Pair-Repeat Allocation Effects in Paired-Acquisition Neural Factorization
Pair-Diversity and Anchor-Repetition Control

Matthew Vaishnav
Independent Researcher
Kitchener-Waterloo, Canada
Repository: paired-acquisition-factorization-allocation

Abstract

Paired-Acquisition Neural Factorization uses paired-region information to learn scanner-aware biological representations. A practical question is whether a fixed pair-presentation budget is better spent on many unique biological pairs or on repeated exposure to fewer anchor pairs. I evaluate this tradeoff under matched total pair-presentation budgets of 6,400 and 12,800 presentations. Across 40 evaluation cells per allocation, increasing biological pair diversity from 50 to 100 or 200 pairs improved the mean universal biological score and factor-separation score. The strongest seed-blocked effect occurred at budget 6,400 for 200 pairs versus 50 pairs, while 200 versus 100 pairs showed a small plateau effect. Doubling the total budget from 6,400 to 12,800 produced positive effects across all allocations. These results support a bounded resource-allocation claim: under matched total budget, Paired-Acquisition Neural Factorization benefits more from biological pair diversity than from repeated anchors alone, with diminishing returns beyond 100 pairs.

1 Introduction

Paired-acquisition objectives depend not only on the total number of pair presentations, but also on how those presentations are allocated. A fixed training budget can be spent on many unique biological pairs with fewer repetitions per pair, or on fewer pairs with more repeated anchor exposures. These choices affect whether the model sees broader biological variation or receives repeated consistency pressure on a smaller set of anchors.

This study isolates that allocation question. The total pair-presentation budget is held fixed while the number of unique biological pairs and requested anchor repetitions are varied inversely.

2 Pair-Repeat Design

I evaluate two total pair-presentation budgets: 6,400 and 12,800. Within each budget, the allocation grid uses 50, 100, and 200 unique biological pairs. Requested anchor repetitions are set so that total pair presentations remain matched within each budget.

Table 1: Matched-budget allocation means.

Budget	Pairs	Reps	Bio score	Factor sep.
6400	50	128	0.408051	0.748740
6400	100	64	0.424338	0.762593
6400	200	32	0.425916	0.766061
12800	50	256	0.448899	0.787414
12800	100	128	0.459528	0.792195
12800	200	64	0.461950	0.798729

3 Allocation Contrasts

Seed-blocked contrasts compare allocation choices within the same total pair-presentation budget. The clearest result is that 200 pairs outperform 50 pairs at budget 6,400, with a positive bootstrap interval and all seed blocks favorable. The 200-minus-100 contrasts are much smaller, indicating a plateau rather than a strong additional gain beyond 100 pairs.

Table 2: Seed-blocked allocation contrasts.

Budget	Contrast	Mean diff.	CI low	CI high	p
6400	100–50	+0.016287	-0.001691	+0.032845	0.113281
6400	200–50	+0.017865	+0.010560	+0.026079	0.001953
6400	200–100	+0.001578	-0.010215	+0.014685	0.822266
12800	100–50	+0.010629	-0.004164	+0.023699	0.195312
12800	200–50	+0.013050	+0.001472	+0.025124	0.076172
12800	200–100	+0.002422	-0.006368	+0.011081	0.607422

4 Budget Doubling

The second question asks whether increasing the total pair-presentation budget improves the biological score. Doubling from 6,400 to 12,800 was positive for every allocation and for the overall matched comparison.

Table 3: Budget-doubling effects: 12,800 minus 6,400.

Scope	Pairs	Mean diff.	CI low	p
By allocation	50	+0.040848	+0.027168	0.003906
By allocation	100	+0.035189	+0.026867	0.001953
By allocation	200	+0.036034	+0.026404	0.001953
Overall	All	+0.037357	+0.030315	0.001953

5 Interpretation

The pair-repeat results support a resource-allocation principle for Paired-Acquisition Neural Factorization. At fixed total pair-presentation budget, increasing unique biological pair diversity improves biological consistency and factor separation more than repeatedly presenting a small anchor set. The clearest allocation contrast is 200 pairs versus 50 pairs at budget 6,400. The much smaller 200-minus-100 contrasts indicate diminishing returns beyond 100 pairs in this grid.

This result should not be interpreted as a universal law for all datasets or encoders. It is a controlled allocation study under the Paired-Acquisition Neural Factorization evaluation design. Its role is to constrain training-budget choices, not to claim diagnostic performance.

6 Reproducibility and Claim Boundary

The accompanying repository contains the frozen CSV evidence tables, a verification script, and this paper source. The supported claim is narrow: under matched pair-presentation budgets, Paired-Acquisition Neural Factorization benefits from biological pair diversity, with a plateau after 100 pairs in the tested grid, and with positive effects from doubling total budget.

7 Conclusion

Under matched total pair-presentation budget, Paired-Acquisition Neural Factorization improved more when the budget was allocated toward broader biological pair diversity than when it was concentrated into repeated anchors. The strongest allocation contrast was 200 versus 50 pairs at budget 6,400. Budget doubling produced positive effects across all allocations.