**Electronic Supplementary Material**

This supplementary material accompanies the research article Pagnotta, Laland, & Coco (in press). “Attentional coordination in demonstrator-observer dyads facilitates learning and predicts performance in a novel manual task.”

Additional materials, including examples of the demonstration videos, the data, and a detailed tutorial with scripts to reproduce the procedure used to fit, compare, interpret, and report the models, can be found in the Open Science Framework project associated with this research, link: https://osf.io/jhtqb.

Contents

[1. Snapshots of the demonstration videos 2](#_Toc37427630)

[2. Control conditions: Is this really a case of ‘social’ learning? 3](#_Toc37427631)

[3. Effects of face visibility on learners’ overt attention. 6](#_Toc37427632)

[4. Learning rate and attentional coordination: model comparison 7](#_Toc37427633)

[5. Additional figures reporting the final models including DET and L 8](#_Toc37427634)

[6. Recurrence Quantification Analysis: models including LAM and TT 10](#_Toc37427635)

[7. Transcriptions of the verbal instructions for each task 12](#_Toc37427636)

# Snapshots of the demonstration videos



**Figure S1** Video frames of the demonstration video for the star puzzle in the conditions noFACE\_noAUDIO (top row) and FACE\_AUDIO (bottom row). The overlaid dots represent the eye tracking data of all participants watching each video. Participants mostly looked at the pieces manipulated by the demonstrator, and they looked at the demonstrator’s face more when it was visible than blurred.

# Control conditions: Is this really a case of ‘social’ learning?

We ran two control conditions to test whether our experimental paradigm may indeed be considered a case of ‘social’ learning by showing participants a still image of the initial frame of the demonstration (see **Figure S2**) with either the audio track included (noVIDEO\_AUDIO) or not (noVIDEO\_noAUDIO) and run the same procedure reported in the main text.

A group of people posing for the camera

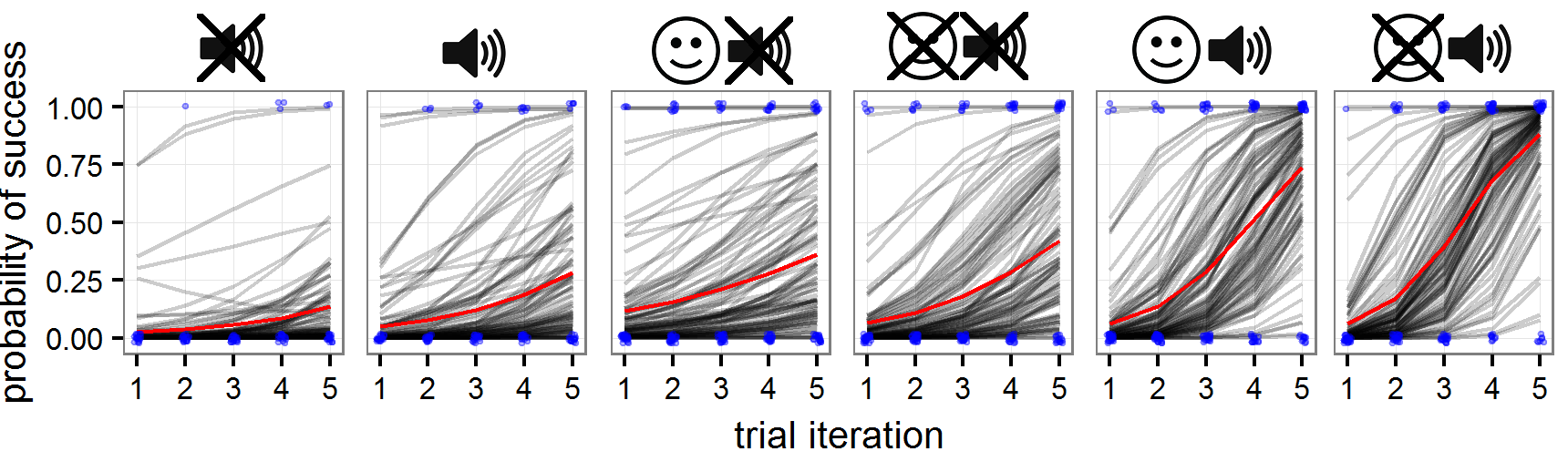
Description automatically generated

**Figure S2** The still images used as the visual stimuli in the control conditions, showing the demonstrator in a neutral position and the puzzle pieces on the table.

We analysed the participant’s performance at solving the puzzles (a binomial variable, correct vs. incorrect) across all conditions of the study: the two controls conditions described above (noVIDEO\_noAUDIO and noVIDEO\_AUDIO) and the four experimental conditions described in the main text (noFACE\_noAUDIO; FACE\_noAUDIO; noFACE\_AUDIO and FACE\_AUDIO). The initial dataset consisted of 795 trials (53 participants x 3 puzzles x 5 iterations). One participant knew one of the puzzles and the corresponding trials were removed from the analysis, therefore the final dataset included 790 trials. We fitted hierarchical logistic models with logit link to the performance outcome as a function of the experimental conditions, iteration, and their interaction. No participant solved any of the tasks in their first attempt so there was no need to include the baseline score as a covariate. **Table S1** reports the effect of iteration on performance (i.e. learning rates), and **Figure S3** shows the corresponding learning curves across conditions.

**Table S1.** Estimated mean and 95% credible intervals for the effect of iteration on the probability of success across conditions.

|  |  |  |  |
| --- | --- | --- | --- |
| **Condition** | | **Effect of iteration**  **(mean and 95% CI)** | |
| **Estimate** | **Odds ratio** |
| **noVIDEO**  **noAUDIO** | A close up of a logo  Description generated with high confidence | 0.82  [0.08, 1.68] | 2.27  [1.09, 5.39] |
| **noVIDEO**  **AUDIO** |  | 0.84  [0.30, 1.44] | 2.32  [1.35, 4.23] |
| **FACE**  **noAUDIO** | A picture containing clipart  Description generated with high confidence | 0.58  [0.19, 0.98] | 1.79  [1.21, 2.67] |
| **noFACE**  **noAUDIO** | A close up of a logo  Description generated with high confidence | 0.91  [0.52, 1.35] | 2.50  [1.69, 3.84] |
| **FACE**  **AUDIO** | A picture containing clipart  Description generated with very high confidence | 1.49  [1.01, 2.05] | 4.44  [2.75, 7.74] |
| **noFACE**  **AUDIO** |  | 1.81  [1.29, 2.38] | 6.10  3.65, 10.76] |



**Figure S3.** Estimates of the probability of success obtained from the posterior distributions of the fitted models as a function of experimental conditions, iteration, and their interaction. The black lines represent 100 simulations and the red line their mean. Simulations were obtained for an average task and an average participant, i.e. without considering the random intercepts of task and participant. The raw performance data (0: no success, 1: success) is shown as jittered blue dots to avoid overlap.

To test whether the learning rates differed systematically across conditions, we examined the posterior distribution from the fitted model. For each sample of the posterior distribution, we computed the difference in the effects of iteration between different sets of conditions (say, noVIDEO\_noAUDIO and noVIDEO\_AUDIO). A difference is considered systematic if the credible interval does not cross the zero. A positive difference means the first term of the difference has a higher estimate, and a negative difference means the second term has a higher estimate.

**Table S2** shows all comparisons performed and the results of this analysis. Compared to participants who saw the still image with no audio (i.e. who could not benefit from the demonstrator, noVIDEO\_noAUDIO), there was no systematic improvement in the learning rate of the participants who listened to the demonstrator’s speech but with no video (i.e., noVIDEO\_AUDIO) or watched the demonstration videos without the audio (i.e., VIDEO\_noAUDIO); see comparisons 1 and 2 in

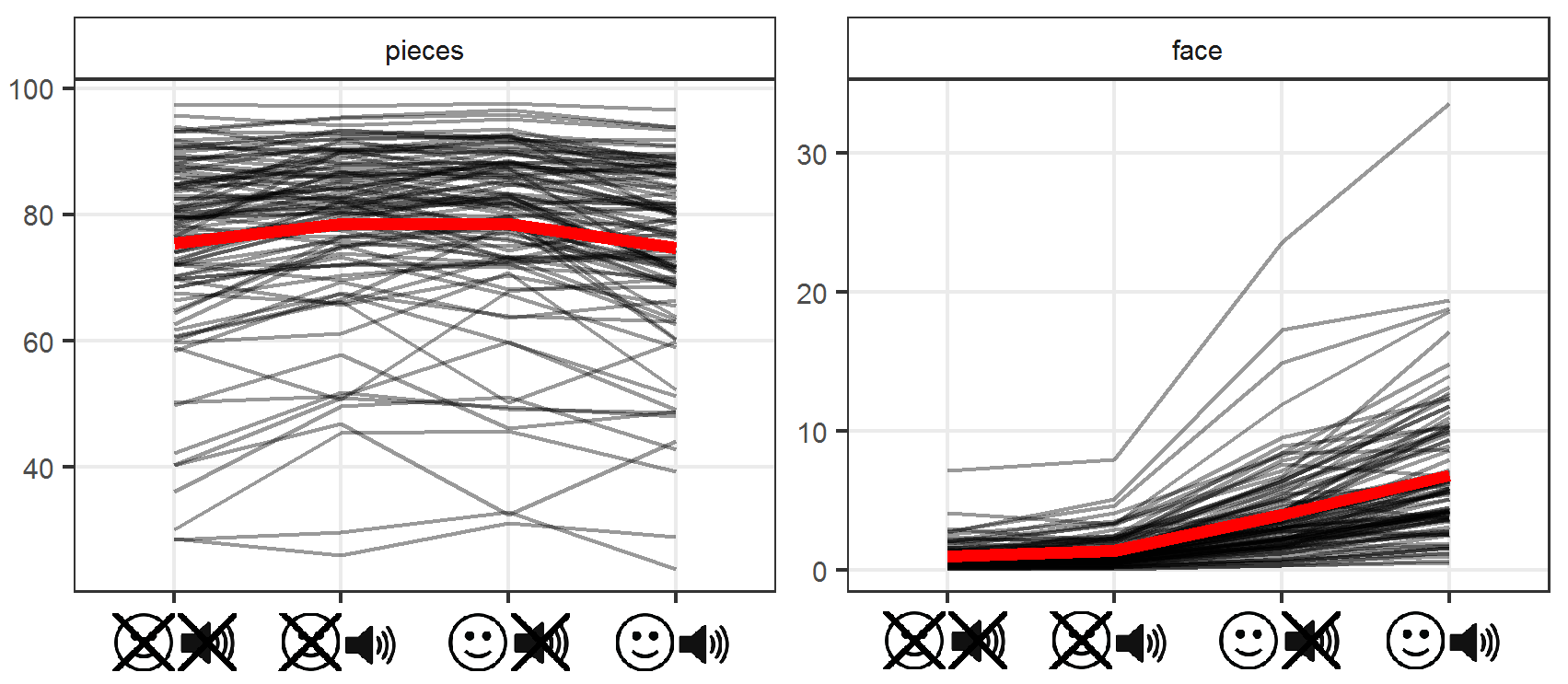
**Table S2** and refer to **Figure S3** for a visualisation. In contrast, participants who watched the demonstration videos *and* listened to his verbal instructions learned faster compared to those who did not watch the videos (see comparisons 3 and 4 in **Table S2**), and to those who watched the videos but without the audio (see comparisons 5 and 6 in **Table S2**). The core result is that learning was facilitated by the presence of the video and the audio simultaneously. Additionally, and in line with the results presented in the main text, there was weak evidence that learning rate was lower for participants who could see the demonstrator’s face compared to face blurred. This is suggested by the model estimating a larger effect of iteration in noFACE\_AUDIO (1.81, 95% CI [1.29, 2.38]) compared to FACE\_AUDIO (1.49, 95% CI [1.01, 2.05]), *p* = .81; and in noFACE\_noAUDIO (0.91, 95%CI [0.52, 1.35]) compared with FACE\_noAUDIO (0.58, 95%CI [0.19, 0.98]), *p* = .88.

**Table S2.** Estimates of the differences on the probability of success between all condition pairs as a function of iteration. Unless indicated, reported values are mean and 95% credible intervals.

|  |  |  |  |
| --- | --- | --- | --- |
| **comparison** | **Conditions being compared** | | **Difference in the effect of iteration (set 2 – set 1)** |
| **set 1** | **set 2** |
| **1** | A close up of a logo  Description generated with high confidence |  | 0.02 [-0.97, 0.93] |
| **2** | A close up of a logo  Description generated with high confidence | A close up of a logo  Description generated with high confidenceA picture containing clipart  Description generated with high confidence | -0.07 [-0.91, 0.79] |
| **3** | A close up of a logo  Description generated with high confidence | A picture containing clipart  Description generated with very high confidence | 0.83, 90% CI [0.10, 1.57] |
| **4** |  | A picture containing clipart  Description generated with very high confidence | 0.81 [0.16, 1.49] |
| **5** | A close up of a logo  Description generated with high confidence |  | 0.89 [0.25, 1.55] |
| **6** | A picture containing clipart  Description generated with high confidence | A picture containing clipart  Description generated with very high confidence | 0.91 [0.28, 1.54] |
| **7** | A close up of a logo  Description generated with high confidence | A picture containing clipart  Description generated with high confidence | -0.34 [-0.91, 0.21] |
| **8** |  | A picture containing clipart  Description generated with very high confidence | -0.32 [-1.07, 0.34] |

# Effects of face visibility on learners’ overt attention.

When learners watched the demonstration videos, they could have attended to two main areas in the display: the demonstrator’s face (visible or blurred) and the puzzle pieces. To test for the effect of face visibility on the allocation of overt attention of our participants, we defined an AOI around his face and AOIs around each of the puzzle pieces, and computed the proportion of time they looked at the face area vs. at the set of pieces. First, we fitted a hierarchical generalised linear model predicting such fixation proportions as a function of the experimental conditions, using a Beta error structure and logit link (because proportions are bounded between 0 and 1). Then, we computed a second model, which included the iteration predictor and its interaction with the experimental conditions. We compared the two models using the Widely Applicable Information Criterion or WAIC (Gelman et al., 2014; McElreath, 2016). As adding iteration did not improve the prediction accuracy of the model, we report the simpler model in what follows. Overall, learners looked at the puzzle pieces more than at the demonstrator’s face (difference in the mean estimates is 74.1%, 90%CI [40.3%, 90.1%], averaged across conditions), but there was no systematic difference across experimental conditions (refer to **Figure S4** for a visualisation). However, learners looked more at the demonstrator’s face when it was visible compared to when it was blurred (difference in the mean estimates between FACE\_noAUDIO and noFACE\_noAUDIO: 3.14%, 95%CI [0.5%, 10.3%]), between FACE\_AUDIO and noFACE\_AUDIO: 5.6%, 95%CI [0.8%, 17.9%]), and even more so when they could listen to his speech (difference between FACE\_AUDIO and FACE\_noAUDIO: 2.9%, 90% CI [0.2%, 8.0%]). This result corroborates our suspicion that the visibility of the demonstrator’s face acted as a distractor because it was not relevant for completing the task at hand, and hence it was detrimental, rather than beneficial, to the learners.



**Figure S4.** Proportion of fixation time participants spent looking at the facilitator’s face and at the puzzle pieces. To indicate uncertainty, reported values are 100 simulations from the fitted models. The average value is shown in red.

# Learning rate and attentional coordination: model comparison

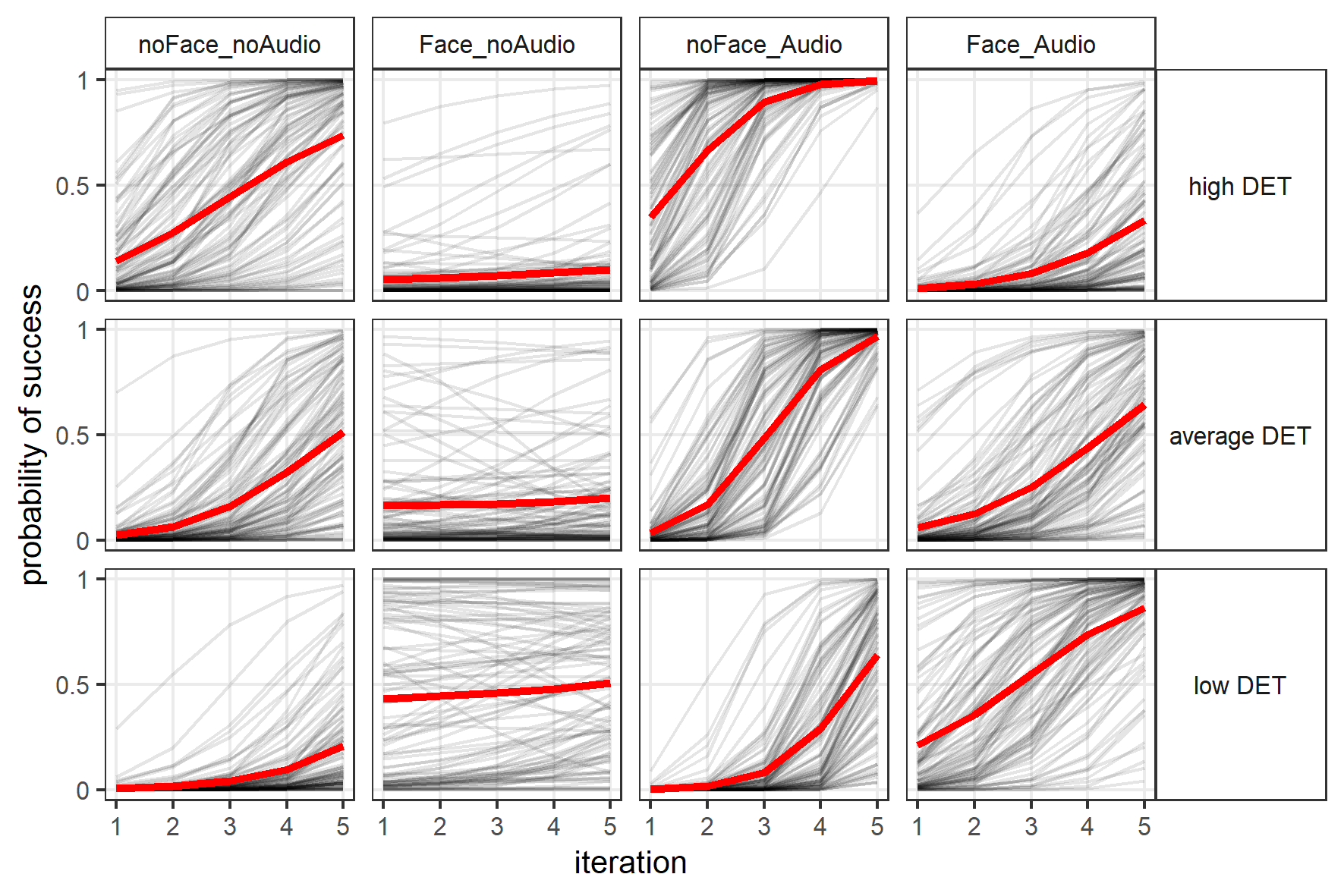
The structure of the regression models is reported in the main text. For each coordination variable, we compared the base model (which does not include the coordination variable) with the two additional models (which include the coordination variable) using the Widely Applicable Information Criterion or WAIC (Gelman et al., 2014; McElreath, 2016). We examined whether adding the coordination variable as a main effect and in interaction with the experimental conditions improved prediction accuracy relative to the base model. Table S3 reports the result of the model comparison.

**Table S3.** Model comparison based on: Widely Applicable Information Criterion (**WAIC**); difference between each WAIC and the lowest WAIC (**dWAIC**); Akaike rescaled WAIC (**weight**); standard error of WAIC estimates (**SE**); standard error of the difference in WAIC between each model and the top-ranked model (**dSE**). Models are ordered from the lowest WAIC (best predictive accuracy) to highest WAIC (worst predictive accuracy). The three final models reported in the main text are in bold font and highlighted in grey.

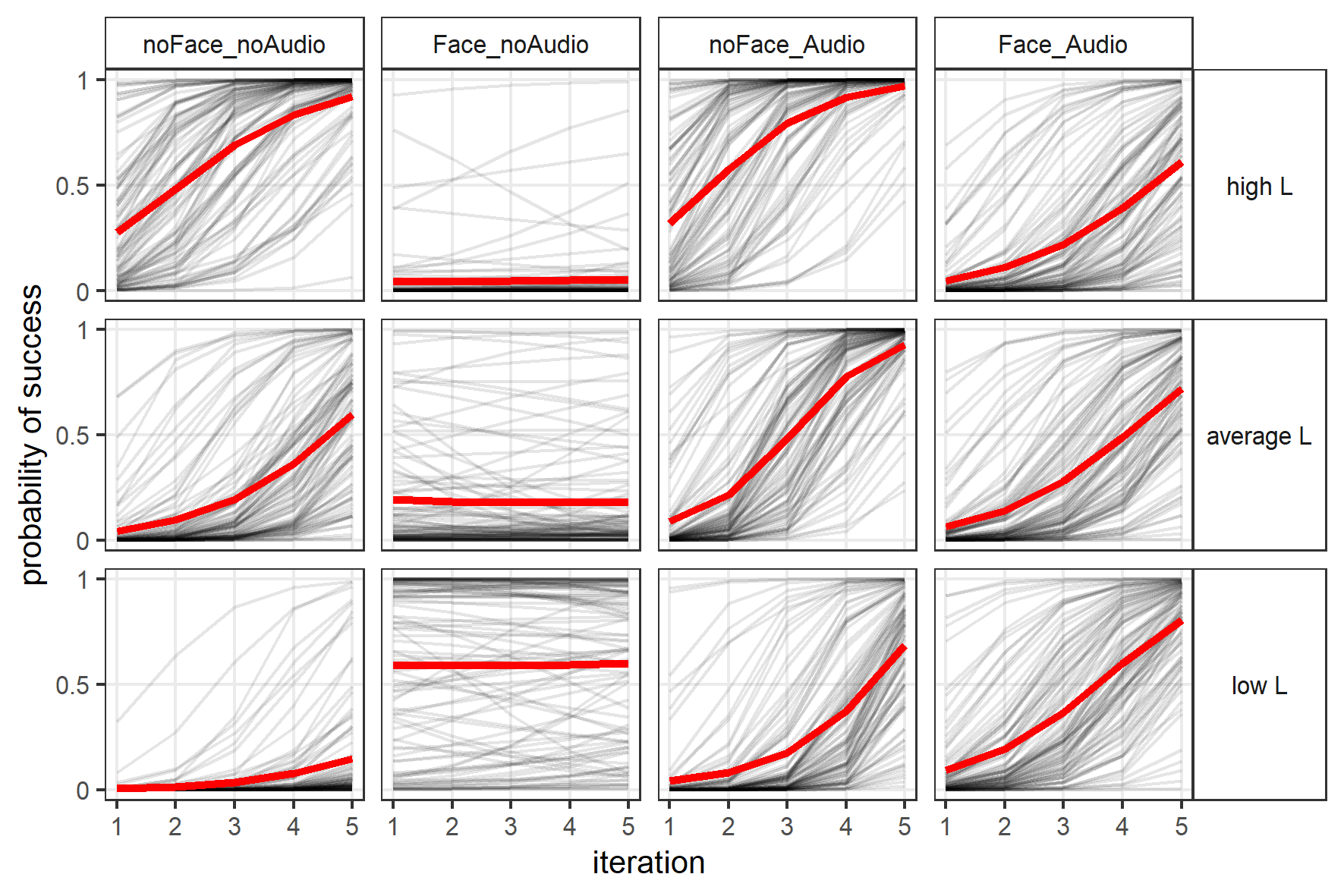
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coordination variable** | **Model** | **WAIC** | **dWAIC** | **weight** | **SE** | **dSE** |
| **RR** | **model 2A** | 319.5 |  | 0.6 | 25.32 |  |
| model 3A | 320.7 | 1.2 | 0.33 | 26.29 | 5.49 |
| base model | 323.9 | 4.4 | 0.07 | 25.27 | 4.27 |
| **DET** | **model 3B** | 307.3 |  | 1 | 25.16 |  |
| base model | 323.9 | 16.6 | 0 | 25.27 | 9.24 |
| model 2B | 325.5 | 18.2 | 0 | 25.55 | 9.33 |
| **L** | **model 3C** | 309.4 |  | 1 | 26.16 |  |
| base model | 323.9 | 14.4 | 0 | 25.27 | 9.22 |
| model 2C | 324.6 | 15.2 | 0 | 25.2 | 9.09 |

Including *RR* as a main effect improved the prediction accuracy relative to the base model but including the interaction between *RR* and the experimental conditions and *RR* did not improve it any further, so we report model 2A in the main text. With respect to *DET* and *L*, instead, including them both as main effect and in interaction with the experimental conditions did improve model accuracy further, so we report model 3B and 3C in the main text.

# Additional figures reporting the final models including DET and L



**Figure S5** Posterior predictions of the final logistic model with the coordination variable DET showing the probability of success (y-axis) as a function of iterations (x-axis) across conditions (columns), while holding RR at either 2 sd below the average (low DET, bottom row), at the average value (average DET, middle row), or at 2 sd above the average (high DET, top row). These simulations are for an average task and average participant. The shaded black lines represent 100 simulations and the thick red lines represent the mean of all simulations within each plot. To see the effect of the different values of DET on performance, the reader should compare the three plots within each column. To see the effect of seeing the demonstrator’s face compared to face blurred, the reader should compare the plots in column 1 with those in column 2, and the plots in column 3 with 4. To see the effect of listening to the demonstrator’s speech compared to no audio, the reader should compare the plots in column 1 with those in column 3, and the plots in column 2 with 4.



**Figure S6** Posterior predictions of the final logistic model with the coordination variable L showing the probability of success (y-axis) as a function of iterations (x-axis) across conditions (columns), while holding RR at either 2 sd below the average (low L, bottom row), at the average value (average L, middle row), or at 2 sd above the average (high L, top row). These simulations are for an average task and average participant. The shaded black lines represent 100 simulations and the thick red lines represent the mean of all simulations within each plot. To see the effect of the different values of L on performance, the reader should compare the three plots within each column. To see the effect of seeing the demonstrator’s face compared to face blurred, the reader should compare the plots in column 1 with those in column 2, and the plots in column 3 with 4. To see the effect of listening to the demonstrator’s speech compared to no audio, the reader should compare the plots in column 1 with those in column 3, and the plots in column 2 with 4.

# Recurrence Quantification Analysis: models including LAM and TT

Following the suggestion of one of the reviewers of our study, we also looked at the RQA variables *LAM* and *TT* as computed from the JRPs. We conducted the same model fitting and selection procedure (based on WAIC) as described in the main text for *DET* and *L*. If we use *LAM*, as the running example, we fitted the following models:

***model\_2****:* *logit(p) = base\_model + b8 \* LAM*

***model\_3:*** *logit(p) = base\_model + (b8+b9\* face+b10\*audio+b11\*face\*audio) \* LAM*

We then compared the base model with the two additional models including LAM (and did the same procedure for *TT*). The final models included *LAM* and *TT* both as main effects and in interaction with the experimental conditions (for *LAM*, model\_3 of the three above). Table S4 reports the effect of *LAM* and *TT* estimated by these final models, which can be compared with the Table S5 which shows the effects of *DET* and *L* estimated by the models originally reported in the manuscript (the effects reported for *DET* and *L* are the same as reported in Table 1 in the main text, copied here for convenience). We also visualise the effects of *LAM* and *TT* in Figure S7, which can be compared with Figure 3 in the manuscript reporting the effects of *DET* and *L*.

We found a positive effect of *LAM* when the face was blurred, and a negative effect when the face was visible, suggesting that seeing the demonstrator’s face, compared to face blurred, was somewhat detrimental to learning. *TT* had instead a positive effect only in the condition noFACE\_noAUDIO. These results are overall similar with what observed with *DET* and *L*, which extends the consistency of our findings (computed from diagonal line structures) to stability measures extracted from vertical line structures.

**Table S4.** Estimated mean values and credible intervals for the relative effects of coordination measures LAM and TT on the probability of success across experimental conditions, obtained from the final models. Values indicating evidence of an effect (i.e. where the credible interval does not cross zero) are in bold to aid reading.

|  |  |  |
| --- | --- | --- |
| **Condition** | **Estimate (mean and credible interval)** | |
| **LAM** | **TT** |
| noFace\_noAudio | **0.95, 90% CI [0.04, 1.89]** | **1.07, 90% CI [0.29, 1.85]** |
| Face\_noAudio | **-1.79, 90% CI [-3.26, -0.07]** | -1.12, 90% CI [-2.26, 0.08] |
| noFace\_Audio | **1.76, 95% CI [0.37, 3.11]** | 0.06, 90% CI [-0.79, 0.97] |
| Face\_Audio | **-1.00, 90% CI [-1.78, -0.15]** | -0.09, 90% CI [-0.97, 0.93] |

**Table S5.** Estimated mean values and credible intervals for the relative effects of coordination measures DET and L on the probability of success across experimental conditions, obtained from the final models as reported in the manuscript. Values indicating evidence of an effect (i.e. where the credible interval does not cross zero) are in bold to aid reading.

|  |  |  |
| --- | --- | --- |
| **Condition** | **Estimate (mean and credible interval)** | |
| **DET** | **L** |
| noFace\_noAudio | **1.14, 95%CI [0.01, 2.29]** | **2.05, 95%CI [0.82, 3.28]** |
| Face\_noAudio | -1.32, 95%CI[-3.03, 0.45] | **-1.82, 90% CI [-3.42, -0.24]** |
| noFace\_Audio | **2.14, 95%CI [0.81, 3.68]** | **1.39, 90% CI [0.02, 2.71]** |
| Face\_Audio | **-1.11, 95%CI [-2.12, -0.17]** | -0.41, 95%CI [-1.11, 0.30] |

A screenshot of a cell phone

Description automatically generated

**Figure S7.** Posterior predictions of the final logistic models showing the probability of success (vertical axis) as a function of coordination (horizontal axis) as captured by the RQA variables (*RR*, top row; *LAM*, middle row; *TT*, bottom row), obtained from JRPs computed by multiplying CRPs for the current and previous iteration, across the four experimental conditions organized along the columns. Coordination variables are standardized (z-scored) with -2 corresponding to 2 SD below the average (low coordination); 0 corresponding to the average value; and 2 corresponding to 2 SD above the average (high coordination). The shaded black lines represent 100 simulations and the thick red lines instead represent the mean of all simulations within each plot.

# Transcriptions of the verbal instructions for each task

**Star puzzle:**

*To complete the star puzzle, you have six identical pieces. Begin by placing two pieces side by side. Balance the second piece on top of the first. The third piece should be placed on opposite sides to create a mirror image. Now carefully lift each segment holding the three pieces together. Rotate one segment slightly so that they fit together, and you are done.*

**Egg puzzle:**

*To complete the egg puzzle, you have two large pieces, two solid-centred pieces, four hollow-centred pieces, and a solid column. Begin by placing a large piece in the palm of your hand, and then balancing the two solid-centred pieces facing each other. Then place the second large piece on the top. Place the puzzle on its end and balance two of the hollow-centred pieces with the hollow facing up. Then turn the puzzle upside down and place the two remaining hollow-centred pieces with the hollow facing up. To complete the puzzle, insert the solid central column; and you are done.*

**Barrel puzzle:**

*To complete the barrel puzzle, you have two large pieces, two C-shaped pieces, two E-shaped pieces, two side pieces, two top pieces, an asymmetric ‘C’, and a bar. To begin, place the C-shaped pieces into the centre of the large pieces and bring the two halves together. Slide one of the E-shaped pieces from the top, turn the puzzle upside down, and slide in the other E-shaped piece. You will have a large hole and a small hole. Grip the side of the small hole and pull. Position one of the side pieces and the asymmetric ‘C’ so that the large section is at the bottom. Place one of the top pieces so that the central bar aligns with the others. Then lock these pieces into place. Position the other side piece and turn the puzzle upside down. Enter the bar half-way and then bring the bar and the remaining piece together, and you are done.*