Figure 3  
Step 1 (Step1TestingRawData20220118.m)

* We depict the raw data points showing the response values as a function of correct values for the inbound turn angle and inbound path length, respectively.

Input: 'PIobjectsYafeiOnly.xlsx', sheet 1, which has 896 rows (28 participants × 4 experiments × 8 paths). Each row has the correct and response locations for three targets.

Output: Figure 3A and 3B in the manuscript. Each figure has all paths (896) by 3 targets = 2688 dots.

Table 1-4

* Averaged Model fitting and validation performance of the 5×2 cross-validation without considering participants’ variable.
* We conduct 5×2 cross-validation modeling using three response locations or home response locations alone for each outbound path.
* Output files save the fitting and validation measures: mseTesting, mseTraining, LogLTesting, LogLTraining, LrTesting, LrTraining, Partial R squareTesting and Partial R square Training.

Step 2 (Step2ModelTraingValidatingMultiObject20220118.m)

Input: 'Split52yafeiOnly.xls', sheets (training1 testing1 etc.). Each sheet has half size (896/2 = 498 rows) of the data of sheet 1 in PIobjectsYafeiOnly.xlsx.

Output: 5 files for using multiple locations, e.g., YafeiOnlyoutput\_obj\_500\_01221.m;

5 files for using home-only locations, e.g., YafeiOnlyoutput\_obj\_500\_021.m;

* We then calculate the averaged results derived from cross-validation based on each model, respectively

Step4getSummary20220121.m

Input: five files for each algorithm (e.g.,YafeiOnlyoutput\_obj\_500\_01221.mat for using three locations)

Output:

Table 1-2: the averaged fitting parameters for each model (e.g., mx1 for the encoding-error model), the models’ performance in fitting (Table 1) and validation (Table 2) [e.g., mTr, m, mlogLTr, mlogL, partRsqTr, partRsq]

e.g., get mse of one fold’s result and calculate rmse and then get averaged value of the two folds from the same cross-validation, and then get the averaged value of the five times cross-validation.

Table 3 (Maximum likelihood ratio (LR): LRs can be directly calculated from the MaxLogLs without adjustment due to difference in parameter numbers.

Table 4 (Alpaydin’s F-test): the difference in the value of RMSE (dRMSE) between them was calculated, generating 5 × 2 difference matrices

Figure 5 (locational residuals), Figure 8 (length and angle residuals) and Table 5 (Lr)

steps 13 and 14

Step13 (step13individualDifferenceGetResiduals20220615.m)

* Get the locational residuals for all 896 paths \* all 3 targets \* 5 cross-validation (show busy figures of ten folds’ results without averaging)

Input Split52yafeiOnlyParticipantID.xls sheets (training1 testing1 etc.) and also the five output files from cross validation (e.g., YafeiOnlyoutput\_obj\_500\_01221.m)

Output residualLocationPID01221.xls for three objects (or residualLocationPID021.xls) sheet=all which has all paths (896) by 3 targets by 5 cross-validations, i.e. 13440 rows. Each row has one target location with correct, response, and model predicted locations. Each row also has distance (location or 2d) residual (difference between response and prediction), x and y residuals.

* We use Step 14 to get a figure showing mean values for all targets. What we did for models with three objects (or single object by changing whichObjects)
* We average the predicted locations across five cross validation for the same target and same path. Then we calculate the residuals for distance, angle, and length based on the averaged predictions of locations in addition to the correct and response locations.
* We chose three objects or a single object by changing whichObjects

Step14 (step14residualAnalysis20220619.m)

Input: residualLocationPID012.xls sheet=all

Output residualLocationPID012.xls sheet=AverageFiveValidation which has 13440/5=2688 rows. In each row we have residuals for distance, angle, and length for each target and path (3\*896=2688).

* We calculated the mean predicted locations of each target in each outbound path averaging across participants (96=4 exp\* 8path\*3targets for multiple locations and 32 for home-only). (2688/28=96) targetRedidual is used to visualize the location, length and angle residusals and the relationship between response errors and prediction errors.

Input residualLocationPID012.xls sheet= AverageFiveValidation

Output residualLocationPID012.xls sheet= targetRedidual which has 96 rows

Save: locresidualLocationPID01221.fig (Figure 5A) and locresidualLocationPID221.fig (Figure 5B) (locational residual = mean predicted location – mean response location)

DresidualLocationPID01221.fig (Figure 8A) and DresidualLocationPID021.fig (Figure 8B)

AresidualLocationPID01221.fig (Figure 8C) and AresidualLocationPID021.fig (Figure 8D).

* The likelihood ratios (Table 5) were computed to compare the models’ performance in predicting inbound length errors and angle errors using r-squared values (showing in Figure 8).

Figure 6 (Confusion matrices in model recovery with fixed parameters)

Step2SimCVMultipTimes20220124.m + Step12 (getSummarySimCVMultipTimes20220627.m + confusionMatrix.m)

* For each model (i.e., three true models), using the fitting parameters (with two algorithms), generating 100 sets of simulated response data (6 folders in total). Adding random noise for every trial (one fixed random sheet for all folds’ stimulated data).

Step2SimCVMultipTimes20220124.m

Input: Split52yafeiOnly.xls (sheets of training1 testing1 etc.) + true parameters and rmse from Step 2.

Output: 100 files of simulated data in each folder

(e.g., SummarySimParaoutput\_M302100.mat in folder ‘m3homeonly’)

* For each true model, we calculated the likelihood ratio between any two models for each of the 100 simulations and classified the likelihood ratios into different categories (Figure S2).
* The best model was determined only when it had likelihood three times higher than both other two models. Figure 6 indicates the frequency of the recovered model being the best model.

getSummarySimCVMultipTimes20220627.m

Input: the summary Array from 100 output files

Output: confustionMatrix.xlsx

* Using the frequency listed in confustionMatrix.xlsx to draw the confusion matrix.

confusionMatrix.m

Input: confustionMatrix.xlsx (sheet of “fixed multiple targets” or “fixed home only”)

Output: “confustionMatrixfixed multiple targets.fig” and “confustionMatrixfixed home only.fig”

Figure 7

Step 12 (distanceParameters.m)

* We calculated the distance of 16 parameters estimated parameters based on real data and simulated data.

Input: confustionMatrix.xlsx (sheet of “fixed multiple targets” or “fixed home only”)

Output: similiarityofParaFixed.fig

Figure 9

Step14 (step14individualDifferenceAnalysis20220619.m)

* We get mean residuals (of 24 targets, 8path\*3targets) for each participant calculating from residualLocationPID012.xls sheet= AverageFiveValidation.
* We get regressions between angles (e.g., response and correct, response functions) and between lengths (e.g., response and correct, response functions) (8 points or 24 points)

Input residualLocationPID012.xls sheet= AverageFiveValidation

Ouput residualLocationPID012.xls sheet= subCorrAverageFive which has 112 rows

Save: regressionLineresidualLocationPID01221.fig (Figures 9A and 9B)

* We visualized the intercept and slope of the response functions (both angle and length). We gave participants category labels (compressionA=1 or 0, compression=1 or 0). 1 and 2 in the latter is separated by the median of slope for the participants with compressionA=1.

(We save subjectRedidual sheet to subjectRedidual.cvs and then added these labels to this file, which will be input for JASP analyses. We manully type the gender information into the sheet= subCorrAverageFive)

Input: residualLocationPID012.xls sheet= subCorrAverageFive

Output: InterSlopeColorPvaluesresidualLocationPID021.fig (Figures 9C and 9D)

Tables 6, 7, and 8

Step 15 (step15separateParticipants20220619.m) + Step2ModelTraingValidatingMultiObject20220613.m (in the folder of “step 16/YesA or /NoA”) + Step4getSummary20220614.m

* Step 15 was used to separate participants into different files (i.e., ‘SplitExp2YafeiNOCompressA.xls’ and ‘SplitExp2YafeiYESCompressA.xls’ ), each with 5by2 sheets. These files can be saved to different folders to do step 2 cross-validation and each group generate 5 output files using two different algorithms (e.g., Exp2Yafeioutput\_obj\_500\_01225.mat).
* Then, step 4 can get the summarized results of cross-validation for each group (Tables 6, 7, and 8).

Step 15 (step15separateParticipants20220619.m)

Input: 'PIobjectsYafeiOnly.xlsx', sheet 1, which has 896 rows, and 'subCategoriesWithGender.xlsx', sheet 'subCorrAverageFive', which has 112 rows corresponding to 112 participants.

('subCategoriesWithGender.xlsx' = sheet ‘subCorrAverageFive’ in residualLocationPID01221.xls + Columns ‘Compress A’ (0/1), ‘Compress D’ (0/1), ‘gender’, etc)

Output: 'PIobjectsYafeiOnlyWithCategories.xlsx', which separates all 896 rows based on different categorization criteria (e.g., sheets ‘noCompressA’ and ‘yesCompressA’). Each sheet contains a part of participants who belonging to that group.

Input: 'PIobjectsYafeiOnlyWithCategories.xlsx'

Output: 'SplitExp2YafeiNOCompressA.xls' and 'SplitExp2YafeiYESCompressA.xls'

* For each group (Yes and No compress group)

Step2ModelTraingValidatingMultiObject20220613.m

Input: 'SplitExp2YafeiYESCompressA.xls' or ‘SplitExp2YafeiNOCompressA.xls’ (e.g., sheets “Training1”)

Output: 5 files for multiple locations, e.g., Exp2Yafeioutput\_obj\_500\_021.mat and 5 files for home locations only, e.g., Exp2Yafeioutput\_obj\_500\_01221.mat

Step4getSummary20220614.m

Input: 5 files for multiple locations, e.g., Exp2Yafeioutput\_obj\_500\_021.mat and 5 files for home locations only, e.g., Exp2Yafeioutput\_obj\_500\_01221.mat

Output: yA01221.mat and yA021.mat for “YesA” group.

nA01221.mat and nA021.mat for “NoA” group.

Tables 9, 10, and 11

Step 16 (step16residualAnalysisDifferentgroup20200619.m)

* Step 16 is just like Step 14 (step14residualAnalysis20220619.m) except for adding a procedure at the beginning to read the “all” sheets from two groups and combine them.
* We got the overall performance by combining the locational residuals of the two compression groups using step 16.

Input: the 'residualLocationPID01221.xls', sheet “all” from each “YesA” folder and “NoA” folder if we used multiple locations; the 'residualLocationPID021.xls', sheet ”all” from each “YesA” folder and “NoA” folder if we used home locations only;

Output: locresidualLocationPID01221.mat for using multiple locations; locresidualLocationPID021.mat for using home locations only in /step16 /results/

* We got the averaged parameters by weighted combining the parameters from the two groups using step 17 (step17getSummary20220619.m)

Input: "NoA/nA021.mat", "YesA/yA021.mat" (using home locations only)

"NoA/nA01221.mat", "YesA/yA01221.mat" (using multiple locations)

Output: results\CrossValidation0.mat and CrossValidation012.mat

“meanx1”, “meanx2”, and “meanx3” are the averaged parameters in Table 9.

Figure 10

(step16residualAnalysisDifferentgroup20200619.m)

* Sheet “targetRedidual” has 96 rows and saves mean residuals for targets (96 targets,4 exp\* 8path\*3targets) averaging across participants (28). (2688/28=96). Sheet “targetRedidual” is used to visualize the location, length, and angle residusals and the relationship between response errors and prediction errors.

Input: the 'residualLocationPID01221.xls', sheet “TargetRedidual” from each “YesA” folder and “NoA” folder if we used multiple locations; the 'residualLocationPID021.xls', sheet ” TargetRedidual” from each “YesA” folder and “NoA” folder if we used home locations only;

Output: “locresidualLocationPID01221.fig” (Figure 10A) “locresidualLocationPID021.fig” (Figure 10B) in /step16 /results/

Figure 11 (Confusion matrices in model recovery with varied parameters)

Step12 (Step12SimCVMultipTimes20220604.m + getSummarySimCVMultipTimes20220627.m + confusionMatrix.m)

* The code (Step12SimCVMultipTimes20220604.m) is based on Step2SimCVMultipTimes20220124.m by adding variable parameter values for each model. For each parameter, we can set a uniform distribution with a range interval of factor\*(1-fixed value) and center of the fixed value.
* For example, the range of the uniform distribution for intercept parameters was twice the distance between the mean intercept and 0.
* For each model (i.e., three true models), using varied parameters (with two algorithms), generating 100 sets of simulated response data (6 folders in total). Adding random noise for every trial (one fixed random sheet for all folds’ stimulated data).

Step12SimCVMultipTimes20220604.m

Input: Split52yafeiOnlyParticipantID.xls (sheets of training1 testing1 etc.) + true parameters and rmse from Step 2.

Output: 100 files of simulated data in each folder

(e.g., SummarySimParaoutput\_M302100.mat in folder ‘m3homeonly’)

* For each true model, we calculated the likelihood ratio between any two models for each of the 100 simulations and classified the likelihood ratios into different categories (Figure S3).
* The best model was determined only when it had likelihood three times higher than both other two models. Figure 11 indicates the frequency of the recovered model being the best model.

getSummarySimCVMultipTimes20220627.m

Input: the summary Array from 100 output files

Output: confustionMatrix.xlsx

* Using the frequency listed in confustionMatrix.xlsx to draw the confusion matrix.

confusionMatrix.m

Input: confustionMatrix.xlsx (sheet of “varied multiple targets” or “varied home only”)

Output: “confustionMatrixvaried multiple targets.fig” and “confustionMatrixvaried home only.fig”

Figure 12

Step 12 (distanceParameters.m)

* We calculated the distance of 16 parameters estimated parameters based on real data and simulated data.

Input: confustionMatrix.xlsx (sheet of “varied multiple targets” or “varied home only”)

Output: similiarityofParaVaried.fig

Figure 13 and Table 12

(step16residualAnalysisDifferentgroup20200619.m)

* Step 16 is just like Step 14 (step14residualAnalysis20220619.m) except for adding a procedure at the beginning to read the “all” sheets ('residualLocationPID01221.xls' or 'residualLocationPID021.xls') from two groups and combine them.
* We got the overall performance by combining the two compression groups using step 16.

Input: the 'residualLocationPID01221.xls', sheet “TargetRedidual” from each “YesA” folder and “NoA” folder if we used multiple locations; the 'residualLocationPID021.xls', sheet ” TargetRedidual” from each “YesA” folder and “NoA” folder if we used home locations only;

* Sheet “targetRedidual” has 96 rows and saves mean residuals for targets (96 targets,4 exp\* 8path\*3targets) averaging across participants (28). (2688/28=96). Sheet “targetRedidual” is used to visualize the location, length, and angle residusals and the relationship between response errors and prediction errors.

Output: DresidualLocationPID01221.fig (Figure 13A) and DresidualLocationPID021.fig (Figure 13B)

AresidualLocationPID01221.fig (Figure 13C) and AresidualLocationPID021.fig (Figure 13D).

* The likelihood ratios (Table 12) were computed to compare the models’ performance in predicting inbound length errors and angle errors using r-squared values (showing in Figure 13).