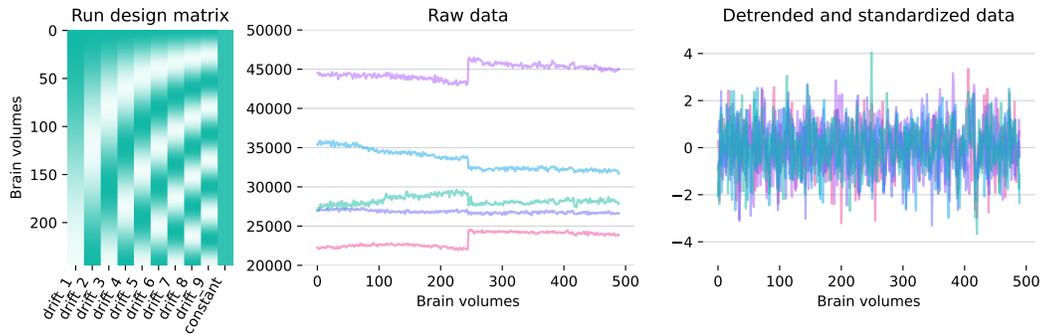# A APPENDIX

A.1 DATA PRE-PROCESSING

Figure S1: **Pre-processing of the Wen 2017 dataset** For each subject and each run, in each vertex, we regress out parts of the signal which can be linearly explained by the design matrix represented on the left, which models cosine drifts of the BOLD signal. The two graphs to the right show time-courses in 5 vertices across 2 different runs before (left) and after (right) they have been pre-processed.
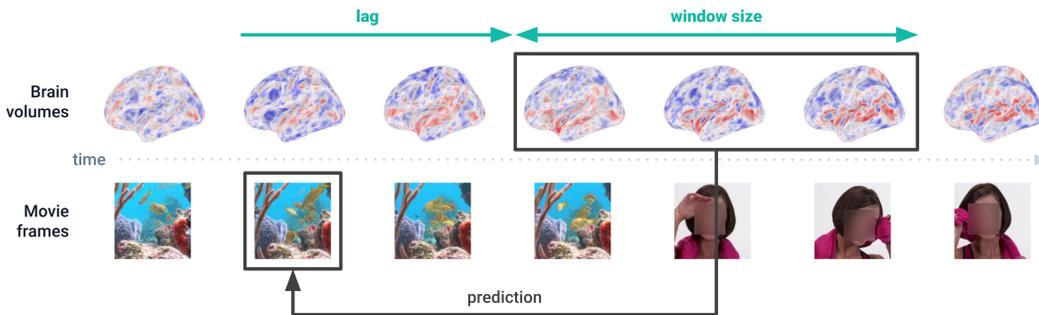

Figure S2: **Lag and window size** In order to decode a movie frame which was seen at time t, one can use brain volumes which were acquired further in time. This delay is referred to as the *lag*. Moreover, one can use several brain volumes to decode a given movie frame. The number of brain volumes used is called the *window size*. Images featuring human faces were blurred.

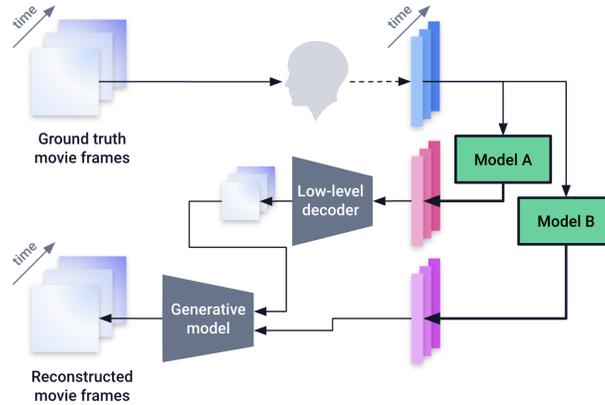## A.2 RESULTS FOR EVERY TYPE OF LATENT REPRESENTATION

14

Figure S3: **Using predicted latents as inputs of generative models** Although out-of-scope for this study, it is possible to use predicted latents as inputs to one or multiple generative models to reconstruct visual stimuli seen by individuals.
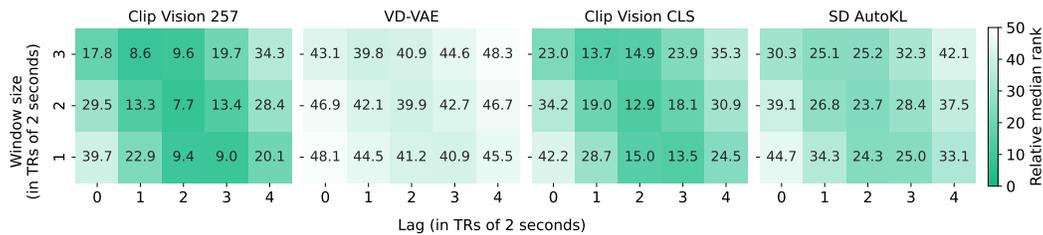

Figure S4: **Relative median rank ↓ of predicted latents averaged across subjects for various time lags and window sizes**
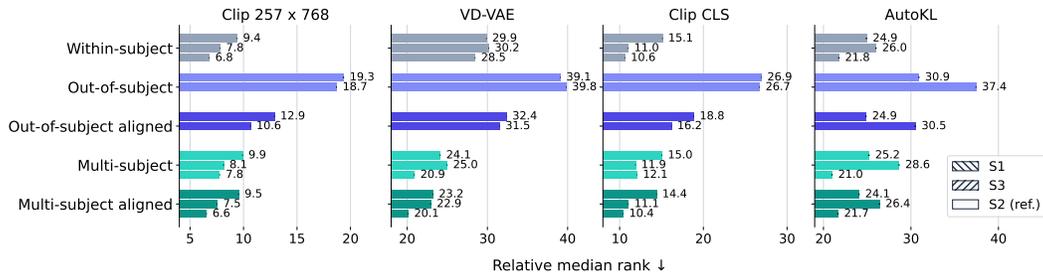

Figure S5: **Effects of alignment** For any type of latent representation, out-of-subject decoding performance, measured through relative median rank ↓, greatly improves when subjects are functionally aligned. Training decoders on multiple subjects also works better when subjects are aligned. These results were averaged across 50 retrieval sets ; all these metrics are reported with a standard error of the mean (SEM) smaller than 0.01.
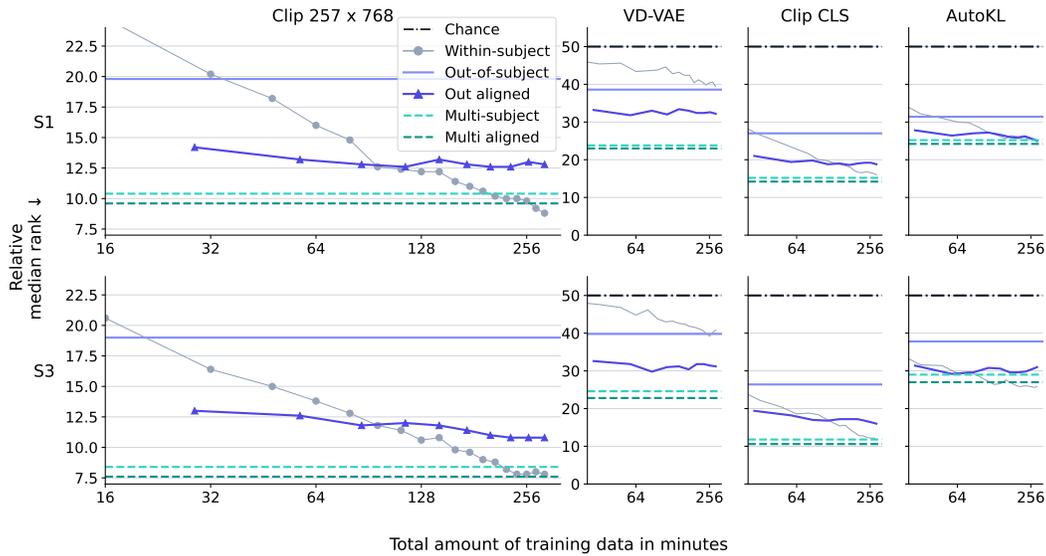
Figure S6: **Performance increases slightly with more alignment data** For any type of latent representation, out-of-subject decoding performance greatly increases with functional alignment even in low data regimes. In high data regimes, out-of-subject decoding does not work as well as fitting single-subject or multi-subject models.
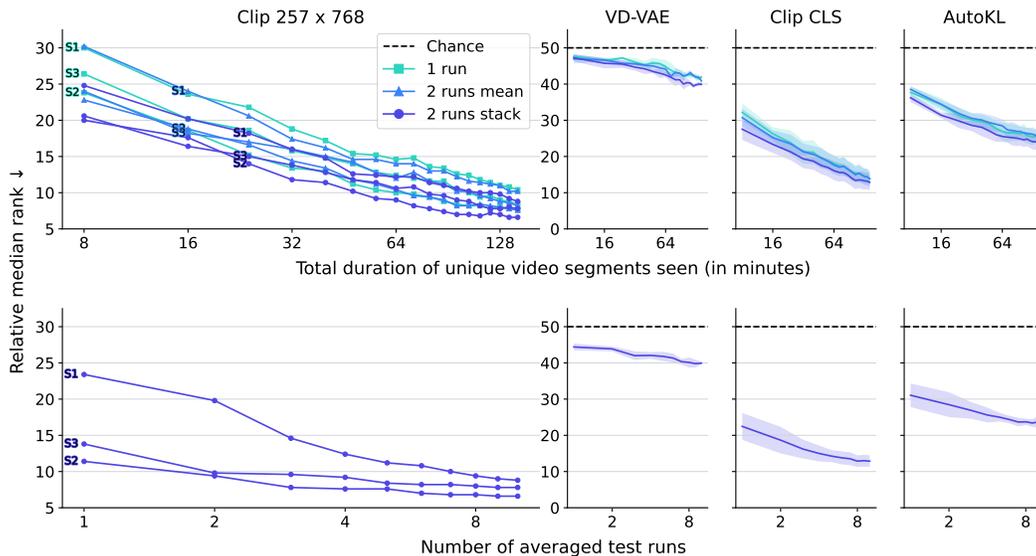

Figure S7: **Scaling studies for all latents** For any type of latent representation, decoding performance increases linearly with exponentially more data. It also seems that, when acquiring data at 3T or more, not repeating stimuli yields the best results. At test time, although repeating stimuli allows to get better metrics, retrieval performance with only one repetition is already reasonable in 2 out of 3 subjects of the Wen 2017 dataset.

Figure S8: Relative median rank ↓ for **CLIP 257 × 768** latents in single- and multi-subject training sets, with and without alignment, tested on within- and across-subjects setups with and without alignment. These results were averaged across 50 retrieval sets ; all these metrics are reported with a standard error of the mean (SEM) smaller than 0.01.

Training data (rows) × **Test data** (columns)

Training data \ Test data	S1	S2	S3	S1 (10%)→S2	S1 (20%)→S2	S1 (30%)→S2	S1 (40%)→S2	S1 (50%)→S2	S1 (60%)→S2	S1 (70%)→S2	S1 (80%)→S2	S1 (90%)→S2	S1 (100%)→S2	S3 (10%)→S2	S3 (20%)→S2	S3 (30%)→S2	S3 (40%)→S2	S3 (50%)→S2	S3 (60%)→S2	S3 (70%)→S2	S3 (80%)→S2	S3 (90%)→S2	S3 (100%)→S2
S1	29.9	34.0	37.4	36.6	35.5	34.3	34.8	35.0	36.0	35.6	35.1	34.5	33.0	37.6	37.8	37.5	36.8	35.6	35.0	37.7	37.3	37.7	38.2
S2	39.1	28.5	39.8	33.1	32.3	33.0	32.2	33.6	33.0	32.8	32.4	32.8	32.4	32.7	31.7	30.2	31.2	31.2	30.8	31.7	31.7	31.5	31.5
S3	40.9	35.1	30.2	37.5	39.0	37.4	36.0	36.4	37.3	37.2	36.4	35.4	36.9	34.2	36.4	34.7	33.1	35.3	34.2	34.7	35.4	36.1	34.8
S1, S2	25.0	21.4	32.0	27.3	26.7	26.0	25.7	26.1	25.4	24.8	24.8	25.1	23.8	30.1	29.2	28.1	28.5	28.7	25.9	25.9	26.0	26.1	26.7
S2, S3	34.9	21.7	24.6	27.4	27.4	26.6	26.8	27.4	25.8	26.6	25.7	26.4	26.3	29.3	27.0	25.2	26.0	26.0	25.2	26.8	24.5	26.0	26.0
S1, S2, S3	24.1	20.9	25.0	27.2	27.7	26.4	26.7	26.3	25.6	26.2	25.5	25.3	25.3	29.1	29.1	27.2	27.5	27.6	26.4	26.0	25.8	26.5	26.5
S1 (10%)→S2, S2	34.7	21.1	38.0	28.7	24.8	26.4	25.8	25.7	25.7	25.1	25.2	25.2	25.0	29.4	27.3	26.6	27.1	27.7	26.6	27.0	26.3	26.7	26.8
S1 (20%)→S2, S2	35.1	21.3	37.5	27.5	27.3	26.2	25.4	25.9	26.0	24.5	24.8	25.2	24.4	28.8	27.6	26.5	27.2	26.8	27.2	26.6	25.3	26.6	25.8
S1 (30%)→S2, S2	35.6	20.8	38.5	26.7	25.8	26.8	24.5	25.6	25.0	24.1	23.9	23.3	23.9	29.3	27.3	27.8	27.6	27.5	26.6	27.1	25.6	25.8	26.1
S1 (40%)→S2, S2	34.4	20.6	38.7	26.9	24.7	24.5	24.6	24.6	24.2	23.8	22.8	23.6	23.4	29.1	27.5	26.8	26.7	26.9	27.0	26.8	26.4	26.6	26.8
S1 (50%)→S2, S2	34.8	20.1	38.1	26.8	23.9	24.7	23.8	25.9	24.4	23.0	22.1	23.0	22.5	28.8	27.4	26.4	26.5	27.1	26.9	26.3	26.4	26.5	26.2
S1 (60%)→S2, S2	34.6	20.6	38.1	25.6	24.7	23.2	24.3	25.0	25.8	22.6	22.9	23.5	23.4	29.2	27.0	25.6	27.2	27.5	27.3	26.6	27.0	26.1	26.1
S1 (70%)→S2, S2	33.7	20.1	38.0	26.1	24.0	24.9	23.0	23.6	23.6	24.6	22.4	22.6	22.8	28.1	26.7	26.3	26.6	27.7	26.6	26.1	26.6	26.0	26.2
S1 (80%)→S2, S2	33.4	20.7	38.1	26.2	24.1	24.6	23.6	24.3	24.0	22.9	25.1	23.5	23.1	29.2	26.6	26.9	27.0	26.9	26.7	26.7	26.6	25.3	26.0
S1 (90%)→S2, S2	32.5	21.4	38.4	26.1	24.3	24.3	23.9	24.4	24.0	22.8	23.3	24.1	22.4	29.8	28.1	26.7	26.1	26.8	26.6	26.7	25.9	26.1	25.8
S1 (100%)→S2, S2	31.4	20.4	38.1	25.1	24.0	23.6	23.5	23.4	23.0	22.6	22.3	22.7	23.6	29.3	26.2	26.2	26.5	26.9	25.7	26.2	25.9	25.0	25.7
S3 (10%)→S2, S2	35.0	21.2	37.5	27.5	26.3	26.5	26.4	26.4	26.5	26.1	25.8	26.3	26.4	30.1	27.2	26.6	26.1	27.1	26.2	26.6	26.0	25.8	26.5
S3 (20%)→S2, S2	35.7	21.1	37.3	27.9	26.5	26.4	25.7	26.3	26.9	25.7	25.6	25.4	25.4	28.6	29.5	26.1	25.9	27.2	26.0	26.0	26.0	26.1	26.2
S3 (30%)→S2, S2	36.1	21.1	36.2	28.2	26.8	26.7	26.4	26.7	25.8	26.8	25.5	25.9	27.0	29.3	27.2	27.9	26.0	27.1	25.6	26.3	25.3	25.9	25.9
S3 (40%)→S2, S2	36.1	21.2	36.7	27.8	28.3	26.8	26.4	26.3	25.6	25.9	25.1	26.7	26.2	28.9	26.2	25.0	26.4	26.4	26.5	25.4	25.0	25.0	25.7
S3 (50%)→S2, S2	35.0	21.1	36.2	27.0	26.5	27.5	26.9	27.0	26.0	26.1	26.1	26.4	26.3	28.0	25.2	24.6	25.4	26.6	24.9	25.6	25.1	25.5	24.9
S3 (60%)→S2, S2	35.4	21.1	35.7	28.2	26.9	27.0	27.0	26.7	26.9	26.8	25.7	25.9	25.8	27.3	26.2	24.5	24.6	25.1	25.7	24.9	25.3	25.3	24.9
S3 (70%)→S2, S2	34.8	21.2	35.9	27.8	26.1	26.9	26.4	28.1	27.4	26.8	26.7	27.3	26.0	28.6	25.2	24.8	24.2	26.0	24.0	24.2	24.6	24.7	24.5
S3 (80%)→S2, S2	36.1	21.5	34.1	27.8	25.9	26.6	26.9	28.3	26.9	26.9	26.9	27.4	26.2	27.4	26.4	24.4	24.8	25.0	23.2	23.9	24.0	23.6	23.6
S3 (90%)→S2, S2	35.1	21.0	33.5	27.0	25.9	26.9	27.0	27.0	27.3	26.6	27.5	27.4	26.3	28.2	25.8	24.4	24.4	24.7	23.6	24.2	23.7	23.8	22.6
S3 (100%)→S2, S2	34.9	21.2	33.4	27.1	26.1	26.6	25.8	27.4	26.9	26.3	25.6	25.7	26.0	27.9	25.8	23.9	23.3	24.8	23.6	22.8	23.4	23.4	23.0
S1 (10%)→S2, S3 (10%)→S2, S2	34.4	20.9	37.9	28.7	25.0	26.2	25.5	26.0	26.5	25.9	25.1	25.1	25.2	30.0	27.6	26.5	26.9	27.9	26.6	27.3	27.1	26.3	27.0
S1 (20%)→S2, S3 (20%)→S2, S2	35.4	21.0	36.9	27.0	27.4	26.0	25.3	25.6	26.6	24.7	24.4	25.0	24.8	29.3	29.4	25.8	26.2	26.7	26.4	26.5	26.6	26.1	25.8
S1 (30%)→S2, S3 (30%)→S2, S2	36.7	20.8	37.6	27.3	25.8	27.2	25.6	25.5	25.2	25.3	24.4	24.5	24.8	29.1	27.4	28.4	26.5	27.3	26.0	27.4	25.3	25.8	25.9
S1 (40%)→S2, S3 (40%)→S2, S2	35.6	20.6	38.9	26.0	25.9	25.0	25.5	24.1	25.0	24.8	23.5	24.7	24.5	29.3	27.0	26.1	26.6	27.2	26.6	26.1	25.6	26.0	25.6
S1 (50%)→S2, S3 (50%)→S2, S2	34.7	20.3	37.8	26.3	24.2	24.9	23.7	25.8	24.7	24.1	23.3	24.2	23.9	29.6	26.1	25.2	25.5	25.6	24.4	25.2	24.6	24.8	25.0
S1 (60%)→S2, S3 (60%)→S2, S2	34.8	20.3	37.9	26.7	24.6	23.9	23.9	24.5	26.0	24.9	23.2	23.7	23.5	28.5	25.4	25.2	25.0	24.8	24.9	25.0	24.9	24.9	25.4
S1 (70%)→S2, S3 (70%)→S2, S2	34.3	19.9	36.3	26.4	24.7	23.9	23.8	24.6	24.4	24.8	24.0	23.8	23.5	29.1	26.2	24.9	24.8	26.0	24.6	23.9	24.8	25.0	24.4
S1 (80%)→S2, S3 (80%)→S2, S2	34.1	20.7	35.6	26.3	24.4	24.8	24.6	25.6	24.8	24.6	24.9	25.2	24.0	28.4	26.7	24.3	25.5	24.8	23.8	22.9	24.2	23.8	23.7
S1 (90%)→S2, S3 (90%)→S2, S2	32.9	21.2	36.1	26.4	23.8	24.1	23.9	25.2	25.7	24.4	23.6	24.5	23.2	28.2	26.4	25.4	24.9	26.0	24.4	23.3	23.5	23.7	24.0
S1 (100%)→S2, S3 (100%)→S2, S2	33.1	20.1	35.0	26.5	24.5	24.2	23.6	23.5	23.6	23.5	23.0	23.4	23.2	28.3	25.9	24.6	23.7	24.7	23.1	22.8	23.1	22.9	22.9

Figure S9: Relative median rank ↓ for **VD-VAE** latents in single- and multi-subject training sets, with and without alignment, tested on within- and across-subjects setups with and without alignment. These results were averaged across 50 retrieval sets ; all these metrics are reported with a standard error of the mean (SEM) smaller than 0.01.

Training data (rows) / **Test data** (columns)

Training data	S1	S2	S3	S1 (10%) → S2	S1 (20%) → S2	S1 (30%) → S2	S1 (40%) → S2	S1 (50%) → S2	S1 (60%) → S2	S1 (70%) → S2	S1 (80%) → S2	S1 (90%) → S2	S1 (100%) → S2	S3 (10%) → S2	S3 (20%) → S2	S3 (30%) → S2	S3 (40%) → S2	S3 (50%) → S2	S3 (60%) → S2	S3 (70%) → S2	S3 (80%) → S2	S3 (90%) → S2	S3 (100%) → S2
S1	15.1	24.5	26.3	25.2	23.9	24.4	25.0	24.9	24.6	25.4	24.4	24.7	24.5	27.3	27.8	27.1	26.4	27.3	26.8	27.2	27.0	26.7	27.1
S2	26.9	10.6	26.7	21.3	19.6	20.1	18.9	18.9	18.8	18.9	19.3	19.3	18.8	19.4	18.3	17.5	16.7	17.0	16.9	17.1	16.6	16.5	16.2
S3	24.3	21.8	11.0	27.5	26.0	25.6	26.2	25.2	25.5	25.6	26.1	25.4	25.4	21.8	20.3	20.0	19.9	19.8	19.2	19.6	19.9	19.8	19.7
S1, S2	14.9	11.1	22.3	20.0	18.4	19.4	18.9	18.2	18.2	18.1	17.9	18.3	18.2	20.5	19.2	18.7	18.3	18.6	17.7	17.7	17.0	17.5	17.5
S2, S3	22.5	11.1	11.2	21.1	19.7	19.7	18.8	18.7	18.8	19.2	18.7	19.3	19.3	18.3	16.4	15.6	15.8	16.0	15.5	15.3	15.1	15.1	15.4
S1, S2, S3	15.0	12.1	11.9	20.1	18.5	19.0	18.9	18.2	18.5	18.6	18.5	18.5	18.7	18.9	18.0	17.7	17.0	17.9	16.6	16.9	16.3	16.3	16.6
S1 (10%) → S2, S2	26.1	10.6	25.8	20.7	19.2	19.5	18.5	18.7	19.2	18.9	19.1	19.2	18.5	18.9	18.2	16.9	17.0	16.4	16.3	16.5	16.5	16.1	16.0
S1 (20%) → S2, S2	25.8	10.5	25.7	19.8	18.3	18.6	18.3	18.2	18.4	17.9	17.8	17.9	18.1	18.9	17.7	17.5	16.7	16.5	16.5	16.2	16.4	16.1	15.9
S1 (30%) → S2, S2	26.1	10.5	25.3	19.9	17.4	16.9	17.1	17.4	17.6	17.3	17.3	17.7	17.6	19.2	17.7	17.2	16.5	16.4	16.2	16.0	15.8	15.9	15.8
S1 (40%) → S2, S2	25.9	10.6	25.3	19.6	17.7	17.7	16.8	17.3	17.3	17.0	17.2	17.2	17.1	19.5	17.5	17.0	16.0	16.1	15.8	15.7	15.5	15.5	15.5
S1 (50%) → S2, S2	25.7	10.5	25.3	18.9	16.8	17.4	16.3	16.4	16.3	16.2	16.4	16.8	16.6	18.7	17.4	16.5	16.3	15.8	15.2	15.6	15.4	15.4	15.4
S1 (60%) → S2, S2	25.3	10.8	24.7	18.6	17.3	17.4	16.3	15.9	16.1	16.2	16.3	16.6	16.1	19.1	17.6	16.9	16.5	16.8	16.2	16.0	15.8	15.8	16.0
S1 (70%) → S2, S2	25.7	10.7	25.0	18.7	17.3	17.5	16.1	15.7	15.5	15.3	15.5	16.0	16.0	18.7	17.4	16.6	16.0	16.6	15.8	16.1	16.1	15.8	15.9
S1 (80%) → S2, S2	25.7	10.7	24.8	18.9	16.5	17.6	16.2	16.4	15.8	15.4	15.5	15.7	15.7	18.9	17.6	16.8	16.4	16.4	16.1	16.1	16.3	15.9	15.7
S1 (90%) → S2, S2	24.4	10.7	25.3	18.9	17.1	18.0	16.9	16.1	16.3	15.9	15.6	15.4	15.4	19.2	17.9	16.9	15.6	16.3	16.0	16.2	16.4	15.9	15.5
S1 (100%) → S2, S2	24.2	10.7	24.7	18.9	16.7	17.9	16.2	15.8	16.4	15.9	15.2	15.3	14.9	19.0	17.6	16.9	15.9	16.0	16.1	16.2	16.1	15.8	15.3
S3 (10%) → S2, S2	26.0	10.6	24.4	21.3	19.2	20.2	18.6	18.7	19.0	18.8	19.4	19.5	18.9	16.1	17.2	16.3	15.8	15.6	15.6	15.9	15.9	15.5	15.2
S3 (20%) → S2, S2	25.7	10.4	23.9	20.8	19.1	19.6	18.1	18.3	18.2	18.3	18.4	18.8	18.4	17.1	15.2	15.9	15.2	15.1	14.9	14.9	15.2	14.8	14.5
S3 (30%) → S2, S2	25.9	10.3	23.4	21.2	19.5	19.4	17.8	18.2	18.4	18.2	18.5	18.6	18.4	17.4	16.4	15.5	15.2	14.9	14.9	15.1	15.1	14.6	14.6
S3 (40%) → S2, S2	26.0	10.1	23.1	20.8	19.5	19.3	18.4	18.4	18.0	18.3	18.4	18.0	18.4	17.4	16.1	15.2	14.8	14.5	14.1	14.2	14.4	14.1	13.8
S3 (50%) → S2, S2	26.5	10.3	22.5	20.8	19.6	19.6	18.7	18.4	18.4	18.1	18.0	18.5	18.7	17.8	16.5	15.1	14.1	13.9	13.5	13.8	13.5	13.3	13.2
S3 (60%) → S2, S2	25.5	10.3	23.4	21.5	19.8	19.6	18.8	18.2	18.3	18.3	17.8	18.5	18.1	17.2	15.7	14.8	14.1	14.1	13.3	13.2	13.3	13.4	13.3
S3 (70%) → S2, S2	25.8	10.6	22.7	20.7	19.3	18.7	18.1	18.0	18.3	18.4	17.9	18.4	18.1	16.8	15.3	14.3	13.8	13.7	12.7	12.2	12.3	12.2	12.2
S3 (80%) → S2, S2	25.9	10.4	21.8	21.2	19.3	19.1	18.2	18.6	18.5	18.2	18.1	18.3	18.5	16.9	15.7	14.1	14.1	13.3	12.8	12.5	11.9	12.0	11.9
S3 (90%) → S2, S2	26.0	10.3	21.6	20.7	19.3	19.7	18.4	18.3	18.5	18.6	18.4	18.9	18.5	16.3	15.7	14.3	13.7	13.9	12.7	12.4	12.5	11.7	11.8
S3 (100%) → S2, S2	25.3	10.4	21.9	20.3	19.3	19.5	18.0	17.8	18.1	17.7	18.0	18.8	18.2	16.1	15.3	14.3	13.2	13.4	12.9	12.0	11.6	11.5	11.0
S1 (10%) → S2, S3 (10%) → S2, S2	25.8	10.7	24.7	21.0	19.5	19.5	18.7	18.6	18.6	18.8	18.8	19.1	18.4	16.1	17.3	16.4	16.0	15.7	15.3	15.8	15.6	15.3	15.3
S1 (20%) → S2, S3 (20%) → S2, S2	25.3	10.5	24.4	20.1	18.2	18.3	17.8	17.8	17.8	17.6	17.7	18.2	18.0	17.8	15.3	16.5	15.7	15.5	15.1	15.3	15.3	14.9	15.0
S1 (30%) → S2, S3 (30%) → S2, S2	25.1	10.4	23.5	19.8	17.8	17.0	17.3	17.4	17.3	17.2	16.9	17.3	17.4	18.0	16.1	15.8	15.2	15.2	15.3	15.0	15.1	15.1	14.7
S1 (40%) → S2, S3 (40%) → S2, S2	25.3	10.4	23.2	19.9	17.4	17.6	16.5	16.8	16.8	16.9	16.8	17.0	16.6	17.8	16.4	15.4	14.9	14.9	14.1	14.3	14.4	14.4	14.2
S1 (50%) → S2, S3 (50%) → S2, S2	24.9	10.3	22.5	19.4	17.3	17.8	17.0	16.1	15.8	15.9	16.3	16.5	16.6	17.7	16.0	15.0	14.6	14.2	13.4	13.7	13.4	13.8	13.8
S1 (60%) → S2, S3 (60%) → S2, S2	25.0	10.6	22.4	19.3	17.8	18.5	16.6	16.0	15.7	15.8	16.2	16.6	15.8	17.3	15.5	15.1	14.4	14.2	13.4	13.0	13.1	13.6	13.3
S1 (70%) → S2, S3 (70%) → S2, S2	24.9	10.5	22.6	18.6	17.3	17.9	16.3	15.5	15.5	15.2	15.0	15.6	15.3	16.2	15.2	14.3	13.5	13.5	13.2	12.3	12.4	12.5	12.4
S1 (80%) → S2, S3 (80%) → S2, S2	25.1	10.7	22.1	19.6	17.1	17.4	16.5	16.1	16.0	15.4	15.3	15.8	15.5	17.1	15.8	14.3	14.3	13.5	13.0	12.8	12.0	12.7	12.0
S1 (90%) → S2, S3 (90%) → S2, S2	24.2	10.6	22.2	19.3	17.3	18.2	16.4	16.1	16.2	15.9	15.5	15.3	15.3	16.7	15.6	14.6	13.8	14.3	13.1	12.8	13.0	12.0	11.6
S1 (100%) → S2, S3 (100%) → S2, S2	23.8	10.4	21.8	18.9	17.5	17.9	16.4	15.9	15.9	15.6	15.4	15.2	14.4	17.0	15.4	14.6	13.5	13.6	13.1	12.5	12.4	12.3	11.1

Figure S10: Relative median rank ↓ for **CLIP CLS** latents in single- and multi-subject training sets, with and without alignment, tested on within- and across-subjects setups with and without alignment. These results were averaged across 50 retrieval sets ; all these metrics are reported with a standard error of the mean (SEM) smaller than 0.01.

Figure S11: Relative median rank ↓ for **AutoKL** latents in single- and multi-subject training sets, with and without alignment, tested on within- and across-subjects setups with and without alignment. These results were averaged across 50 retrieval sets ; all these metrics are reported with a standard error of the mean (SEM) smaller than 0.01.