0.237**	0.189
location_10s	DINOv2	**35.0**	30.2	**65.6**	59.5	**0.335**	0.288
location_10s	SigLIP2	**33.3**	30.3	**62.4**	57.7	**0.308**	0.291
location_10s	V-JEPA2	29.8	29.8	55.9	**58.9**	0.269	**0.291**
location_20s	CLIP	**16.6**	12.7	**34.3**	32.0	**0.119**	0.110
location_20s	DINOv2	**18.8**	16.5	**40.8**	36.9	**0.157**	0.144
location_20s	SigLIP2	17.9	**19.6**	**39.2**	38.1	0.144	**0.179**
location_20s	V-JEPA2	**17.1**	15.7	35.3	**36.4**	0.139	**0.140**
Teammate							
aliveCount	CLIP	27.8	**35.6**	68.6	**74.0**	0.264	**0.344**
aliveCount	DINOv2	32.4	**37.6**	72.4	**80.6**	0.310	**0.368**
aliveCount	SigLIP2	31.1	**34.9**	73.4	**74.9**	0.299	**0.339**
aliveCount	V-JEPA2	30.0	**38.7**	69.8	**80.0**	0.285	**0.380**
Enemy							
aliveCount	CLIP	28.1	**34.2**	68.3	**74.2**	0.255	**0.331**
aliveCount	DINOv2	31.0	**39.0**	73.9	**81.3**	0.299	**0.386**
aliveCount	SigLIP2	31.4	**36.3**	71.7	**77.1**	0.301	**0.357**
aliveCount	V-JEPA2	28.9	**34.9**	70.1	**79.3**	0.277	**0.343**
Global							
roundOutcome	CLIP	21.9	**23.1**	60.9	**64.1**	0.188	**0.202**
roundOutcome	DINOv2	22.8	**24.1**	62.2	**67.2**	0.188	**0.216**
roundOutcome	SigLIP2	23.6	**24.8**	62.5	**67.9**	0.200	**0.224**

Table 19 continued

Task	Encoder	Top-1 (%)		Top-3 (%)		Macro-F1	
		B	C	B	C	B	C
roundOutcome	V-JEPA2	20.7	**24.4**	61.8	**65.0**	0.173	**0.219**

Table 20: **Multi-label location scores.** Notation as in Table 18. Exact-match and Hamming accuracy in percent; macro-F1 in native units.

Task	Encoder	Exact (%)		Hamming (%)		Macro-F1	
		B	C	B	C	B	C
Teammate							
location_0s	CLIP	0.0	**5.6**	87.1	**87.7**	0.004	**0.237**
location_0s	DINOv2	3.8	**5.6**	87.7	**88.0**	0.173	**0.249**
location_0s	SigLIP2	2.5	**6.0**	87.5	**88.0**	0.096	**0.248**
location_0s	V-JEPA2	1.2	**7.1**	87.3	**88.2**	0.038	**0.290**
location_5s	CLIP	0.0	**2.5**	87.0	**87.4**	0.003	**0.178**
location_5s	DINOv2	0.9	**2.2**	87.3	**87.5**	0.112	**0.189**
location_5s	SigLIP2	0.5	**2.1**	87.2	**87.4**	0.066	**0.167**
location_5s	V-JEPA2	0.0	**3.1**	87.1	**87.7**	0.016	**0.199**
location_10s	CLIP	0.1	**0.4**	86.9	86.9	0.002	**0.083**
location_10s	DINOv2	0.4	**1.1**	86.9	**87.0**	0.057	**0.102**
location_10s	SigLIP2	0.3	**0.7**	87.0	87.0	0.030	**0.084**
location_10s	V-JEPA2	0.0	**0.7**	87.0	87.0	0.005	**0.119**
location_20s	CLIP	0.0	0.0	**87.8**	87.7	0.000	**0.027**
location_20s	DINOv2	0.1	0.1	**87.7**	87.6	0.021	**0.031**
location_20s	SigLIP2	0.0	0.0	**87.8**	87.7	0.005	**0.035**
location_20s	V-JEPA2	0.0	**0.2**	**87.8**	87.5	0.000	**0.056**
Enemy							
location_0s	CLIP	0.0	**5.1**	85.1	**85.7**	0.023	**0.238**
location_0s	DINOv2	4.1	**5.6**	85.4	**85.7**	0.181	**0.252**
location_0s	SigLIP2	1.6	**5.7**	85.3	**85.7**	0.092	**0.233**
location_0s	V-JEPA2	0.5	**6.3**	85.0	**86.3**	0.043	**0.278**
location_5s	CLIP	0.0	**1.3**	85.3	**85.9**	0.020	**0.181**
location_5s	DINOv2	1.1	**1.6**	85.9	**86.2**	0.154	**0.200**
location_5s	SigLIP2	0.2	**2.2**	85.8	**86.1**	0.094	**0.199**
location_5s	V-JEPA2	0.0	**1.8**	85.6	**86.2**	0.050	**0.216**
location_10s	CLIP	0.0	**1.0**	85.1	**85.4**	0.010	**0.118**
location_10s	DINOv2	0.1	**0.4**	85.2	**85.5**	0.081	**0.120**
location_10s	SigLIP2	0.1	**0.2**	85.3	**85.5**	0.058	**0.111**
location_10s	V-JEPA2	0.1	**0.6**	85.2	**85.5**	0.016	**0.126**
location_20s	CLIP	0.1	**0.5**	**86.1**	85.8	0.003	**0.044**
location_20s	DINOv2	0.3	0.3	**85.9**	85.8	0.035	**0.050**
location_20s	SigLIP2	0.2	**0.5**	86.0	86.0	0.021	**0.054**
location_20s	V-JEPA2	0.0	**0.6**	**86.1**	85.9	0.007	**0.066**

Table 21: **Regression-task scores.** Notation as in Table 18. Native units; lower MSE and MAE and higher R^2 are better.

Task	Encoder	MSE		MAE		R^2	
		B	C	B	C	B	C
Self							
speed	CLIP	**0.025**	0.031	**0.126**	0.139	**0.466**	0.344
speed	DINOv2	**0.027**	0.030	**0.127**	0.137	**0.420**	0.359
speed	SigLIP2	**0.021**	0.028	**0.114**	0.131	**0.541**	0.402
speed	V-JEPA2	**0.016**	0.027	**0.096**	0.130	**0.658**	0.424

Table 22: **Location macro-AUROC and probe variability.** Means ± sample SD across probe seeds, with n the baseline/CECL repeat counts and Δ the CECL-minus-baseline difference. Encoder checkpoints and data are fixed, so the SDs are not independent representation-training uncertainties.

		Teammate location				Enemy location			
h (s)	Encoder	n	Baseline	CECL	Δ	n	Baseline	CECL	Δ
0	CLIP	4/4	0.690 ±0.001	**0.773** ±0.002	+0.083	4/4	0.668 ±0.001	**0.757** ±0.002	+0.089
	DINOv2	4/4	0.753 ±0.002	**0.781** ±0.004	+0.028	3/3	0.729 ±0.003	**0.781** ±0.001	+0.052
	SigLIP2	4/4	0.742 ±0.000	**0.780** ±0.003	+0.038	4/4	0.716 ±0.001	**0.777** ±0.001	+0.061
	V-JEPA2	4/4	0.692 ±0.001	**0.798** ±0.001	+0.106	4/4	0.675 ±0.001	**0.796** ±0.001	+0.120
5	CLIP	4/4	0.689 ±0.001	**0.738** ±0.001	+0.048	4/4	0.682 ±0.001	**0.746** ±0.003	+0.064
	DINOv2	4/4	0.722 ±0.001	**0.759** ±0.006	+0.037	3/3	0.727 ±0.002	**0.768** ±0.003	+0.041
	SigLIP2	4/4	0.720 ±0.001	**0.748** ±0.001	+0.028	4/4	0.722 ±0.001	**0.756** ±0.004	+0.035
	V-JEPA2	4/4	0.687 ±0.001	**0.764** ±0.001	+0.077	4/4	0.687 ±0.002	**0.773** ±0.003	+0.086
10	CLIP	4/4	0.656 ±0.001	**0.695** ±0.004	+0.038	4/4	0.650 ±0.001	**0.711** ±0.003	+0.061
	DINOv2	4/4	0.692 ±0.004	**0.715** ±0.004	+0.023	3/3	0.691 ±0.001	**0.730** ±0.002	+0.039
	SigLIP2	4/4	0.692 ±0.001	**0.706** ±0.004	+0.015	4/4	0.686 ±0.001	**0.722** ±0.003	+0.036
	V-JEPA2	4/4	0.648 ±0.001	**0.710** ±0.006	+0.062	4/4	0.646 ±0.002	**0.733** ±0.003	+0.087
20	CLIP	4/4	0.626 ±0.000	**0.644** ±0.002	+0.019	4/4	0.613 ±0.001	**0.648** ±0.003	+0.035
	DINOv2	3/3	0.648 ±0.003	**0.654** ±0.003	+0.006	3/3	0.635 ±0.003	**0.667** ±0.002	+0.032
	SigLIP2	4/4	0.640 ±0.002	**0.662** ±0.002	+0.023	4/4	0.641 ±0.001	**0.660** ±0.006	+0.019
	V-JEPA2	4/4	0.618 ±0.004	**0.656** ±0.002	+0.038	4/4	0.608 ±0.003	**0.663** ±0.003	+0.055

Table 21 continued

Task	Encoder	MSE		MAE		R^2	
		B	C	B	C	B	C
Teammate							
speed	CLIP	0.044	0.044	0.182	**0.169**	0.161	**0.162**
speed	DINOv2	0.050	**0.048**	0.183	**0.175**	0.051	**0.087**
speed	SigLIP2	**0.045**	0.046	0.177	**0.173**	**0.153**	0.121
speed	V-JEPA2	**0.042**	0.047	**0.173**	0.174	**0.202**	0.103

F.1 Location ranking audit

We locate the original validation-selected probe records by model, target, horizon, UI mask, checkpoint arm, and run configuration. The 246 records cover all 32 location comparisons in both arms, and re-averaging their F1 values reproduces every corresponding source-table entry at the recorded precision; no model is re-evaluated and no checkpoint is re-selected. Each comparison and arm has three or four probe seeds from the recorded set {1, 2, 3, 42}, and the complete records and counts are supplied as CSV. Table 22 reports the resulting macro-AUROC. Across arms there are 123 same-seed pairs. All paired AUROC differences are positive, and one paired F1 difference is negative. Because encoders and test samples are shared, these counts are consistency diagnostics rather than binomial significance tests.

Macro-AUROC is threshold independent but does not replace macro average precision on sparse labels. The saved aggregate logs contain neither full score vectors nor validation predictions, so they cannot support reconstructing AP, optimized thresholds, predicted label cardinality, label-prior scores, or match-clustered intervals. The supplied run table describes probe randomness around a fixed representation. Only independent encoder training and a held-out-match protocol would address the other sources of uncertainty.

F.2 Label-prior bounds

This check bounds the macro-F1 of a trivial always-positive predictor from label prevalence alone. At prevalence p, that predictor has $F_1 = f(p) = 2p/(1+p)$. For $L = 23$ labels with $a \le p_\ell \le b$, let $\mu = L^{-1}\sum_\ell p_\ell$, which equals mean positive cardinality divided by L. Since $f''(p) = -4/(1+$

$p)^3 < 0$, concavity gives the chord and Jensen bounds

$$\frac{b-\mu}{b-a}f(a) + \frac{\mu-a}{b-a}f(b) \;\le\; \frac{1}{L}\sum_{\ell=1}^{L} f(p_\ell) \;\le\; f(\mu). \tag{2}$$

The bounds do not require label independence. The lower bound need not be attainable with exactly 23 labels; it is conservative rather than an estimate of the score.

Table 23 applies Equation (2) to the *pre-split pools* of Table 10. We widen prevalence endpoints and mean cardinality by half their printed units, use the smallest (largest) allowed mean for the lower (upper) bound, and round outward. The supplied script derives all eight rows from these summaries.

Table 23: **Always-positive macro-F1 bounds.** Computed from pre-split label summaries with rounding uncertainty; these are not measured test baselines.

	Teammate			Enemy		
h (s)	Labels/ex.	Lower	Upper	Labels/ex.	Lower	Upper
0	3.0	0.199	0.235	3.49	0.229	0.264
5	3.01	0.207	0.232	3.48	0.231	0.264
10	2.99	0.204	0.231	3.45	0.229	0.262
20	2.80	0.195	0.218	3.15	0.215	0.242

The 20-second teammate interval, 0.195–0.218, cannot be ranked against CECL's 0.037 test mean because the populations differ, and a valid baseline comparison needs the same test rows. Missing test labels and prediction vectors prevent reconstructing held-out prior scores, AP, and validation-selected thresholds. Constant positive outputs have no ranking resolution, so this check does not explain away the AUROC gains.

G Restricted evaluation

This section documents the restricted evaluations of Section 5.4, which test whether gains survive once the observation window no longer reveals the target directly. The diagnostic set has five map–encoder pairs, namely SigLIP2 on Dust2, Inferno, and Mirage, and DINOv3 and V-JEPA2 on Mirage. DINOv3 is an additional diagnostic backbone, distinct from the DINOv2 encoder of the 33-task benchmark. The saved diagnostic report covers 120 prediction dumps (six tasks, five pairs, two arms, and two runs per arm) and averages runs within an arm before taking the CECL-minus-baseline difference. These counts must not be replaced by the four-encoder or three-seed counts of other experiments.

Subset definitions. For bomb state, we keep windows in which the plant happened before the window or had not happened by the target time, excluding every window that contains the plant. For alive counts, we require an earlier death and no death inside the window, and HUD masking already removes the timer and kill feed. For forecasts, we keep examples whose occupied regions at the target time are disjoint from those at the end of the window, which rules out persistence.

Recomputing full-split metrics from these predictions matches the saved per-run metrics with a median absolute difference of 0.00020 and a maximum of 0.00328. Against earlier best-checkpoint records the maximum difference is 0.00443, because the restricted dumps use the last checkpoints. Small numerical differences and checkpoint selection are therefore separated explicitly from comparisons across experiment sets.

Bomb-state correction. Restricting to plants before the observation window leaves only positive labels, so its +0.0911 difference is positive-class recall, not accuracy on a two-class task. A broader filter that combines pre-window plants with unplanted examples retains 2,544 unique rows (1,238 positive and 1,306 negative), with mean difference +0.0562. Twenty-three of these rows still contain an in-window plant because event and target ticks differ. Excluding those rows yields the strict event-free result in the main paper, +0.0570, positive for five of five pairs, with median $n = 748$ and pair differences from +0.0247 to +0.1083. Its near-equality to the full-split gain is a survival check, not evidence that the gain increases when all visual traces disappear.

Alive-count metric. Requiring an earlier death removes the no-death class, and macro recall over the original fixed class set answers a different question from macro recall over present classes. The saved fixed-class differences are +0.0260 for enemies and +0.0144 for teammates. The present-class differences are +0.0313 and +0.0180, with ranges [+0.0075, +0.0427] and [−0.0196, +0.0606]. The main table uses present-class recall and states that convention; it should not be read as an exactly class-matched causal comparison to the full split.

Forecast restriction. For occupied region sets $S(t)$, we retain rows with $S(t_{\text{end}}) \cap S(t_{\text{target}}) = \varnothing$. This excludes unchanged-region copying but leaves all other predictable structure. The retained samples are small, so uncertainty and heterogeneity across pairs matter. Table 24 reports the underlying forecasts; Mirage examples repeated across encoders are not additional independent clips.

Table 24: **Disjoint-region forecast subsets.** Native macro-F1 of each arm, CECL-minus-baseline difference Δ, and retained test size n; bold marks the better arm.

Map	Encoder	Baseline	CECL	Δ	n
Enemy location, 5 s					
Dust2	SigLIP2	0.0408	**0.0703**	+0.0295	75
Inferno	SigLIP2	0.1644	**0.1782**	+0.0138	90
Mirage	DINOv3	0.1655	**0.1933**	+0.0278	45
Mirage	SigLIP2	0.1193	**0.2072**	+0.0879	45
Mirage	V-JEPA2	0.0813	**0.1671**	+0.0858	45
Enemy location, 10 s					
Dust2	SigLIP2	0.0081	**0.0491**	+0.0410	168
Inferno	SigLIP2	0.1312	**0.1529**	+0.0217	160
Mirage	DINOv3	**0.0698**	0.0678	−0.0020	117
Mirage	SigLIP2	0.0359	**0.1111**	+0.0752	117
Mirage	V-JEPA2	0.0158	**0.0926**	+0.0768	117
Teammate location, 10 s					
Dust2	SigLIP2	0.0118	**0.0278**	+0.0160	183
Inferno	SigLIP2	0.0915	**0.1132**	+0.0216	189
Mirage	DINOv3	**0.0681**	0.0393	−0.0288	146
Mirage	SigLIP2	0.0168	**0.0560**	+0.0392	146
Mirage	V-JEPA2	0.0000	**0.0672**	+0.0672	146

H TEAMMATE SIMILARITY

This section details the similarity analysis of Section 5.5. Probes show what a representation makes linearly readable, so here we ask directly, in the frozen feature space, what CECL makes synchronized teammates share.

Setup. We use the SigLIP2 Mirage checkpoint of the supporting setting (last three blocks, one temporal block, round split) and its unadapted counterpart, evaluated on the 163 test rounds of 44 matches that CECL never trained on (16,314 five-second clips). For each clip, we compute the cosine similarity to each synchronized teammate and to distractor clips of the same side from other rounds in the same 10-second round phase, the pseudo-pairs used by arm H. The score is the AUROC of teammate over distractor similarity, where 0.5 is chance. Intervals are 95% bootstrap intervals over matches (1,000 resamples). Grouped ridge probes on the same features reproduce the gain–cost direction, with teammate-occupancy AUROC rising from 0.741 to 0.763 and self-speed R^2 falling from 0.765 to 0.749.

Distance and timing. Without adaptation, teammate similarity is explained by co-location. Nearby teammates (under 600 units) score 0.783, whereas teammates more than 1,500 units apart score 0.409, below chance, so distant teammates look less alike than unrelated clips from the same phase. After CECL, distant teammates score 0.698, a gain of 0.289 [0.268, 0.307], and the gain grows with distance (Figure 6a). The similarity is specific to the moment. For distant pairs, it falls from 0.705 for synchronized windows to below chance at offsets beyond about 27 seconds, while the unadapted encoder stays near 0.42 at every offset (Figure 6b). The same pattern holds for frame-averaged backbone features, so it does not come from the temporal block alone. Because distractors

share side and coarse phase, and teammates this far apart are unlikely to share a view, co-visible objects, side, and coarse phase are unlikely explanations. Finer-grained phase cues are not excluded, and this diagnostic compares only the baseline and CECL; evaluating the S and H arms in the same space remains to be done.

Neighborhood structure. Local neighborhoods change differently. Among the ten nearest neighbors from other rounds, agreement on the observer's own region falls from 0.67 to 0.46 and agreement on side rises from 0.84 to 0.94, while teammate-occupancy agreement barely moves (0.245 to 0.250). The linear gains therefore reflect a direction in the space rather than a reorganization of local neighborhoods. A two-dimensional projection would mainly show side, so we do not report one. Scripts and source numbers are in `scripts/teammate_similarity`.

I WITHIN-MAP TASK EFFECTS

This section lists the per-task SigLIP2 effects behind the three-map comparison of Section 5.1. We recover the 28 tasks common to both SigLIP2 arms on Dust2, Inferno, and Mirage from saved validation-loss-selected probe checkpoints. Each task–map result averages seeds 1 and 2 per arm, giving 84 results from 336 runs. These results use the supporting temporal-encoder configuration of Table 13, with one fixed CECL encoder checkpoint per map, so they supply neither independent encoder repeats nor a control for the added aggregation module.

The recovered task scores reproduce all twelve means and improved-task counts of the earlier mixed-metric map summary at its printed precision. Table 2 replaces that summary with eleven target–metric groups, and Table 25 lists every task difference. An accompanying per-task file contains absolute arm means, and `map_probe_runs.csv` contains anonymized run-level scores. Keeping metrics separate matters. For example, teammate location improves on Inferno while teammate speed deteriorates, a distinction obscured by averaging their F1 and R^2 differences together.

J SEMANTIC PROBE SENSITIVITY

This section details the concept-ranking probe of Section 5.6 and its sensitivity to prompt templates. The vocabulary contains five groups of 50 concepts. For each of 100 sampled test clips per map, cosine similarity between visual and frozen text embeddings ranks the 250 concepts, and group changes summarize baseline-to-final mean rank differences. The template bank contains 43 singleton templates in seven families, namely bare/minimal (3), generic visual (8), video/temporal (8), gameplay (8), CS2-specific (6), caption/declarative (6), and instruction/question (4). Examples include a bare concept, a screenshot of the concept, footage of the concept, and a CS2 round with the concept. Four mean-pooled, renormalized ensembles are additional sensitivity checks and are excluded from the 129-pair singleton summary.

Baseline-rank control. Within each map–template pair, we fit per-concept change as a quadratic function of baseline rank and subtract the fitted value. A group's mean residual then compares its change with the change expected at its initial rank. The permutation check preserves group membership counts within five equal-count baseline-rank strata, and 2,000 within-stratum resamples yield a two-sided empirical p value with a plus-one correction. Counts of p values below 0.05 are descriptive robustness summaries across overlapping probes, without a multiple-testing correction or an independence claim.

Outcome. After conditioning on initial rank, Self declines under every template–map combination, and Teammate rises in 121 of 129. Global and Enemy pass the stratified check in fewer than half the pairs, and Spatial is inconsistent. These outcomes replace the stronger interpretation drawn from a single prompt, and no single-template spatial rise is used as a headline claim. The rank statistic measures relative language alignment, not calibrated recognition accuracy, attention, or beliefs. The dense 250-concept trajectory plot is supplied as a separate vector supplement for zoomed inspection of individual concepts.

Table 25: **All 28 within-map SigLIP2 task effects.** CECL minus baseline in native units, averaged over probe seeds 1 and 2. Accuracy (Acc.), class-macro recall (Recall), macro-F1 (F1), and R^2 are kept separate. Encoders use the later temporal configuration. Absolute scores and all 336 run metrics are supplied in the accompanying CSVs.

Task	Metric	Dust2	Inferno	Mirage
Self				
death_5s	Acc.	+0.001	+0.008	−0.013
death_10s	Acc.	+0.009	−0.003	+0.012
inCombat	Acc.	+0.026	−0.018	+0.017
kill_0s	Acc.	+0.022	−0.012	+0.038
kill_5s	Acc.	+0.014	−0.015	+0.013
kill_10s	Acc.	+0.015	+0.001	+0.044
location_0s	Recall	+0.054	−0.012	+0.015
location_5s	Recall	+0.027	−0.002	−0.021
location_10s	Recall	−0.006	−0.011	−0.031
speed	R^2	−0.169	−0.262	−0.172
Teammate				
aliveCount	Recall	+0.040	+0.012	+0.034
inCombat	Acc.	+0.044	−0.010	+0.058
location_0s	F1	+0.088	+0.067	+0.158
location_5s	F1	+0.048	+0.025	+0.093
location_10s	F1	+0.016	+0.020	+0.058
speed	R^2	−0.219	−0.224	−0.247
Enemy				
aliveCount	Recall	+0.037	+0.021	+0.038
location_0s	F1	+0.098	+0.096	+0.157
location_5s	F1	+0.056	+0.027	+0.093
location_10s	F1	+0.036	+0.022	+0.064
Global				
anyKill_0s	Acc.	+0.007	+0.002	+0.031
anyKill_5s	Acc.	−0.005	+0.011	−0.003
anyKill_10s	Acc.	+0.037	−0.019	−0.017
bombPlanted	Acc.	+0.030	+0.033	+0.107
bombSite	Acc.	−0.032	+0.068	−0.018
roundOutcome	Recall	+0.001	+0.026	+0.015
roundWinner	Acc.	−0.059	+0.003	+0.087
willPlant	Acc.	−0.077	−0.020	+0.054

J.1 TEMPLATE BANK

Table 26 lists the 43 singleton templates behind the 129 map–template pairs; ensemble variants are not included in that count. The placeholder `{concept}` is replaced by each vocabulary phrase of Appendix J.2.

Table 26: **Singleton template bank.** Templates are grouped into seven families, with family sizes in parentheses.

Key	Template
Bare/minimal (3)	
bare	{concept}
bare_period	{concept}.
bare_article	a {concept}
Generic visual (8)	
photo	a photo of {concept}
photo_article	a photo of a {concept}
image	an image of {concept}
picture	a picture of {concept}
screenshot	a screenshot of {concept}
photo_cropped	a cropped photo of {concept}
photo_lowres	a low resolution photo of {concept}

Table 26 continued

Key	Template
photo_bright	a bright photo of {concept}
Video/temporal (8)	
video	a video of {concept}
video_showing	a video showing {concept}
this_video_shows	this video shows {concept}
paper	this video shows "{concept}"
video_frame	a video frame of {concept}
clip	a clip of {concept}
footage	footage of {concept}
recording	a recording of {concept}
Gameplay (8)	
gameplay	gameplay showing {concept}
gameplay_screenshot	a gameplay screenshot of {concept}
videogame_screenshot	a video game screenshot of {concept}
ingame_footage	in-game footage of {concept}
fp_game_view	a first person game view of {concept}
game_scene	a game scene with {concept}
esports_footage	esports footage of {concept}
fps_screenshot	a first person shooter screenshot showing {concept}
CS2-specific (6)	
cs_screenshot	a Counter-Strike gameplay screenshot showing {concept}
cs_fp_scene	a first person Counter-Strike scene with {concept}
cs_footage	gameplay footage of {concept}
cs2_match	a Counter-Strike 2 match showing {concept}
cs2_round	a CS2 round with {concept}
cs_pro_match	a professional Counter-Strike match showing {concept}
Caption/declarative (6)	
there_is	there is {concept}
scene_contains	the scene contains {concept}
this_is	this is {concept}
example_of	an example of {concept}
is_visible	{concept} is visible
showing	showing {concept}
Instruction/question (4)	
what_happening	what is happening: {concept}
describe_scene	describe the scene: {concept}
caption	caption: {concept}
scene_description	scene description: {concept}

J.2 Concept vocabulary

The five groups contain 50 phrases each. They are fixed textual probes, not labels of what is visible in a clip, and some Self phrases refer to HUD elements removed from the images. Vocabulary selection remains a limitation even when prompt templates are swept.

Self (50). first person view; first person shooter perspective; my point of view; what I see in front of me; my hands holding a weapon; my gun in frame; holding a rifle; holding a pistol; reloading my weapon; switching weapons; my crosshair on screen; aiming down sights; looking through scope; my health bar; my ammo count; I am running; I am walking slowly; I am sprinting; I am crouching; I am jumping; I am standing still; I am moving forward; I am moving backward; moving at high speed; moving cautiously; stopped completely; quick movement; I am shooting; I am about to fire; I am taking aim; I am throwing a grenade; I am planting the bomb; I am checking angles; I am clearing a room; I am watching a doorway; I am camping a position; I am throwing a flashbang; I am throwing smoke; I am throwing molotov; I am using utility; I am in danger; I am safe; I am exposed; I am hidden; I am vulnerable; I am protected; I am surrounded; I am in a good position; I am trapped; I am outnumbered.

Teammate (50). teammate visible; teammates nearby; teammate in frame; ally player on screen; friendly player visible; multiple teammates visible; teammate in the distance; teammate close by; one teammate visible; two teammates visible; three teammates visible; four teammates visible; full team present; teammate is running;

teammate is crouching; teammate is holding position; teammate is advancing; teammate is retreating; teammate is shooting; teammate is in a firefight; teammate is engaging enemy; teammate is taking fire; teammate got a kill; teammate died; teammate is covering me; teammate is watching flank; teammate is throwing utility; teammates working together; coordinated team push; team executing strategy; teammates stacking a site; team split across map; teammates regrouping; trading kills with teammate; teammates covering angles; teammates crossfire setup; teammate holding bombsite; teammates on same site; teammates spread out; teammate in position; teammate flanking; teammates are alive; teammate just died; teammate is last alive; teammates are winning; teammates are losing; team has numbers advantage; team is outnumbered; team morale is high; team is struggling.

Enemy (50). enemy visible; enemy player spotted; hostile in sight; opponent on screen; enemy in frame; enemy in the distance; enemy up close; enemy behind cover; enemy exposed; one enemy visible; two enemies visible; multiple enemies; enemy team present; enemies everywhere; no enemies visible; enemy hiding; enemy is running; enemy is pushing; enemy is retreating; enemy is rotating; enemy is flanking; enemy is rushing; enemy is sneaking; enemy is shooting at me; enemy is aiming at me; enemy is throwing utility; enemy is planting bomb; enemy is defusing bomb; enemy got a kill; enemy died; enemy threat nearby; dangerous enemy position; enemy has advantage; enemy is vulnerable; enemy is dangerous; enemy is weak; enemy is strong; enemy in power position; enemy caught off guard; enemy is camping; enemy is holding angle; enemy is peeking; enemy is baiting; enemy is lurking; enemy aggression; enemy playing passive; enemy is reloading; enemy is distracted; enemy is isolated; enemy team coordinating.

Global (50). bomb is planted; bomb not planted yet; bomb is being planted; bomb is being defused; bomb is about to explode; bomb timer running; bomb at A site; bomb at B site; bomb dropped on ground; carrying the bomb; round just started; early round; mid round; late round; round is almost over; round ending soon; overtime round; match point; round is being won; round is being lost; close round; one-sided round; comeback happening; clutch situation; active combat; firefight in progress; intense gunfight; shots being fired; explosions happening; chaotic battle; quiet moment; calm before storm; lull in action; kills happening; players dying; trade kills; multi-kill; ace in progress; eco round; full buy round; force buy round; saving money; team is broke; team has money; expensive weapons visible; pistol round; anti-eco round; half buy round; low economy; high economy.

Spatial (50). bombsite A; bombsite B; middle area; spawn area; connector area; main entrance; back entrance; sniper nest; heaven position; hell position; indoor area; outdoor area; enclosed space; open space; corridor or hallway; large room; small room; narrow passage; wide area; near a wall; behind cover; near a box; on stairs; near a door; near a window; under an arch; in a corner; defensive position; offensive position; power position; weak position; exposed position; safe position; chokepoint; crossfire position; good cover available; no cover nearby; partial cover; head glitch spot; one-way angle; well lit area; dark area; shadowy corner; bright outdoor; dim indoor; smoke blocking view; flash effect; clear visibility; obstructed view; long sightline.

K ViZDoom control details

This section specifies the ViZDoom control of Section 5.7, which tests whether same-timestep cross-agent alignment can help closed-loop coordination in a setting where policy learning is possible. It does not replace the CS2 representation benchmark, since the environment, action interface, agent count, and optimization loop all differ. We therefore use it as a narrow control for embodied coordination rather than as a broad MARL benchmark.

Cooperative task. We implement a two-agent Remedy Rush Cooperative scenario in ViZDoom (Figure 14). The map continuously seeds linked pickup pairs, each with two armor pickups placed near one another but far enough apart that a single agent cannot reliably collect both within the short response window. Agents receive egocentric first-person observations and have no explicit communication channel. When either item of a pair is collected, the partner item despawns after 5 game tics unless it is also collected. The first collected item gives reward $+1$, and the partner item, if collected before despawning, gives an additional $+2$, so a solo pair yields total reward 1 and a coordinated pair yields 3. This design preserves individual reward progress while making temporally aligned two-agent behavior more valuable. Tables 27–29 list the environment, training, and algorithm settings.

Objective. Both methods optimize the clipped MAPPO objective (Schulman et al., 2017; Yu et al., 2022) with a centralized critic and decentralized visual policies. For minibatch index b and agent



Figure 14: **A separate closed-loop coordination test.** Two first-person views show linked armor pickups in Remedy Rush Cooperative; the right panel is a debugging view of trajectories and pickup locations, not an agent input. The short collection window rewards coordinated behavior, but the experiment does not measure the CS2 representation shift.

Table 27: **ViZDoom task settings.**

Setting	Value
Scenario	`remedy_rush_coop_multi_agent`
Agents	2
Engine and networking	ViZDoom multiplayer, async mode, netmode 1, ticrate 35
Rendering	RGB array; screen resolution 160×120
Policy observation	96×72 RGB, frame stack 1
Action repeat	4 skipped frames
Episode timeout	1000 environment steps
Object metadata	Enabled, maximum 64 objects
Parallel collection	16 environments; CPU rollout collection and storage
Evaluation	Deterministic policy evaluation; returns min–max normalized over both algorithms and all seeds

Table 28: **ViZDoom training settings.** Shared by MAPPO and CECL-MAPPO.

Setting	Value	Setting	Value
Seeds	3 per method	Backbone	Nature CNN
Reported points	5M, 10M env. frames	CNN channels	[32, 64, 64]
Parallel environments	16	CNN kernels	[8, 4, 3]
Rollout length	1024 steps	CNN strides	[4, 2, 1]
Minibatch size	4096	CNN paddings	[0, 0, 0]
PPO epochs per update	10	MLP hidden units	[512]
Learning rate	5×10^{-5}	Training device	Auto-selected accelerator
Discount γ	0.99	Sampling, buffer device	CPU
GAE λ	0.9		

a, let $r_{b,a}(\theta) = \pi_\theta(u_{b,a} \mid o_{b,a}) / \pi_{\theta_{\text{old}}}(u_{b,a} \mid o_{b,a})$ denote the PPO importance ratio, $\hat{A}_{b,a}$ the GAE advantage, and $\hat{R}_b$ the bootstrap return. The minimized loss is

$$\mathcal{L}_{\text{MAPPO}} = -\frac{1}{BA}\sum_{b=1}^{B}\sum_{a=1}^{A}\min\Big(r_{b,a}(\theta)\hat{A}_{b,a},\ \text{clip}\big(r_{b,a}(\theta), 1-\epsilon, 1+\epsilon\big)\hat{A}_{b,a}\Big) + \frac{c_v}{B}\sum_{b=1}^{B}\big(V_\theta(s_b) - \hat{R}_b\big)^2 - \frac{c_H}{BA}\sum_{b=1}^{B}\sum_{a=1}^{A}\mathcal{H}\big[\pi_\theta(\cdot \mid o_{b,a})\big],$$

where s_b is the centralized critic input and $o_{b,a}$ is the egocentric observation available to the decentralized policy. CECL-MAPPO adds a contrastive alignment term over same-timestep agent embeddings,

$$\mathcal{L}_{\text{CECL-RL}} = -\frac{1}{BA}\sum_{b=1}^{B}\sum_{a=1}^{A}\log\frac{\exp\big(\text{sim}(h_{b,a}, h_{b,\bar{a}})/\tau\big)}{\sum_{(b',a')\neq(b,a)}\exp\big(\text{sim}(h_{b,a}, h_{b',a'})/\tau\big)},$$

Table 29: **Algorithm-specific ViZDoom settings.** Shared entries are held fixed between MAPPO and CECL-MAPPO; a dash marks a setting that does not apply.

Setting	MAPPO	CECL-MAPPO
Algorithm	`mappo`	`contrastive_mappo`
Shared-parameter critic	true	true
Clip ϵ	0.2	0.2
Entropy coefficient	0.01	0.01
Critic coefficient	1.0	1.0
Critic loss	L2	L2
TD(λ) parameter	0.9	0.9
Policy scale mapping	`biased_softplus_1.0`	`biased_softplus_1.0`
Tanh-normal policy	true	true
Minibatch advantage normalization	false	false
Contrastive coefficient	–	0.05
Contrastive embedding dimension	–	128
Contrastive temperature	–	0.1
Contrastive hidden dimension	–	512
Contrastive embedding normalization	–	true

where $h_{b,a}$ is the normalized projected embedding of agent a, $\bar{a}$ is the other agent in the two-agent task, and other minibatch embeddings serve as negatives. The total CECL-MAPPO loss is

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{MAPPO}} + \lambda_{\text{CL}} \mathcal{L}_{\text{CECL-RL}}.$$

Evaluation. The primary metric is the deterministic evaluation return of the learned policy, separate from stochastic training rollouts. For visualization and aggregate comparison, returns are min–max normalized over the union of MAPPO and CECL-MAPPO runs,

$$\tilde{R}_{m,s,t} = \frac{R_{m,s,t} - R_{\min}}{R_{\max} - R_{\min}},$$

where m indexes method, s seed, and t evaluation time. Table 6 reports the saved interquartile-mean (IQM) aggregate with bootstrap intervals over seeds, following Agarwal et al. (2021). With three scores, this implementation's IQM equals their mean, and the small seed count limits the interpretation of its intervals. The available report contains the two normalized checkpoints, not the underlying raw returns, per-seed trajectories, or learning curves, so we draw no conclusion about the full learning dynamics.

Uncertainty and scope. At 5M frames, normalized means are 0.565 and 0.734, medians 0.563 and 0.731, and optimality gaps 0.435 and 0.266 for MAPPO and CECL-MAPPO. At 10M, IQMs are 0.685 [0.673, 0.693] and 0.679 [0.672, 0.684], with overlapping intervals. The seed bootstrap has only three source runs per method, so its narrow intervals should not be read as broad evidence of robustness. A five-seed replication and a second task have not been completed, and neither game-to-game transfer nor cross-domain replication of the representation shift is tested.