Supplementary Materials

Contents

1 Experiment Design

We describe important details about the experiment in this section, but the full experiment may be found in the GitHub repository. Although the full code for the experiment is available, we also provide a PDF of screenshots in `figures/screenshots.pdf`. Note that the screenshots were generated for one possible participant and what might be shown at the start of each screen. The screenshots were taken within a development environment and contain artifacts that would not have appeared for actual participants, such as the accumulated compensation always being $0.00, the screen number, and "You solved out of 89 puzzles" on the final screen which should contain an actual number of puzzles solved. Some screens had short pop-up messages during the tutorial in response to participants' actions that are not depicted in the screenshots. Screens with the blue `Submit` button had a required task and most would not allow the participant to proceed unless they provided valid responses. Although the PDF shows screenshots for every screen in the experiment, not all screens were shown to all participants depending on their responses. We describe the control flow below.

1. Screen 2 (diagnostic survey): if the participant either correctly solved the diagnostic puzzle in Screen 1 and/or did not respond with "None" to the question "About how many Sudoku puzzles have you successfully completed?", skip to Screen 120.
2. Screen 5 (agreement): participants were allowed to proceed if their responses matched the prompt by at least 85%.
3. Screens 6-14 (tutorial): participants could only proceed after submitting correct responses (where relevant). Hints were provided through short pop-up messages if they submitted incorrect responses.
4. Screens 18-42 (practice phase): participants could only proceed after inputting the correct digit in the goal cells.
5. Screens 44-107 (test phase): if participants correctly solved a puzzle, they were allowed up to 10 seconds before automatically proceeding to the next screen (this period may be skipped). If they submitted an incorrect response (including a blank) or did not submit anything for 2 minutes, they were required to wait 10 seconds before automatically proceeding to the next screen.
6. Screen 108 (questionnaire attention check): participants could only proceed after answering all 3 questions, regardless of correctness.
7. Screens 109-119 (questionnaire): see Section 1.4

1.1 Diagnostic Test and Survey

At the beginning of the experiment, in order to filter out any participants who had prior experience in solving Sudoku puzzles, we presented a simple diagnostic test and a survey. Participants were not informed about the purpose of the diagnostic material.

The diagnostic puzzle was a 4x4 Sudoku grid as shown in **Figure 1**. Participants were told that the puzzle was a 4x4 variant of Sudoku and were instructed in how to interact with the program interface. They were offered $0.25 to complete the puzzle with a $0.01 penalty for each incorrect attempt. All cells needed to be correct for the puzzle to be considered solved.

After the diagnostic puzzle, participants were presented with the following survey:

1. Have you heard of Sudoku before?
   (a) Yes
   (b) No
   (c) Not sure
2. Have you ever attempted to solve a Sudoku puzzle?
   (a) Yes
   (b) No
   (c) Not sure
3. About how many Sudoku puzzles have you successfully completed?

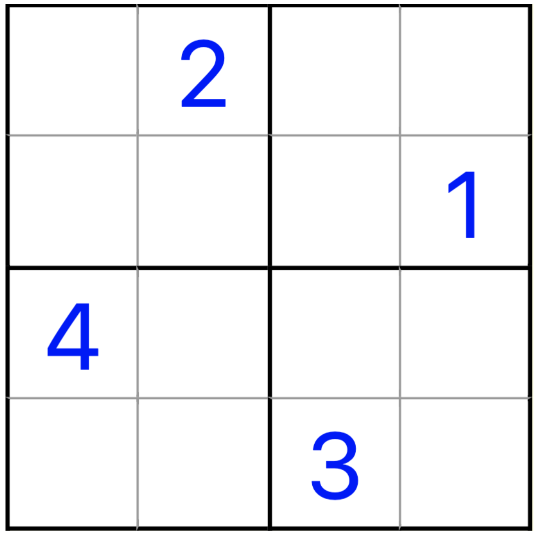


Figure 1: Diagnostic Puzzle

(a) None
(b) 1 to 3
(c) 4 to 6
(d) 7 to 9
(e) 10 or more

Only participants that did not solve the diagnostic puzzle and also responded that they had never successfully completed any Sudoku puzzles proceeded to the rest of the experiment. For all others, the program skipped to the demographics survey and terminated thereafter. Of the 1983 people that originally entered the study, only 600 had not solved the diagnostic puzzle. 327 of these had solved at least one Sudoku puzzle, leaving 273 participants in the experiment.

1.2 Tutorial

At the start of the experiment, participants were given a brief description of Sudoku with the following sentence: "Sudoku is a puzzle with a 9x9 grid of numbers where each row, column, and 3x3 box must contain exactly one of each number from 1 to 9." Beyond this initial statement, we took as much care as we could to avoid terms referring to variables and roles. For exact phrasing used in the experiment, please refer to the experiment screenshots on the Online Supplementary Materials.

As noted in the main text, each participant was randomly assigned a house type (row or column), a house index (between 1 and 9), a cell index (between 1 and 9), and 2 disjoint sets of 4 digits (all between 1 and 9) which would define the parameters of the tutorial puzzles, practice phase puzzles, and the control conditions in the test phase. Using these features, we randomly generated a Hidden Singles puzzle to use during the tutorial (see the example for one participant in **Figure 2a)** following the same generative procedure as puzzles in the practice phase, test phase, and questionnaire. (see Methods in main manuscript). The tutorial puzzles always had one highlighted row or column based on the assigned house type and house index, and one target cell based on the cell index.

Based on this puzzle, we provided two preliminary exercises to help solidify the general concepts of Sudoku before proceeding to the tutorial itself. First, each participant saw the row or column corresponding to their assigned house type and index, with two copies of a digit, one of which occupied the assigned goal cell (Figure 2b). A participant seeing this particular assignment would then read that the column "contains two copies of the digit 1, forming a contradiction" and be asked to select both. Next, the participant would be shown the same grid with the digit in the goal cell removed (Figure 2c) with the statement "one of the 1's forming the contradiction has been removed. Fill in the missing number in the green cell so that the column contains every number from 1 to 9," in this case a 3. Incorrect responses would invoke feedback noting that the participant's response

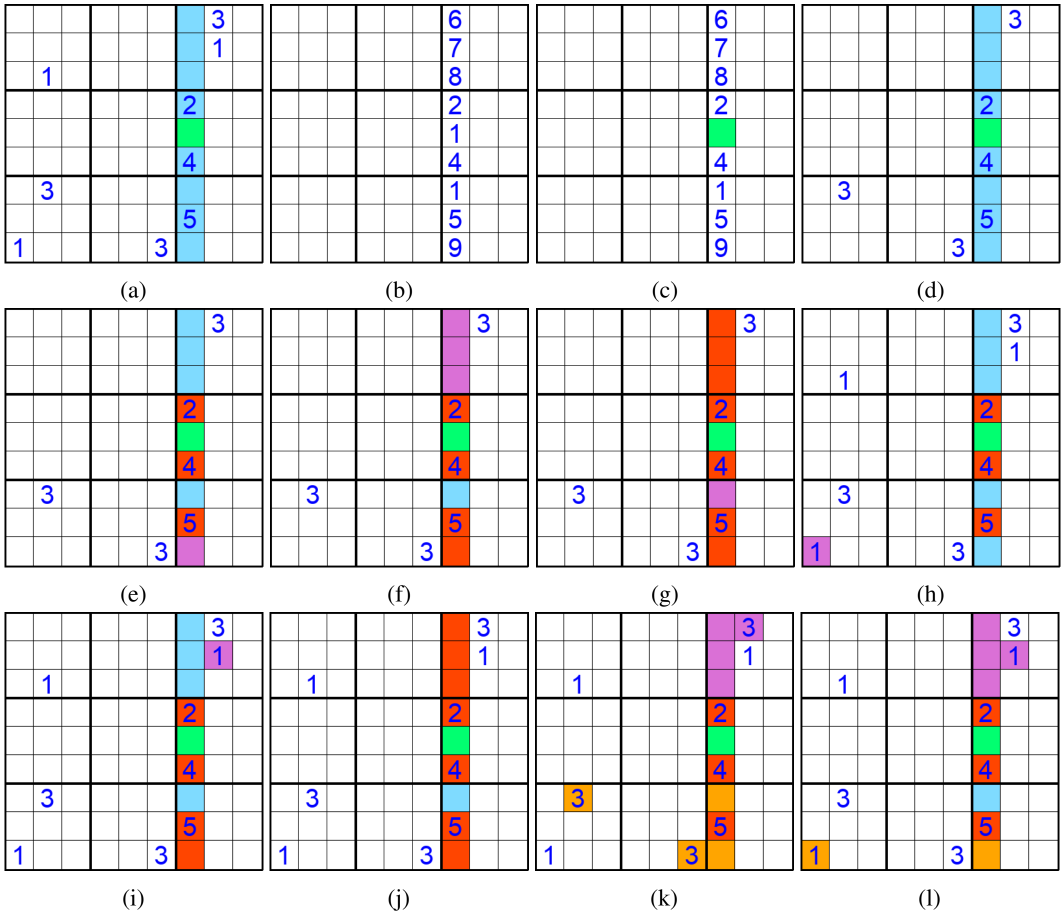


Figure 2: (a) The tutorial puzzle randomly generated for this participant. (b) The contradiction exercise. The goal is to select the two cells containing 1s. (c) The Full House exercise. The goal is to fill in the green cell with 3. (d) The tutorial puzzle with distractors removed. (e-g) Positive example tasks. (h-j) Negative example tasks. (k-l) Summary grids for both positive and negative examples.

already exists in the row or column. Note that the digit used in the first exercise (1) and the digit used in the second exercise (3) were both sampled from the participant's training digit set and were subsequently used as the distractor and target digits for the remainder of the tutorial.

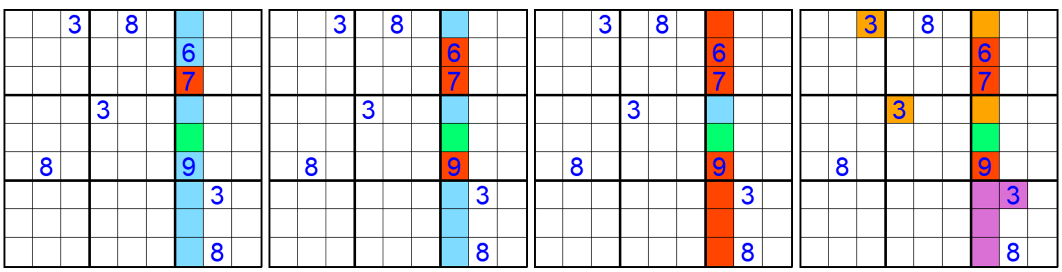
1.2.1 Hidden Singles Tutorial

The tutorial on the Hidden Single technique began with a screen with only three digits in the blue column and three copies of the digit 3 outside this column, positioned to prevent a 3 from occurring in any cell other than the green goal cell, as in Figure 2d. The participant was asked to focus on the green cell and solve for its value, with the statement "since the blue column must contain exactly one 3, let's see if we can determine if the only place a 3 can go is in the green cell." Across a series of screens, cells that could not contain a 3 were ruled out and highlighted in various colors as shown in Figures 2c-g. The tutorial walked the participants through solving the puzzle by identifying the clues on the grid that constrained each of the blue cells. As shown in **Figure 2e**, the participant would be asked to find the instance of the target digit (in this case 3) that constrains the purple cell, for which the correct answer would be the 3 directly left of the purple cell. They would continue this until all blue cells were eliminated (progressively highlighted red) and it was visually apparent that the only cell the target digit could be in is the goal cell, highlighted in green. At this point, the participant would be presented with the statements "We have successfully eliminated every cell in the blue row except the green cell as potential candidates for 3. Fill in the green cell with the correct digit to solve this puzzle." Here, they would need to fill in the goal cell with the target digit to proceed.

Next, three 1's were added to the grid, i.e. the distractors, positioned to prevent a 1 from occurring in four of the empty blue cells but leaving both the green goal cell and one of the blue cells as possible placements for this digit, as in Figure 2a. This time, participants were asked to identify the cells that each clue constrains. For example, in **Figure 2i**, they would need to select all 3 blue cells in the 3x3 box that the purple cell is in. This provided an alternative method of approaching the puzzle, starting with the clues to eliminate blue cells instead of checking if a blue cell could be eliminated as a possible location for a digit. After inevitably failing to solve for the distractor (one blue cell could not be eliminated), participants were shown color-coded grids for both the target (**Figure 2k**) and distractor (**Figure 2l**) side-by-side, indicating the types of constraints on the grid and showing how all blue cells could be eliminated for the target but not for the distractor. Thus, by example, participants were shown how to select one of the two digits occurring three times in the puzzle as the correct digit to place in the goal cell, and to rule out the other of these digits as the correct answer.

1.3 Practice Phase

During the practice phase, participants were given detailed feedback whenever they gave incorrect responses (Figure 3). The feedback was customized to each puzzle and the type of response. Thus, if the participant gave an in-house response, only the cell containing their input digit would be highlighted red. If the response was an absent digit, all the cells containing digits would be highlighted red. If the response was a distractor, all blue cells would be made red except the cell that is unconstrained by any of the distractor clues. Once the participant corrected an incorrect response, all cells in the house and cells containing target clues were highlighted with different colors to indicate the different constraints. No feedback was given for correct first responses except that they were correct.



(a) (b) (c) (d)

Figure 3: Sample feedback sequences during practice phase. (a) Feedback for submitting 7, an in-house response. (b) Feedback for submitting 4, an absent response. (c) Feedback for submitting 8, a distractor response. (d) Feedback for a correcting a previously incorrect response.

The following text would accompany the feedback.

- In-house feedback (Figure 3a): 7 cannot be at the[1] because 7 already exists in the same column in the red cell.
- Absent feedback (Figure 3b): It is not certain that 4 must be at the green cell because 4 may potentially be in a blue cell. The red cells cannot be 4 because they already contain digits.
- Distractor feedback (Figure 3c): It is not certain that 8 must be at the green cell because 8 may potentially be in the blue cell. Note that neither the green cell nor the blue cell share the same row, column, or box with a 8.
- Target feedback (Figure 3d): 2 is correct! We can be certain that the green cell must contain 2 because no other cell in its column can be a 2. The red cells cannot be 2 because they already contain digits. The empty purple cells cannot be 2 because they share they share the same box with a 2. The empty orange cells cannot be 2 because they share rows with other 2s.

[1]This typo was made in the actual experiment

1.4 Strategy Survey

After completing the text phase described in the main text, participants were first instructed on the contents of the following segment and asked 3 attention-check questions. Next, they were given a puzzle sharing the same house type and digit set as the tutorial, but with the goal cell located in the center box and were asked to select and enter the digit that must go in the goal cell. Without providing feedback on correctness, we then displayed the puzzle and their response for the remainder of the questionnaire, allowing the participants to refer to it as necessary. We asked the following questions in order of increasing specificity using free-response questions to elucidate general responses without additional prompting and more specific multiple-choice questions to formalize their strategies.

1. "How confident do you feel that your answer is correct, expressed as a percentage?" (Responses were allowed between 0 and 100 in increments of 5.)
2. "Explain as clearly as possible the steps you went through to choose your answer. Please be as detailed as possible so that someone else could replicate your strategy by following your response."
3. "There are two numbers in the puzzle that occur three times outside of the row/column containing the target cell. Which of the following best describes how you chose between the two candidate numbers to consider?"
   (a) I noticed something in the puzzle that initially made one candidate seem more likely to be correct than the other.
   (b) I arbitrarily chose between the two candidates because they seemed equally promising to consider.
4. "What did you notice in the puzzle that initially made one candidate seem more likely to be correct than the other?"
5. "Please select the cell(s) that initially made one candidate seem more likely be correct than the other." (Participants could select one or more cells in the grid.)
6. "Please explain how the cell(s) you selected initially made one seem more likely to be correct than the other." (Previously selected cells were shown.)
7. "After you selected a candidate to consider, did you check further to determine whether that candidate was actually correct or not?"
   (a) Yes, I checked to see whether the candidate was actually correct.
   (b) No, I just submitted my original guess without checking any further.
8. "What did you do to determine if that candidate was actually correct?"
9. "Which of the following best describes the way you determined whether or not the candidate was actually the correct answer?"
   (a) I checked whether the candidate I chose could go in any of the empty blue cells in the row/column.
   (b) I looked at other numbers in the puzzle until I noticed something that helped me decide whether or not the candidate was correct.
10. "Please provide any additional information or clarifications to any of your previous responses so that we can most accurately understand as best we can how you solved this puzzle."

Participants that did not respond to the puzzle with the target or distractor skipped ahead to Question 10 after Question 2. Participants that responded with option (b) in Question 3 skipped ahead to Question 10. Participants that responded with option (b) in Question 7 skipped ahead to Question 10.

1.5 Demographics Survey

Following the strategy questionnaire, we also asked the participants about their demographic information including their age, gender, highest level of education, and mathematical topics they have taken courses in. All participants, regardless of diagnostic test and survey results, were asked the following questions about their educational backgrounds. Participants that were filtered out from the diagnostics were presented the demographics survey immediately after the diagnostic survey screen.

"What is your highest level of education (including currently pursuing)?" Participants were allowed to select one of the following:

- Have not graduated high school
- High school graduate, diploma or equivalent
- Associate degree
- Bachelor's degree
- Master's degree
- Professional degree (e.g. M.D., J.D.)
- Doctoral degree

"Degree status" Participants were allowed to select one of the following:

- Currently pursuing
- Completed

"Which of the following mathematics topics have you taken a course in? Select all that apply." Participants were allowed to select zero or more of the following:

- High school algebra
- High school geometry
- Trigonometric functions
- Single-variable calculus
- Multi-variable calculus
- Linear algebra
- Probability & statistics
- Discrete mathematics
- Formal logic

2 Reported Regressions

Here, we provide the exact formulas used in the regressions reported in the main text and the full list of fitted coefficient values.

In this section, t refers to the trial number (between 1 and 64) and s refers to the subject. The $(1 + \log_2 t|s)$ included in each regression indicates participant-level random effects. All other terms are fixed effects. Response time models were fitted using only correct trials.

2.1 Overall Accuracy

We used logistic mixed effects models for predicting the correctness of each trial for both phases.

$$P(\text{correct}_{t,s}) \sim \log_2 t + (1 + \log_2 t|s)$$

Table 1: Accuracy regression coefficient estimates and 95% highest density intervals. Reported numbers in logits.

Phase	Term	Estimate	HDI-L	HDI-U
Practice	intercept	-0.493	-0.702	-0.280
Practice	$\log_2$(trial)	0.344	0.287	0.404
Test	intercept	0.453	0.266	0.645
Test	$\log_2$(trial)	0.147	0.108	0.186

2.2 Digit Sets (DS)

Accuracy models

$$P(\text{correct}_{t,s}) \sim DS + \log_2 t + (1 + \log_2 t|s)$$

Table 2: Digit set: accuracy regression coefficient estimates and 95% highest density intervals. Reported numbers in logits.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	2.405	1.745	3.182	2.903	2.334	3.537
$\log_2$(trial)	0.086	-0.142	0.308	0.091	-0.035	0.217
DS	-0.098	-0.472	0.268	0.020	-0.248	0.286

Duration models

$$\log(\text{duration}_{t,s}) \sim DS + \log_2 t + (1 + \log_2 t|s)$$

Table 3: Digit set: duration regression coefficient estimates and 95% highest density intervals. Reported numbers in $\log_2$ seconds.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	4.453	4.293	4.611	3.945	3.777	4.116
$\log_2$(trial)	-0.126	-0.168	-0.084	-0.076	-0.100	-0.053
DS	-0.022	-0.086	0.043	0.000	-0.032	0.032

2.3 Goal Position (GP)

Accuracy models

$$P(\text{correct}_{t,s}) \sim GP + \log_2 t + (1 + \log_2 t|s)$$

Table 4: Goal position: accuracy regression coefficient estimates and 95% highest density intervals. Reported numbers in logits.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	2.223	1.525	3.042	3.151	2.539	3.829
$\log_2$(trial)	0.089	-0.143	0.305	0.092	-0.030	0.209
GP	0.165	-0.240	0.570	-0.314	-0.668	0.019

Duration models

$$\log(\text{duration}_{t,s}) \sim GP + \log_2 t + (1 + \log_2 t|s)$$

Table 5: Goal position: duration regression coefficient estimates and 95% highest density intervals. Reported numbers in $\log_2$ seconds.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	4.354	4.198	4.515	3.877	3.708	4.055
$\log_2$(trial)	-0.126	-0.168	-0.086	-0.076	-0.099	-0.053
GP	0.115	0.040	0.189	0.089	0.050	0.128

2.4 House Type (HT)

Accuracy models

$$P(\text{correct}_{t,s}) \sim HT + \log_2 t + (1 + \log_2 t|s)$$

Table 6: House type: accuracy regression coefficient estimates and 95% highest density intervals. Reported numbers in logits.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	2.508	1.842	3.308	2.867	2.286	3.497
$\log_2$(trial)	0.089	-0.138	0.312	0.089	-0.033	0.215
HT	-0.302	-0.666	0.059	0.095	-0.170	0.364

Duration models

$$\log(\text{duration}_{t,s}) \sim HT + \log_2 t + (1 + \log_2 t|s)$$

Table 7: House type: duration regression coefficient estimates and 95% highest density intervals. Reported numbers in $\log_2$ seconds.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	4.292	4.136	4.447	3.911	3.746	4.082
$\log_2$(trial)	-0.128	-0.168	-0.087	-0.076	-0.098	-0.053
GP	0.319	0.258	0.379	0.055	0.022	0.088

2.4.1 Accounting for Tutorial House Type

We checked to see if the effects were moderated by whether the participants were taught the Hidden Single technique using rows or column. Specifically, we added a new term C indicating whether the participant's tutorial used columns, yielding the following.

Accuracy models

$$P(\text{correct}_{t,s}) \sim HT + C + HT * C + \log_2 t + (1 + \log_2 t|s)$$

Table 8: House type accounting for tutorial house type: accuracy regression coefficient estimates and 95% highest density intervals. Reported numbers in logits.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	2.688	1.937	3.490	2.969	2.334	3.652
$\log_2$(trial)	0.104	-0.111	0.321	0.092	-0.027	0.215
HT	-0.405	-0.918	0.114	0.018	-0.367	0.407
C	-0.419	-1.169	0.353	-0.207	-0.857	0.453
HT*C	-0.183	-0.528	0.876	0.151	-0.391	0.680

Duration models

$$\log(\text{duration}_{t,s}) \sim HT + C + HT * C + \log_2 t + (1 + \log_2 t|s)$$

Table 9: House type accounting for tutorial house type: duration regression coefficient estimates and 95% highest density intervals. Reported numbers in $\log_2$ seconds.

	Trials 1-16			Trials 17-64		
Term	Estimate	HDI-L	HDI-U	Estimate	HDI-L	HDI-U
intercept	4.318	4.119	4.523	4.014	3.798	4.225
$\log_2$(trial)	-0.128	-0.169	0.-0.089	-0.076	-0.100	-0.053
HT	0.331	0.246	0.421	0.041	-0.003	0.083
C	-0.056	-0.316	0.206	-0.205	-0.463	0.062
HT*C	-0.023	-0.150	0.101	0.028	-0.036	0.090

2.5 Education

In the questionnaire, we asked for highest education pursued, whether completed or in-progress. However, because we found that some education levels were extremely rare in our dataset (e.g. PhD), we converted them into years of education with the following mapping: *Incomplete High School* -> 10, *High School* -> 12, *Associate's Degree* -> 14, *Bachelor's Degree* -> 16, *Master's Degree* -> 18, *Professional Degree* -> 20, *PhD* -> 21.

2.5.1 Education model

$$\text{num\_solved} \sim \text{education}$$

Table 10: Education-only regression coefficient and R^2 estimates and 95% highest density intervals.

Term	Estimate	HDI-L	HDI-U
intercept	36.129	20.514	51.166
education	1.455	0.448	2.453
R^2	0.032	0.003	0.077

2.5.2 All math model

num_solved $\sim$ alg + geom + trig + sv_calc + mv_calc + linalg + pr_stat + disc + logic

Table 11: Education-only regression coefficient and R^2 estimates and 95% highest density intervals.

Term	Estimate	HDI-L	HDI-U
intercept	38.757	32.505	44.885
alg	9.684	2.874	16.781
geom	9.813	3.785	16.034
trig	3.139	-2.368	8.724
sv_calc	4.617	-2.156	11.540
mv_calc	-1.006	-8.985	6.611
linalg	0.392	-5.576	6.540
pr_stat	2.619	-2.267	7.237
disc	6.361	-3.138	15.396
logic	-1.280	-8.287	5.729
R^2	0.207	0.132	0.279

2.5.3 Algebra and geometry model

$$\text{num\_solved} \sim \text{alg} + \text{geom}$$

Table 12: Algebra and geometry regression coefficient and R^2 estimates and 95% highest density intervals.

Term	Estimate	HDI-L	HDI-U
intercept	40.722	34.966	46.395
alg	9.131	2.146	16.065
geom	12.704	7.070	18.691
R^2	0.152	0.086	0.226

2.5.4 Education, algebra and geometry model

$$\text{num\_solved} \sim \text{education} + \text{alg} + \text{geom}$$

Table 13: Education, algebra, and geometry regression coefficient and R^2 estimates and 95% highest density intervals.

Term	Estimate	HDI-L	HDI-U
intercept	25.476	10.401	39.890
education	1.046	0.099	1.995
alg	9.591	2.573	16.401
geom	11.479	5.605	17.328
R^2	0.168	0.098	0.240

2.5.5 All math and education model

num_solved $\sim$ education + alg + geom + trig + sv_calc + mv_calc + linalg + pr_stat + disc + logic

Table 14: Education and math regression coefficient and R^2 estimates and 95% highest density intervals.

Term	Estimate	HDI-L	HDI-U
intercept	30.14	15.36	45.15
education	0.61	-0.39	1.56
alg	9.93	2.98	16.81
geom	9.40	3.39	15.53
trig	2.70	-3.22	8.55
sv_calc	4.64	-2.18	11.55
mv_calc	-1.28	-8.87	6.83
linalg	0.40	-6.02	6.69
pr_stat	2.10	-2.73	6.96
disc	6.20	-3.09	15.17
logic	-1.62	-8.52	5.51
R^2	0.213	0.139	0.285

3 Unreported Regressions

Here, we describe the regressions that were not included in the main article but were committed in the pre-registration. Due to the large number of parameters, we do not include their estimates here. For the coefficients of these models, see the spreadsheet in the project repository.

As in the previous section, t refers to the trial number (between 1 and 64) and s refers to the subject. The $(1 + \log_2 t|s)$ included in each regression indicates participant-level random effects. All other terms are fixed effects. Response time models were fitted using only correct trials.

3.1 Digit Sets (DS)

One additional regression was conducted for the digit sets analysis which included an interaction term between the treatment and practice effects. This model was fitted using data from all 64 test phase trials.

Accuracy model

$$P(\text{correct}_{t,s}) \sim DS + log_2 t + DS * log_2 t + (1 + log_2 t|s)$$

Duration model

$$\log(\text{duration}_{t,s}) \sim DS + log_2 t + DS * log_2 t + (1 + log_2 t|s)$$

3.2 Goal Position (GP)

One additional regression was conducted for the goal position analysis which included an interaction term between the treatment and practice effects. While this model was not committed in the pre-registration, we include it for completeness. This model was fitted using data from all 64 test phase trials.

Accuracy model

$$P(\text{correct}_{t,s}) \sim GP + log_2 t + GP * log_2 t + (1 + log_2 t|s)$$

Duration model

$$\log(\text{duration}_{t,s}) \sim GP + log_2 t + GP * log_2 t + (1 + log_2 t|s)$$

3.3 House Index (HI) and Cell Index (CI)

These regressions were the original formulations for analyzing the effect of moving the goal cell. Due to the transposition of the grid when applying the house type (HT) condition and the high perceptual dissimilarity, we intended to perform separate regressions for puzzles with the HT condition applied and for puzzles without the HT condition applied. However, post pre-registration, we had decided that the effect of interest was less about exactly which axis the goal cell had translated across and more about the presence of a change at all. Moreover, we had originally planned to account for an interaction between HI and CI, but decided that the interaction complicated the interpretation. Thus, we collapsed house index and cell index conditions as the goal position (GP) condition in favor of improved power and interpretability, considering the condition to be applied if one or both of the house index or cell index had been applied.

Note that each of the four regressions below have been fitted separately using trials with and without house type conditions.

Accuracy: 16-trial and 48-trial models

$$P(\text{correct}_{t,s}) \sim HI + CI + HI * CI + (1 + log_2 t|s)$$

Duration: 16-trial and 48-trial models

$$P(\text{duration}_{t,s}) \sim HI + CI + HI * CI + (1 + log_2 t|s)$$

Accuracy: 64-trial model

$$P(\text{correct}_{t,s}) \sim HI + CI + HI * CI + HI * log_2 t + CI * log_2 t + HI * CI * log_2 t + (1 + log_2 t|s)$$

Duration: 64-trial model

$$P(\text{duration}_{t,s}) \sim HI + CI + HI * CI + HI * log_2 t + CI * log_2 t + HI * CI * log_2 t + (1 + log_2 t|s)$$

3.4 House Type (HT)

One additional regression was conducted for the house type analysis which included an interaction term between the treatment and practice effects. This model was fitted using data from all 64 test phase trials.

Accuracy model

$$P(\text{correct}_{t,s}) \sim HT + log_2 t + HT * log_2 t + (1 + log_2 t|s)$$

Duration model

$$\log(\text{duration}_{t,s}) \sim HT + log_2 t + HT * log_2 t + (1 + log_2 t|s)$$

4 Hidden Markov Model

4.1 Aggregate Model

The aggregate model was used to fit the hidden Markov model parameters using the aggregate data across all solvers using the following functional form:

$$P(\text{response}_t) = \pi \mathbf{A}^{t-1}\mathbf{B} \tag{1}$$

where $P(\text{response}_t)$ is the inferred distribution of responses in a given trial t, π is the initial strategy distribution across all participants, $\mathbf{A}$ is the transition matrix between strategies, and $\mathbf{B}$ is the response emission matrix. Rows in $\mathbf{A}$ indicate source strategies and columns indicate destination strategies. Rows in $\mathbf{B}$ indicate strategies and columns indicate responses. Strategies are presented in the following order: uniform guess, avoid direct contradictions, prevalent digits, and successful. Responses are presented in the following order: in-house, absent, distractor, and target. Note the 0's in the lower-triangle of $\mathbf{A}$, representing the assumption that participants always use the best strategy available to them. In $\mathbf{B}$, the three fitted ϵ terms in $\mathbf{B}$ represent probabilities of responses of types that are as errors under each strategy. Except for the accidental selection of the distractor when using a successful strategy, most error terms were found to be negligible. Non-error responses are treated as distributed evenly across all digits that are candidates under the given strategy.

$$\pi = [\pi_1 \quad \pi_2 \quad \pi_3 \quad \pi_4]^T \tag{2}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$$\mathbf{B} = \begin{bmatrix} 3/9 & 4/9 & 1/9 & 1/9 \\ \epsilon_1 & (1-\epsilon_1)\frac{4}{6} & (1-\epsilon_1)\frac{1}{6} & (1-\epsilon_1)\frac{1}{6} \\ \epsilon_1 & \epsilon_2 & (1-\epsilon_1-\epsilon_2)\frac{1}{2} & (1-\epsilon_1-\epsilon_2)\frac{1}{2} \\ \epsilon_1 & \epsilon_2 & \epsilon_3 & 1-\epsilon_1-\epsilon_2-\epsilon_3 \end{bmatrix} \tag{4}$$

The four π, nine a, and three e parameters were fitted to the aggregate data by minimizing the cross-entropy loss over 2000 epochs using Adam gradient descent. Separate models were fitted for solvers and non-solvers. The resulting π, $\mathbf{A}$ and $\mathbf{B}$ matrices for solvers (subscripted with s) and shown first, followed by the corresponding matrices for non-solvers (subscripted with ns).

$$\pi_{\mathbf{s}} = [0.053 \quad 0.286 \quad 0.293 \quad 0.368]^T \tag{5}$$

$$\mathbf{A_s} = \begin{bmatrix} 0.001 & 0.989 & 0.009 & 0.001 \\ 0 & 0.276 & 0.720 & 0.004 \\ 0 & 0 & 0.772 & 0.228 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

$$\mathbf{B_s} = \begin{bmatrix} 0.333 & 0.444 & 0.111 & 0.111 \\ 0.001 & 0.666 & 0.167 & 0.167 \\ 0.001 & 0.003 & 0.498 & 0.498 \\ 0.001 & 0.003 & 0.046 & 0.951 \end{bmatrix} \tag{7}$$

$$\pi_{\mathbf{ns}} = [0.145 \quad 0.357 \quad 0.476 \quad 0.022]^T \tag{8}$$

$$\mathbf{A_{ns}} = \begin{bmatrix} 0.617 & 0.004 & 0.374 & 0.111 \\ 0 & 0.781 & 0.209 & 0.010 \\ 0 & 0 & 0.992 & 0.008 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

$$\mathbf{B_{ns}} = \begin{bmatrix} 0.333 & 0.444 & 0.111 & 0.111 \\ 0.008 & 0.661 & 0.165 & 0.165 \\ 0.008 & 0.033 & 0.480 & 0.480 \\ 0.008 & 0.033 & 0.143 & 0.816 \end{bmatrix} \tag{10}$$

4.2 Incremental Transition Variants

In estimating the participants' response sequence likelihoods under the incremental transition hypothesis, we tried 3 different variants, the best of which is reported in the main manuscript. We describe the two other variants below.

As described in the main manuscript, the incremental hypothesis assumes a weighted superposition of multiple strategies and iteratively applies a matrix multiplication with the fitted transition matrix to generate strategy weight vectors for subsequent trials. Naturally, this means that the fitted strategy distribution vector for trial 1 can also be interpreted as a weight vector. However, the data are extremely unlikely under this variant of the incremental transition hypothesis that treats the initial strategy distribution as a superposition of strategies with weights that all participants share at the outset, since there were several participants that demonstrated perfect or near perfect accuracy in the practice phase (e.g. Subject 38 as shown in the main manuscript). Such profiles are more likely under the discrete transition model, since under this model, more than a third of participants are thought to begin the practice phase using a successful strategy. Therefore, we more formally contrasted the discrete model with two incremental model versions that assign individuals to one of the four possible strategies for trial 1, i.e. one strategy having weight of 1 and all others 0, allowing those not starting with a successful strategy to transition incrementally toward better strategies.

Although the discrete hypothesis simulation has variation in strategy trajectories built in to the generative method, the incremental hypothesis simulation always applies matrix multiplication using the same fitted transition matrix. Thus, although the final responses are variable due to sampling using the emission matrix weights, the simulation would implicitly assume that all participants share one of four strategy trajectories (one for each starting strategy). Indeed, this method produced an even higher Bayes factor, 143.359, almost 3x more in favor of the discrete hypothesis than the method described in the main manuscript. This prompted the use of the Dirichlet distributions as described in the Methods section of the main manuscript and in the section below.

4.3 Simulating Samples

To estimate the Dirichlet distribution parameters for sampling transition matrices, we calculated the expected number of times each transition was observed based on the fitted values of $\pi_\mathbf{s}$, $\mathbf{A_s}$ and $\mathbf{B_s}$:

$$88 * \sum_{t=1}^{25} \pi_s {\mathbf{A_s}}^{t-1} = \begin{bmatrix} 0.005 & 4.623 & 0.043 & 0.004 \\ 0 & 11.369 & 29.642 & 0.178 \\ 0 & 0 & 187.37 & 55.349 \\ 0 & 0 & 0 & 1999.4 \end{bmatrix} \tag{11}$$

To estimate the Dirichlet distribution parameters for sampling emission matrices, we calculated the expected number of times each emission was observed:

$$88 * \sum_{t=1}^{25} \pi_s {\mathbf{A_s}}^{t-1} \mathbf{B_s} = \begin{bmatrix} 1.558 & 2.078 & 0.519 & 0.519 \\ 0.040 & 27.433 & 6.858 & 6.858 \\ 0.235 & 0.676 & 120.9 & 120.9 \\ 1.933 & 55.670 & 91.372 & 1900.5 \end{bmatrix} \tag{12}$$

Figure 4 shows the non-target response distributions using the two transition types along with the actual participants' responses.

Each row in the above matrices represent the parameters of separate Dirichlet distributions. Thus, each set of sampled HMM parameters would consist of a single categorical sample from the π distribution and 8 samples from the 8 Dirichlet distributions.

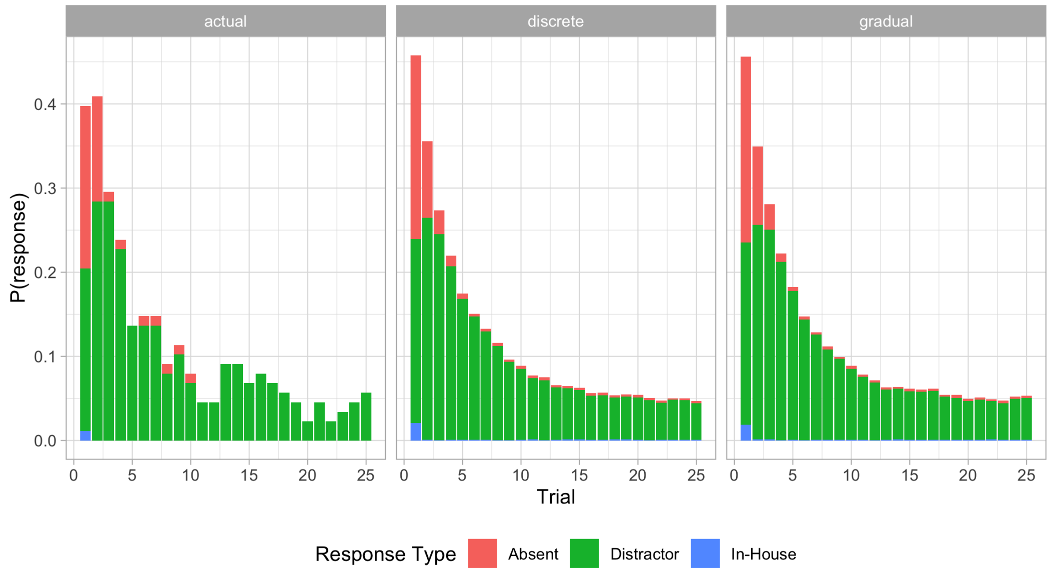


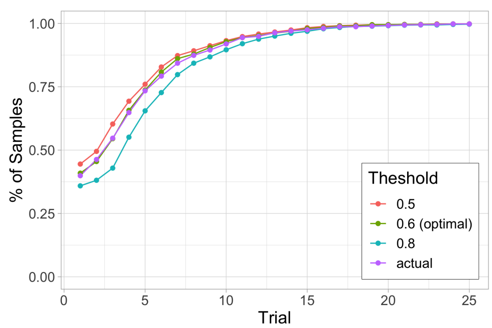
Figure 4: The distribution of non-target responses. Left panel shows actual participants' data, center panel shows samples generated using discrete transitions, and the right panel shows samples generated using incremental transitions.

4.4 Individualized Model Accuracy

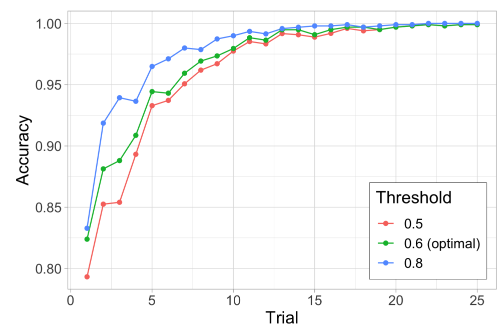
We tested the individually fitted model accuracies using 1000 simulated sample response trajectories with discrete transitions. For each sample, we fitted a separate model and found the posterior probabilities of latent strategies. We optimized our decision criterion based on how similar the predictions were to the actual frequencies of the successful strategy by minimizing the cross-entropy as shown in Equation 13 where y_t is the actual frequency of successful strategy at trial t and p_t is the frequency of samples inferred to be using the successful strategy at trial t. Using this method, we found that the optimal threshold is 60%. We also attempted maximizing accuracy which produced a negligibly different optimal threshold of 58%. Figure 5a shows how closely the inferred distribution of the successful strategy matches the actual simulated data.

$$BCE = \sum_{t=1}^{25} -(y_t \log(p_t) + (1 - y_t) \log(1 - p_t)) \tag{13}$$

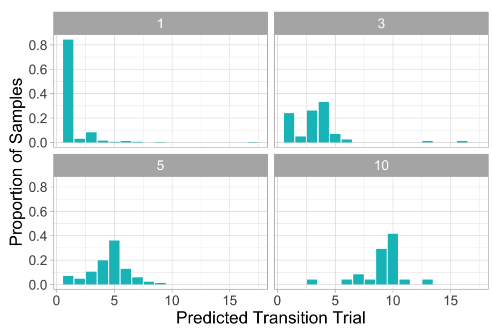
We looked at the proportion of samples that actually used the successful strategy given that the model inferred that the probability of the successful strategy on the first trial on was greater than 60% (i.e. $P(\text{successful}_1 \mid \hat{P}(\text{successful}_1) \geq 60\%)$, and found that 82.4% of actually did so (Figure 5b). We note that while the model was not always correct in predicting when transitions occurred, its inferences were distributed close to the ground truth (Figure 5c). Lastly, we found that the model accuracy was generally the lowest on trials adjacent to when strategy transitions occurred with an AUC measure of 86.0% for the successful strategy (Figure 5b).



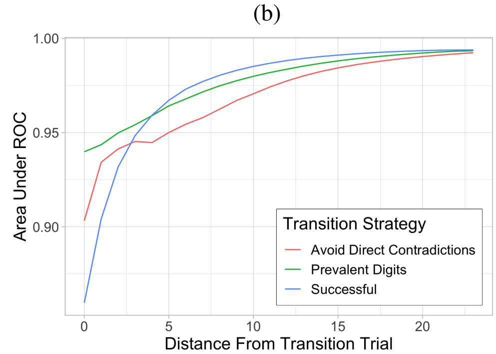
(a)



(b)



(c)



(d)

Figure 5: Individualized model results using 1000 sample trajectories with discrete transitions. (a) Percent of samples inferred to be using the successful strategy evaluated using different decision threshold values. Actual percent of samples using successful strategy also shown. (b) Percentage of samples that actually used the successful strategy, given that the model inferred were using the successful strategy, i.e. $P(\text{successful} \mid \hat{P}(\text{successful}) \geq \text{threshold})$. (c) Distribution of inferred trial that the successful strategy was first used, given the actual trial that the successful strategy was first used. Top-left panel shows samples that used the successful strategy from trial 1, top-right panel shows samples that used the successful strategy from trial 3, etc. Predictions made using a decision threshold of 0.6. (d) The model AUC for different strategies. Line colors indicate the strategy that was transitioned into. X-axis indicates absolute distance from the trial where the strategy changed and each value contains trials on both sides. For example, 0 includes both the last trial of the previous strategy and the first trial of the new strategy.

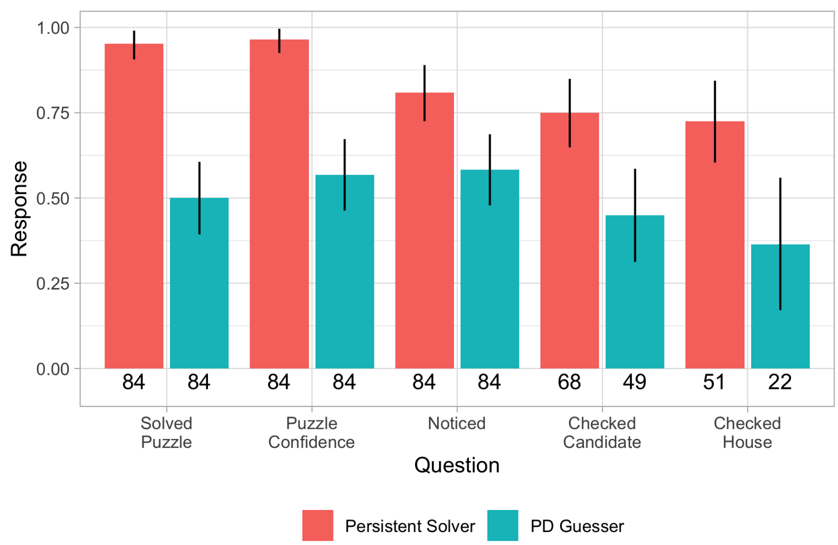


Figure 6: Quantitative responses by group. Total number of responses at the bottom of each bar for each question and group. Each participant was asked 3 attention check questions, producing 252 responses total. *Attention Check* bars indicate proportion of responses that responded correctly. *Solve Puzzle* bars indicate proportion of participants that solved the puzzles.

5 Questionnaire

5.1 Accuracy and confidence

On the final puzzle presented at the start of the questionnaire phase, 80 of the 84 persistent-solvers and 42 of the 84 PD-guessers solved this new puzzle correctly. We also asked the participants their level of confidence that their answers were correct expressed as a percentage. Persistent-solvers reported an average confidence of 96.43% (95% HDI = [92.45%, 99.59%]) while PD-guessers reported 56.73% (95% HDI = [46.29%, 67.23%]) for a difference of 39.7% (95% HDI = [28.60%, 51.18%]).

5.2 Multiple choice questions

Comparing the responses between the persistent-solvers and PD-guessers, we found significant differences across all 3 questions (Table 15). Specifically, when asked about how they chose which of the two modal digits they considered, 80.95% of persistent-solvers and 58.33% of PD-guessers responded that they had noticed something in the puzzle that made one candidate seem more likely versus chose arbitrarily. Among those that noticed something, 75.00% of persistent-solvers and 45.90% of PD-guessers responded that they had further checked to see if their chosen candidate was correct versus submitted without checking. Finally, among those that checked, 72.55% of persistent-solvers and 36.36% of PD-guessers responded that they checked to see if the candidate could go in another cell in the house versus looking for information in other numbers.

Table 15: Questionnaire multiple choice responses rates, differences between persistent solvers (PS) and PD guessers (PDG), and 95% highest density intervals (HDI) of the differences.

Measure	Mean PS	Mean PDG	Mean Diff	Diff HDI-L	Diff HDI-U
Attention Check	90.08%	82.54%	7.54%	1.68%	13.55%
Solved Puzzle	95.24%	50.00%	45.24%	32.75%	55.98%
Puzzle Confidence	96.43%	55.95%	40.48%	28.60%	51.18%
Noticed	80.95%	58.33%	22.62%	8.83%	35.78%
Checked Candidate	75.00%	45.90%	30.10%	12.24%	45.54%
Checked House	72.55%	36.36%	36.19%	12.33%	56.94%

5.3 Free response ratings

5.3.1 Participant selection

For our ratings of questionnaire results, we focused our analysis on subsets of the solver and non-solver groups based on phase two performance. Among the solvers, we selected a subset we call *persistent solvers* who were both classified as solvers at the end of the practice phase and remained so in the test phase. To identify these participants, we fitted a logistic regression to the test phase data, similar to the logistic regression fitted to the practice phase data in the original solver classification. Only participants with predicted accuracy for the 64th trial of at least 80% were kept for the questionnaire analyses. Among the non-solvers, we focused on a subset we call *PD-guessers* based on the fact that (a) they exhibited evidence of exceeding chance performance by the end of the test phase and (b) were choosing predominantly between the two prevalent digits. We used the same logistic repression as above to select those with predicted accuracy of at most 60%, and also required that they had only solved 3 to 5 of the last 8 puzzles in the test phase and that they had selected the target or the distractor in at least 58 of the 64 test phase puzzles. Applying these filters yielded 84 of 88 persistent solvers and 84 of 183 PD-guessers for the questionnaire analyses.

5.3.2 Rating options

The following options were available for identifying whether or not the responses indicated awareness of error. This was only done for puzzles that were not solved correctly.

1. **Yes with explanation:** The participant indicates a clear realization that the answer they entered was not correct, and explains why they decided that the other answer was correct.
2. **No or yes w/o explanation:** The participant's response was incorrect, but they did not explicitly report realizing that they had answered incorrectly, or the response suggests they think they were wrong without certainty or clarity about the reason.

The following options were available for identifying prevalent digits mentioned in the responses.

1. **Both:** The participant mentioned both prevalent digits by name, or otherwise led you to believe they realize that they have to choose among these two digits.
2. **Target:** The participant mentioned the target, by name or in some other way, but did not mention the distractor, and did not indicate that there were two prevalent digits.
3. **Distractor:** The participant mentioned the distractor, by name or in some other way, but did not mention the target.
4. **Neither:** The participant did not mention either the target or the distractor, either by name or some other way, and there is no indication of awareness that there are two prevalent digits.
5. **Vague or uncertain:** The participants' answer could be signaling that they were selecting between the two prevalent digits or that they were focusing on a digit that cannot go elsewhere or the distractor, but they do so vaguely or in a way that it is hard to be certain.

The following options were available for identifying the basis for choosing between the two prevalent digits. The options were presented in this order to the raters with the alphabetical labels A-J. The V1-V3, U1-U2, I1-I3, and M coding scheme was adopted for presentation purposes in the main manuscript and were not used during the actual ratings.

A **Chose digit that cannot go elsewhere (V1):** The participant states that a digit cannot go in any of the empty blue cells and gives that digit as the answer. As a shortcut variant of category A, participant states that they found a cell in the target house that was only constrained by one of the two prevalent digits and chose that digit as the answer. To be valid, this must be a cell that is not in fact constrained by the distractor.

B **One PD can go elsewhere, chose the other (V2):** The participant states that one of the prevalent digits can go in one of the empty blue cells and concludes that the other prevalent digit is the answer.

C **Found cell where one PD could go and the other could not (V3):** The participant mentions finding a blue cell that can contain one of the two prevalent digits and not the other, and chooses the digit that cannot go in that cell as the answer.

D **Potentially valid but general or not fully specified (U1):** The participant's response provides incomplete information about how they came to choose their answer, but could be a general or under-specified description of a valid procedure.

E **Explicit guess (I1):** The participant indicates that they know they are guessing (either completely at random or between the two prevalent digits)

F **Irrelevant basis for choice (I2):** The participant indicated a basis for choosing the answer that was not related to the logic of the puzzle.

G **Chose the most frequent digit (I3):** The participant indicates they chose the digit that occurred the most frequently in the puzzle, apparently not realizing that there were always two digits that both occurred exactly 3 times.

H **Unclear, confused or missing basis (U2):** The answer attempts to provide information about how the choice was made, but is unclear, incorrect, confusing, or fails to specify the procedure used to select between the PDs.

I **Did not answer the question (M):** Response does not address the question.

J **Other:** None of the above.

5.3.3 Rating Design

One author (JLM) went through all 168 participants' responses to the first free-response question in an effort to develop a set of categories into which individual participants' responses could be sorted, with access to the specific puzzle each participant had just attempted to solve, the participant's chosen response, and summary scores characterizing the participant's performance both on the practice and test phases of the experiment. In so doing, he noted that persistent solvers typically (a) referenced the target digit or both the target and the distractor; (b) provided one of three types of responses that seemed to describe a valid solution strategy given the rules of Sudoku and further constraints on the constructions of the puzzles used, and, (c) in the small number of cases in which a solver incorrectly chose the distractor, explicitly reported realizing that they had selected the wrong digit and explained why their choice was wrong. The author also noted that non-solvers were (a) less likely to mention either or both prevalent digits even if they almost always chose one of the two prevalent digits, (b) rarely described a valid solution strategy, and (c) in the cases where they chose the distractor, never expressed a clear understanding that the response was incorrect. The author began to develop a set of scoring criteria that could be used to corroborate these impressions. He developed preliminary versions of the awareness of error, mention of prevalent digits, and basis for choice rating categories, including preliminary versions of the set of options for raters to choose among with made up illustrative examples. The other author (AJHN) checked the categories and descriptions for clarity and distinctness with minimum reference to the actual data. The two authors then met to refine the categories, their short names, their longer descriptions, and the illustrative examples in an attempt to span the range of variation of responses that fell into each of the categories. The authors jointly agreed on the criteria used to identify the persistent solvers as a subset of the solvers and the PD-guessers as a subset of the non-solvers. Note that this excluded several individuals who might have been late or partial solvers as well as others who chose responses other than one of the two PDs on more than 6 (corresponding to more than 10%) of the test trials.

Next, both authors rated a pilot set of 20 participants' responses. As with the final ratings discussed below, these ratings were carried out with access to the specific puzzle each of the participants solved and the answer the participant gave, but without reference to information about the participant's classification as a persistent solver or a PD-guesser or any other information about the participant including their performance in the training or test phases of the experiment. To make the pilot set representative without inspecting the actual responses, we sampled 9 *persistent solvers* that solved the puzzle, 1 *persistent solver* that did not solve the puzzle, 6 *PD-guessers* that solved the puzzle, and 4 *PD-guessers* that did not solve the puzzle. After independently rating all 20 participants, the authors further discussed and refined the categories, descriptions, and examples. In this process we noted that some responses were difficult to categorize with certainty, but that in these cases the uncertainty was restricted to two alternative possibilities. Accordingly we adjusted the scoring system to allow raters to provide a second choice category to capture these cases.

For the final ratings, we wished to measure the extent to which persistent solvers (a) referenced the target digit or both the target and the distractor in their response to the first free response question;

(b) indicated that they had employed one of the three valid solution strategies; and (c), reported a realization of their error and explanation of why it was wrong on cases where they entered the distractor as their response instead of the target. We wanted to demonstrate that this determination could be made reliably and in the absence of any other information beyond the specific puzzle the participant received during the questionnaire phase, the digit choice the participant entered as their solution, and their response to the first free response question about how they solved the problem. We further wished to measure the extent to which PD-guessers exhibited these same tendencies, and if not, to document the distribution of responses they provided related to all three of these questions. Because author JLM had previously considered all of the participants responses with additional information, this author did not contribute to the final ratings. Instead we sought an individual who would be motivated to work conscientiously to identify the strategy used by each participant without having had any opportunity to consider any other information about each participant beyond the specific puzzle the participant saw, their digit choice, and their answer to the first free response question. A Stanford computer science Master's student (LKS) pursuing research on human problem solving in collaboration with author JLM met this criterion. LKS had participated in lab meetings in which the behavioral choice results from the study had been discussed, including the classification of participants as solvers or non-solvers, but had not been exposed to details of the performance of any of the participants. LKS and author AJHN then served as the two raters for the full set of participants. Although AJHN had previously considered some of the participants free responses with other information about their performance, we reasoned that if agreement between the two raters was high, this would be a sufficient indication that the ratings could be made reliably without access to additional information about the participants. LKS was paid $30 per hour for his participation. He first completed the experiment to familiarize himself with the study, and was classified as a persistent solver, achieving 100% accuracy in the practice phase and 100% accuracy in the test phase of the experiment. He then reviewed the rating instructions and guide thoroughly before rating the 20 pilot participants. Subsequently JLM, AJHN, and LKS met to discuss the ratings criteria and discrepancies between AJHN's and LKS's ratings of these participants. As a result of this discussion we made some final refinements in the category descriptions and examples until all three people were satisfied that they were as clear as possible and that all three had a common understanding of the categories. AJHN and LKS then adjusted their ratings in accordance with the refined categories. We did not require the two raters to necessarily agree perfectly on their first choice rating category, allowing disagreements to reflect the ambiguities that were present in some of the responses.

Finally AJHN and LKS proceeded to rate the responses of the remaining 148 persistent solvers and PD-guessers. To rate each participant, the rater first looked at the specific puzzle the participant had received and judged whether or not the participant's response was correct (all of these participants choices were one of the PDs). The rater entered this judgment in a spreadsheet, which then checked whether the rater's judgment of the participant's response was correct and alerted the rater in the rare case that the rater was incorrect in judging the correctness of the participant's response (this occurred 0 times for rater LKS and 1 time for AJHN). The rater proceeded to rate the participants' free responses, proceeding through the ratings in the order listed above. Raters were asked to complete their ratings in several sessions spread out over a one-week period, and to proceed slowly and to work for no more than an hour at a single sitting, though they could resume work after a short break. Raters were also asked to check their ratings and to reconsider cases where they were initially uncertain, and were allowed to adjust any of their ratings after completing a first pass rating the full set. We expected there would be some cases of inconsistency in the ratings, and did not seek to resolve these inconsistencies, so that for all but 20 of the rated participants, the final ratings of the two raters were reached without influence from the other rater.

5.3.4 Consistency between raters

For each participant, as an attention check, raters first determined whether or not the participant's answer to the puzzle was correct. Both raters scored very highly on this with accuracies of 97% and 100%. Raters were also asked to determine whether or not the incorrect participants expressed awareness of their errors. Of the 4 persistent solvers and 42 PD-guessers that did not correctly solve the puzzles, both raters agreed on all 46 identifications.

We did not have ground truth for mentions of the two prevalent digits (PDs). However, this task had clear boundaries between categories and generally less ambiguity, allowing us to compare PD ratings

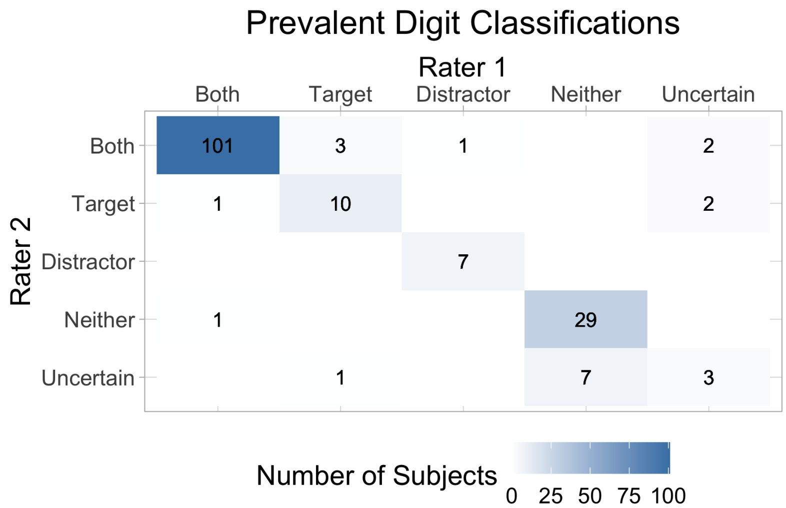


Figure 7: Prevalent Digit classifications.

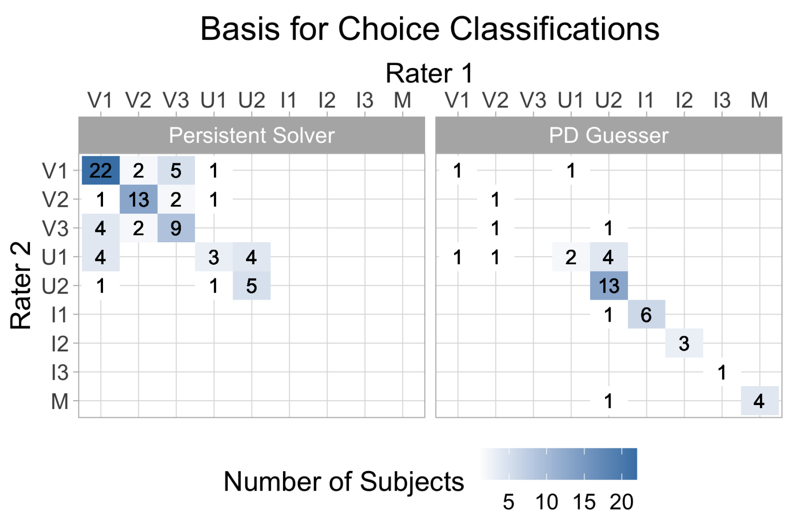


Figure 8: Basis for choice classifications. Only first choices shown.

as a measure of consistency between the two raters and found a very high agreement rate of 89.3% across all 168 participants (**Figure 7**).

Identifying bases for choice was overall a much tougher task, often due to incomplete ideas or ambiguous phrasing from the participants. Even when the responses were clear and relatively easy to understand, the different valid bases could be viewed as different ways of describing the same logic, making classification difficult for some responses. Moreover, some participants mentioned multiple strategies, sometimes accounting for counterfactual scenarios where the first PD they considered ended up being the distractor but acknowledging it could just as likely have been the target. Therefore, we allowed a second option in case a response mentioned multiple legitimate strategies or was borderline unclear.

As mentioned in the main manuscript, 69.6% of first choice ratings agreed between raters (Figure 8). This was consistent even when looking at the 80 persistent-solvers and 42 PD-guessers that correctly solved the questionnaire puzzles with 68.0% agreement (Figure 9). When comparing whether ratings were valid, unclear, invalid, or missing, the raters were 86.9% in agreement overall and 89.3% among correct responders.

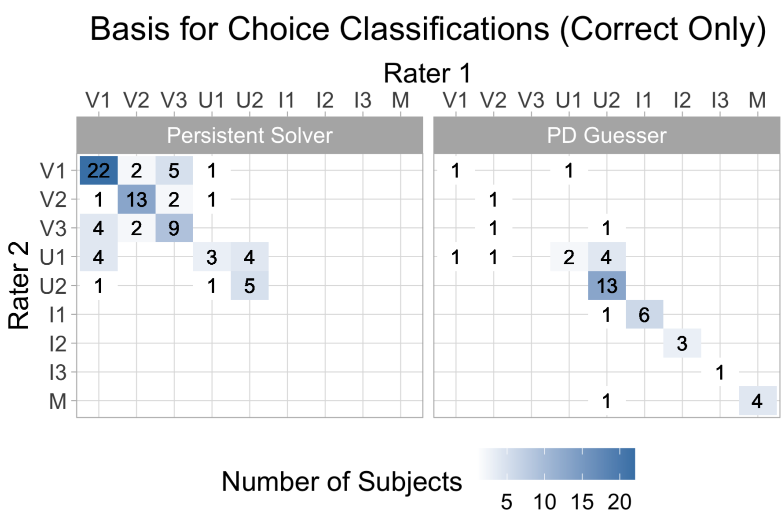


Figure 9: Basis for choice classifications among the 80 persistent-solvers and 42 PD-guessers that correctly solved the questionnaire puzzles. Only first choices shown.

6 Recurrent Relational Network

Here, we describe our replication of the Recurrent Relational Network. The main architecture was implemented following the description in the original article as we understood it, but with simplifications. Next we describe the architecture and note our deviations from the original.

6.1 Model Architecture

The Recurrent Relational Network (RRN) uses a local message passing scheme in a graph where each cell in the Sudoku grid is a separate node. Cells that share houses (row, column, or box) are considered neighbors and their nodes share edges between them. Each node i at time step t is represented by a hidden state vector h_i^t, where $h_i^1 = x_i$ is an embedding of the cell's initial state (a blank or a clue). In the original paper, the initial cell state is given by

$$x_i = \text{MLP}(\text{concat}(\text{embed}(d_i), \text{embed}(\text{row}_i), \text{embed}(\text{column}_i)))$$

where d_i is the initial content of the cell, if any. In our implementation, we found the cell coordinate information detrimental to performance and simplified the embedding to $x_i = \text{embed}(d_i)$.

At each step, for each cell i and its neighbor j, a vector representing the message from j to i is computed using $m_{ij}^t = \text{MLP}(\text{concat}(h_i^{t-1}, h_j^{t-1}))$. Then the messages from all of i's neighbors are summed to $m_i^t = \sum_j m_{ij}^t$. The subsequent hidden state vector of each cell is updated:

$$h_i^t, c_i^t = \text{LSTM}(\text{MLP}(\text{concat}(x_i, m_i^t)), c_i^{t-1})$$

In our implementation, however, we found a simple linear layer to be sufficient replacements for the two MLPs.

An output vector is also calculated for each step and cell using a linear decoding layer: $o_i^t = \text{linear}(h_i^t)$, which is then used to calculate the cross-entropy loss with the solution for the cell.

6.2 Replication of Results

We attempted to replicate the findings of Palm et al. using the same dataset and parameters as the original network. However, due to limited computational resources, we had to use our simplified architecture and fewer parameters. Specifically, we used a digit embedding size of 10 (9 digits + blank; original model used 16) and reduced the MLPs to single linear layers. Moreover, we reduced the number of training epochs to 100 and the batch size to 20. The models were trained on the same 180000 training puzzles and 18000 test puzzles as the original paper. Our replication attempt resulted in 97.0% of all cells solved and 74.2% of test puzzles fully solved for all 81 cells. While this is much lower than the 96.6% reported in the original paper, it demonstrates that our simplified architecture can achieve a significant level of success while retaining key features of the original architecture.

6.3 Adapting to Hidden Singles

To compare the model to human performance, we trained and tested the RRN model using the Hidden Singles puzzles similar to what we provided for our participants. Specifically, we created a set of puzzles with the same Digit Set, House Index, Cell Index, and House Type to train the model. We then generated 64 puzzles with varied features to test for the model's ability to generalize. Because the majority of the grid was empty, we only calculated the cross-entropy loss for the goal cell and the initial hint cells. We found the latter to be a critical auxiliary signal for the model to successfully train.

The model was written using PyTorch and trained using Adam with batch size = 100, learning rate = 0.001 (original model used 0.0002), and L2 regularization of 0.0001. We used a digit embedding size of 10 and a hidden layer size of 48 (original model used 96). We found these parameters sufficient given the significantly lower complexity of the Hidden Singles puzzles compared to full Sudoku puzzles.

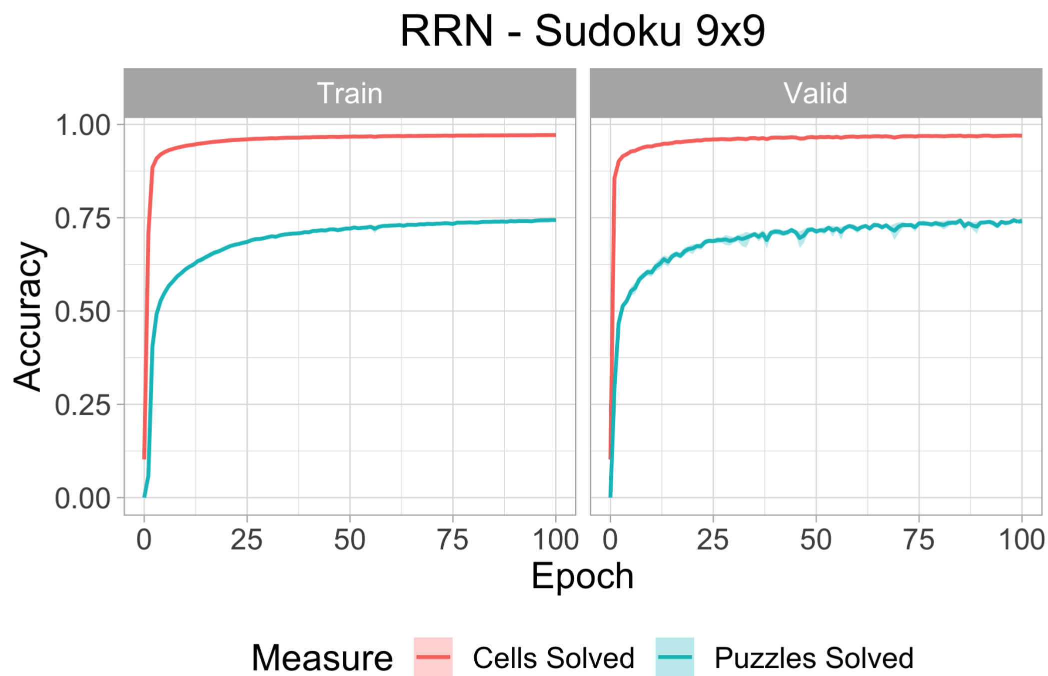


Figure 10: Replication results of Palm et al. on full 9x9 Sudoku puzzles. Results show averages across 10 different model instances.

As noted in the main article, the model trained gradually over tens of thousands of puzzle presentations with no discernible stages suggesting discrete strategies. It struggled to generalize even to control puzzles (puzzles sharing features with training set) when provided with fewer than 300 puzzles in the training set. Moreover, the model demonstrated perfect transfer for House Index, Cell Index, and House Type, but no transfer at all for Digit Set. The invariance to only the positional features is a result of having it built into the architecture of the model.

6.4 Inducing Digit Invariance

If weight sharing across spatial features produces spatial invariance, it stands to reason that weight sharing across the digit feature can produce digit invariance. Following this logic, we expanded the Palm RRN by defining a node for each cell and digit (x, y, digit), thus producing $9^3 = 729$ nodes. We defined edges between any two nodes that represent different digits of the same cell or two nodes that represent the same digit of two cells sharing a house. For example, nodes for (3, 3, 1) and (3, 3, 6) would share an edge because they both represent the cell at (3, 3). Nodes for (3, 3, 9) and (3, 6, 9) would share an edge as they share the same house at Row 3 and the same digit, but (3, 3, 7) and (3, 6, 9) would not share an edge since they do not share the same digit.

Since individual nodes represent digits, we no longer needed a layer to embed numbers. The output layer was also modified such that each node maps to a scalar logit, and the 9 digit nodes of a cell would together be softmaxed to produce the output probability vector. Due to memory resource constraints, the hidden state and message vector sizes were reduced to 16, the batch size reduced to 10, and the learning rate reduced to 0.0001.

7 Additional Figures

7.1 Solvers vs Non-solvers

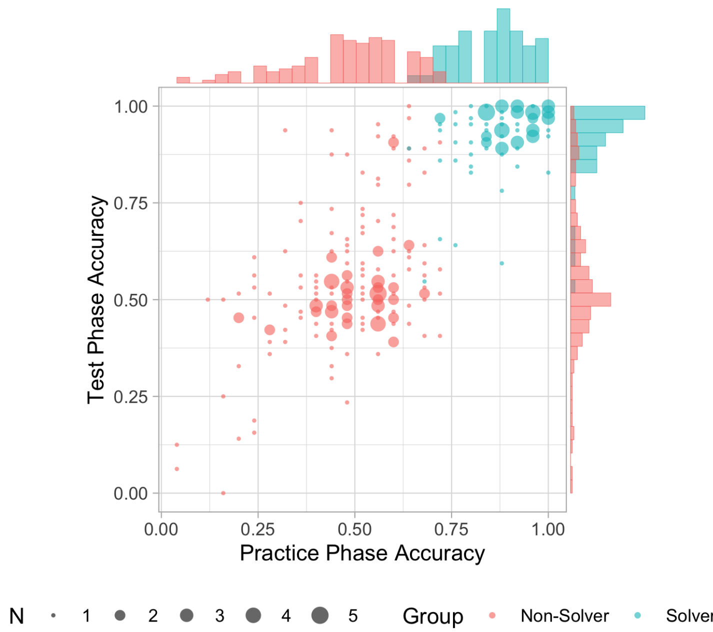


Figure 11: Overall accuracy during the practice and test phases. Points represent individual participants. Multiple participants sometimes solved the exact same numbers of puzzles in both the practice and test phases, indicated by the size of the points. Marginal histograms show proportions by group, not raw counts.

7.1.1 Practice phase accuracy

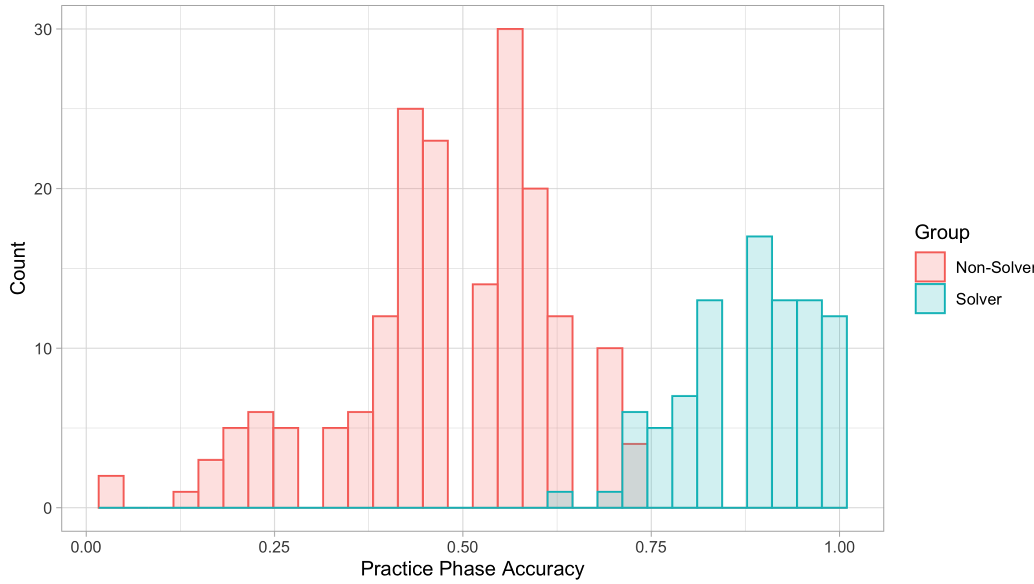


Figure 12: Histogram of overall accuracy during the practice phase.

7.1.2 Test phase accuracy

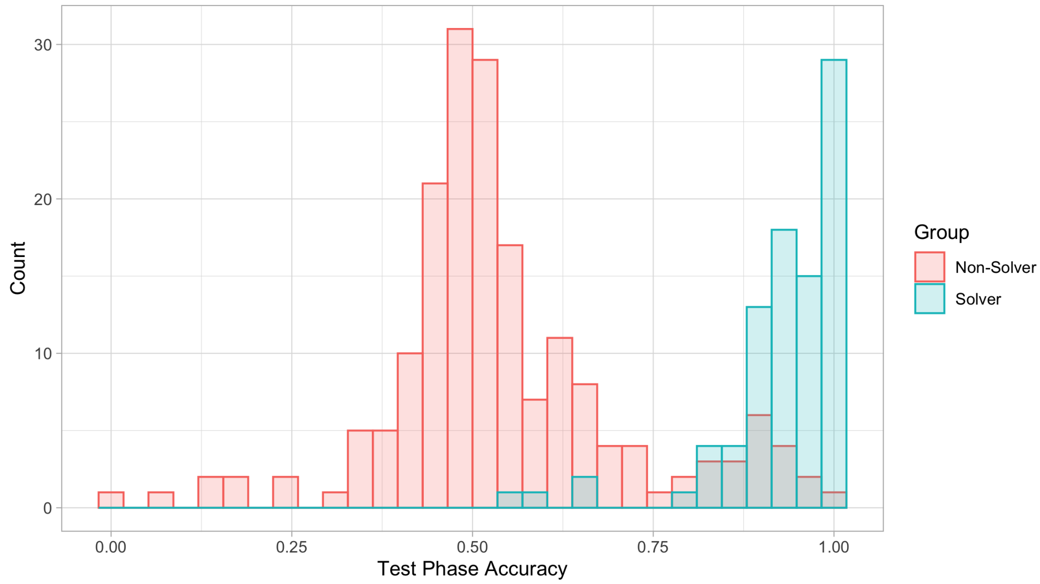


Figure 13: Histogram of overall accuracy during the test phase.

7.2 Test Phase Results

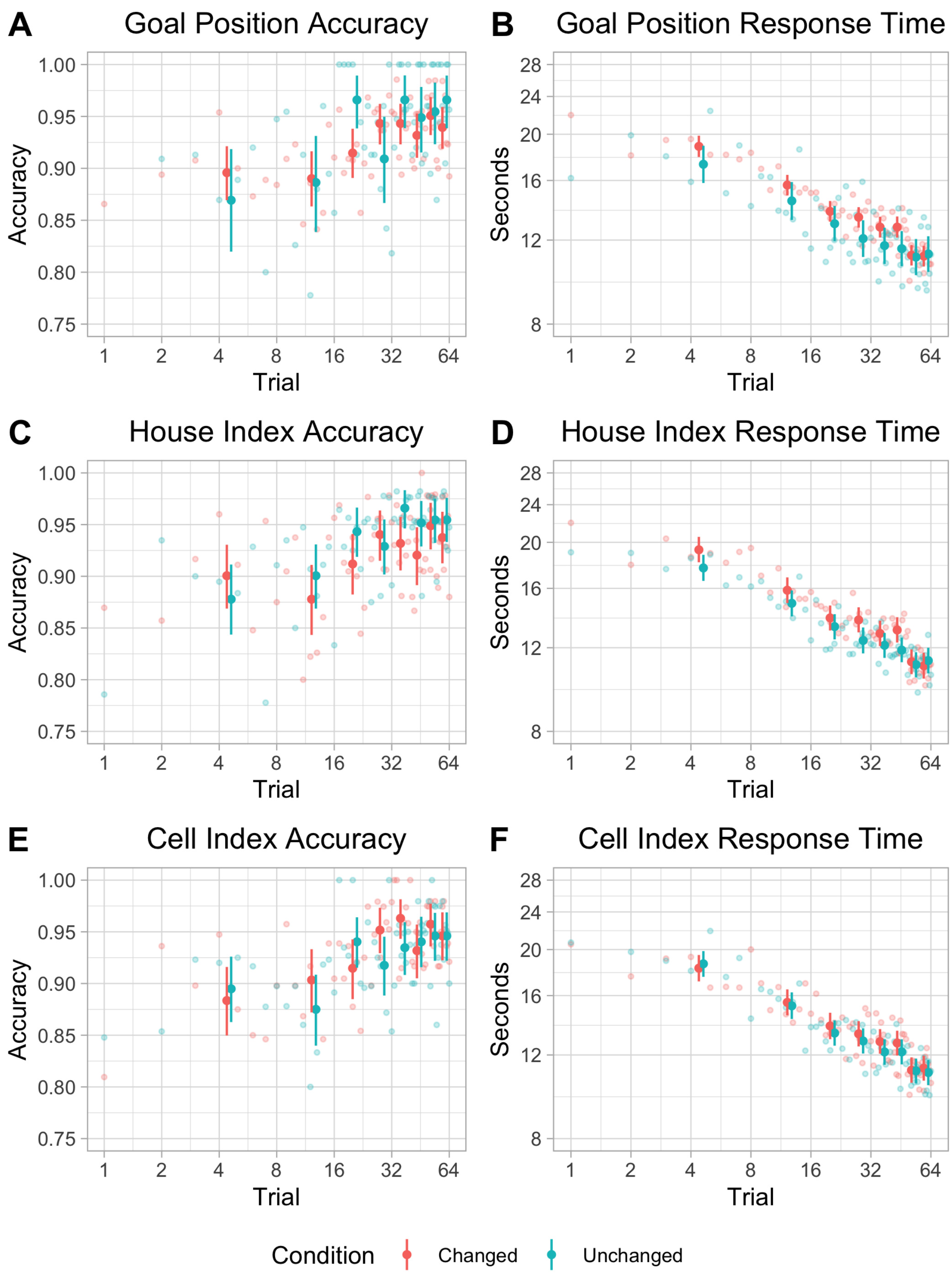


Figure 14: Mean accuracy and response times. Darker points indicate means and error bars indicate 95% highest density intervals for sets of 8 trials. Lighter points indicate means at individual trials. Values for response times computed in log-space. Note that each participant only had 16 puzzles with goal positions unchanged compared to the 48 puzzles with goal positions changed. Only trials with correct responses included for response time plots.