Appendix

Target movement

Let us denote $\phi = (1/30) \cdot s + \alpha$ as a time step constant, where s denoted the current step number of a 250 long episode, and α marks a constant offset. Unless stated otherwise, our tests state the offset at the constant value of 3. Thus, the location of the target takes the exact form

$$\Psi_t[0] = 3 \cdot \sin(1.8 \cdot \phi) + \\ 3.4 \cdot \sin(1.8 \cdot \phi) + 2.5 \cdot \sin(1.82 \cdot \phi) + 4.3 \cdot \sin(2.34 \cdot \phi) \quad (5)$$

$$\Psi_t[1] = 3 \cdot \sin(1.1 \cdot \phi) + \\ 3.2 \cdot \sin(3.6 \cdot \phi) + 3.8 \cdot \sin(2.5 \cdot \phi) + 4.8 \cdot \sin(1.48 \cdot \phi) \quad (6)$$

Experimental Results

Other connectivity modality: note we tested several learning step sizes for each modality and came to the conclusion that 1500 is a fair amount for each one.

Main Experiment: The evaluation was performed on 50 episodes for the FC and on 26 episodes for the WESL task. Game units as learning sessions with 1500 steps each, across 10 learners, after outlier removal.

To confirm the difference in results between the connected and not-connected settings, a series of statistical tests were performed as follows; For each of the four hypotheses, a t-test was conducted on the Frechet distance values of the learner when connected to a virtual teacher, with low-high modality, versus the ones in the not-connected settings. The p-values for the FC task across the four hypotheses are described in Tables 2 and 3.

Robustness tests:

To assess the robustness of our AI teacher, we carried out extensive tests under various real-world scenarios. We introduced learners with varying skills, learning rates, and attention levels, and had the AI teach them two tasks. Noise was added to simulate occasional learner errors, with types ranging from Gaussian to random and varying degrees of intensity. All 7 game units were tested, representing different session lengths. Results are shown in Tables 4 and 5.

Game Unit	Mean Not-connected	Mean Connected	p-value
1	0.235	0.524	0.016
2	0.437	0.794	$< 0.5 \cdot 10^{-4}$
3	0.491	0.814	< 0.0005
4	0.427	0.874	$< 0.1 \cdot 10^{-5}$
5	0.53	0.873	< 0.0005
6	0.662	0.906	0.001
7	0.617	0.892	0.005

Table 2: H1 results on FC task. Columns 2 and 3 show the median Frechet distance at each game unit across all repetitions, and column 4 shows the p-values for pairwise t-tests.

Game Unit	Not-connected	Connected	p-value
1	0.404	0.5	0.23
2	0.51	0.69	0.0001
3	0.487	0.716	0.0014
4	0.537	0.754	$< 0.5 \cdot 10^{-4}$
5	0.528	0.739	$< 0.5 \cdot 10^{-6}$
6	0.56	0.79	$< 0.5 \cdot 10^{-5}$

Table 3: H1 results on WESL task. Columns 2 and 3 show the median Frechet distance at each game unit across all repetitions, column 4 shows the p-values for pairwise t-tests.

Noise Variance	Noise Type	Game Unit	Distance Not Connected	Distance Connected
4	normal	4	0.62	0.91
4	normal	5	0.63	0.92
4	uniform	4	0.63	0.92
4	uniform	5	0.62	0.91
6	normal	4	0.67	0.91
6	normal	5	0.71	0.93
6	uniform	4	0.69	0.91
6	uniform	5	0.72	0.89
8	normal	4	0.61	0.9
8	normal	5	0.65	0.91
8	uniform	4	0.63	0.89
8	uniform	5	0.69	0.89
10	normal	4	0.61	0.9
10	normal	5	0.7	0.92
10	uniform	4	0.51	0.9
10	uniform	5	0.7	0.92

Table 4: Test results of learners with different capabilities on the FC task.

Noise Variance	Noise Type	Game Unit	Distance Not Connected	Distance Connected
4	normal	4	0.48	0.67
4	normal	5	0.55	0.72
4	uniform	4	0.58	0.73
4	uniform	5	0.58	0.73
6	normal	4	0.56	0.72
6	normal	5	0.54	0.75
6	uniform	4	0.55	0.7
6	uniform	5	0.56	0.74
8	normal	4	0.56	0.69
8	normal	5	0.59	0.72
8	uniform	4	0.56	0.68
8	uniform	5	0.58	0.74
10	normal	4	0.57	0.68
10	normal	5	0.56	0.73
10	uniform	4	0.54	0.68
10	uniform	5	0.6	0.74

Table 5: Test results of learners with different capabilities on the WESL task.