

HyperStorylines: Interactively Untangling Dynamic Hypergraphs

Report of CI values and plots

May 3th, 2021

1. Data wrangling

1_aggregate_data.R -> Load the file that contains the answers (user_answers_ms.csv) and creates two files with the aggregated data, one for the training trials and one for the main ones.

2. Analysis for all tasks

2.1. Scripts

Final plots are stored in plots/2_all_tasks/.
Inside that folder, intermediary results are in the raw/ folder.

For the three first scripts the phase (exploration [called main in the files] or advanced [called training in the files]) is selected manually at the top of the script. For the last script, the phases are selected automatically in a loop.

-2-all_tasks_CI_ERROR.R -> computes Confidence Intervals (mean and difference between tools) for ERROR for both phases for all tasks collectively.

-2-all_tasks_CI_TIME.R -> computes Confidence Intervals (mean and difference between tools) for TIME for both phases for all tasks collectively

-2-all_tasks_create_table_phases.R -> Creates table with plots of Confidence Intervals for TIME, ERROR, and means for CONF and DIFF for all tasks and all phases

2.2. CI values:

Exploratory Phase

	Technique	Mean time	Tech Compared	Mean Differences
All Tasks	PAOH HSL	93.39 [82.91, 120.85] 92.44 [73.03, 124.45]	PAOH-HSL	0.94 [-37.43, 33.28]

	Technique	Mean error	Tech Compared	Mean Differences
All Tasks	PAOH	8.33 [1.39, 12.5]	PAOH-HSL	-1.38 [-11.11, 4.17]

	HSL	9.72 [4.17, 12.5]	
--	-----	-------------------	--

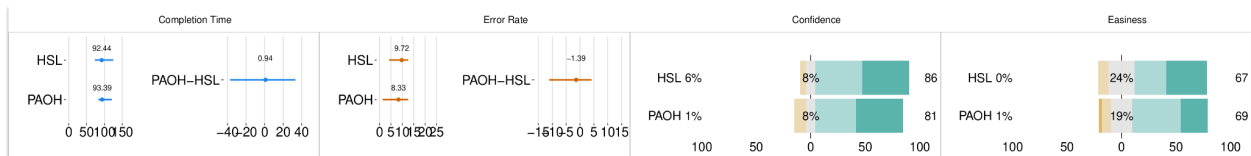
Advanced Phase

	Technique	Mean time	Tech Compared	Mean Differences
All Tasks	PAOH	84.67 [70.57, 99.72]	PAOH-HSL	17.51 [4.46, 35.59]
	HSL	67.16 [55.82, 78.87]		

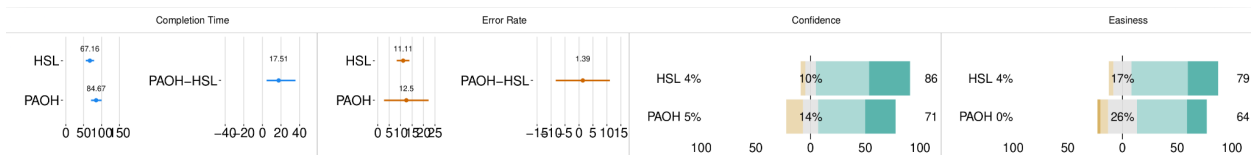
	Technique	Mean error	Tech Compared	Mean Differences
All Tasks	PAOH	12.5 [2.78, 22.22]	PAOH-HSL	1.39 [-8.33, 11.11]
	HSL	11.11 [8.33, 13.89]		

2.3. Plots

Exploratory Phase



Advanced Phase



3. Analysis per task

3.1. Scripts

The following scripts compute results per each individual task.

Final plots are stored in plots/3_per_task/.

Inside that folder, intermediary results are in the raw/ folder.

For all the following scripts the phase (exploration [called main in the files] or advanced [called training in the files]) is selected manually at the top of the script.

- 3-tasks_separated_CI_ERROR.R -> computes Confidence Intervals (mean and difference between tools) and plots for ERROR for both phases per task
- 3-tasks_separated_CI_TIME.R -> computes Confidence Intervals (mean and difference between tools) and plots for TIME for both phases
- 3-per_task_create_table.R -> Creates table with plots of Confidence Intervals for TIME, ERROR, and means for CONF and DIFF the results per task, per phase

3.2. CI Values

Exploratory Phase

	Technique	Mean time	Tech Compared	Mean Differences
Find Relationships	PAOH HSL	76, 01 [67.71, 81.28] 68.59 [56.93, 90.16]	PAOH-HSL	7.42 [-6.83, 20.04]
Characterize relationships	PAOH HSL	54.57 [35.3, 106.32] 62.97 [45.92, 91.49]	PAOH-HSL	-8.40 [-46.8, 37.23]
Find similar entities	PAOH HSL	101.34 [82.31, 132.07] 128.39 [102.99, 157.51]	PAOH-HSL	-27.05 [-62.96, 26.82]
Find massive events	PAOH HSL	141.62 [109.72, 168.37] 109.83 [78.33, 187.48]	PAOH-HSL	31.79 [-81.63, 78.42]

	Technique	Mean error	Tech Compared	Mean Differences
Find Relationships	PAOH	0 [0, 0]	PAOH-HSL	-55.56 [-27.78, 0]
	HSL	5.56 [0, 11.11]		
Characterize relationships	PAOH	5.56 [0, 11.11]	PAOH-HSL	-5.56 [0, -27.78]
	HSL	11.11 [0, 22.22]		
Find similar entities	PAOH	16.67 [0, 22.22]	PAOH-HSL	5.56 [-22.22, 16.67]
	HSL	11.11 [1, 22.22]		
Find massive events	PAOH	11.11 [0, 22.22]	PAOH-HSL	0 [-22.22, 11.11]
	HSL	11.11 [0, 22.22]		

Advanced Phase

	Technique	Mean time	Tech Compared	Mean Differences
Find Relationships	PAOH	69.8 [51.73, 91.39]	PAOH-HSL	23.95 [0.91, 45.02]
	HSL	45.84 [36.81, 57.46]		
Characterize relationships	PAOH	80.6 [61.91, 95.48]	PAOH-HSL	32.69 [17.91, 48.32]
	HSL	47.91 [41.73, 57.89]		
Find similar entities	PAOH	97.68 [78.96, 121.42]	PAOH-HSL	2.74 [-28.25, 36.95]
	HSL	94.94 [74.79, 121.42]		
Find massive events	PAOH	90.61 [74.46, 106.75]	PAOH-HSL	10.67 [-15.74, 20.91]
	HSL	79.94 [59.61, 92.05]		

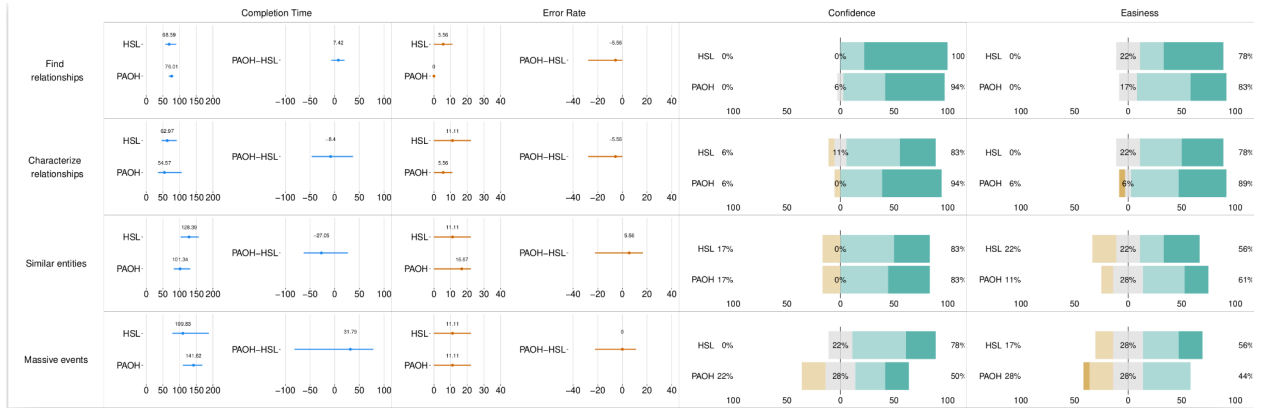
	Technique	Mean error	Tech Compared	Mean Differences
Find Relationships	PAOH	11.11 [0, 22.22]	PAOH-HSL	11.11 [0, 22.22]
	HSL	0 [0, 0]		
Characterize relationships	PAOH	22.22 [0, 38.89]	PAOH-HSL	5.55 [-22.22, 33.33]
	HSL	16.67 [0, 22.22]		
Find similar entities	PAOH	16.67 [0, 22.22]	PAOH-HSL	5.55[-16.67, 16.67]
	HSL	11.11 [0, 22.22]		
Find massive events	PAOH	0 [0, 0]	PAOH-HSL	-16.67 [-33.33, -11.11]
	HSL	16.67 [0, 22.22]		

3.3. Plots

Exploratory Phase



Advanced Phase



4. Analysis of learning effect per task

4.1. Scripts

- 4-per_task_per_phase_TIME.R -> compute Confidence Intervals for TIME for each phase for each vis
- 4-per_task_per_phase_TIME_table.R -> Merge Confidence Intervals for TIME per phase for each vis in a table
- 4-per_task_per_phase_ERROR.R -> compute Confidence Intervals for ERROR for each phase for each vis
- 4-per_task_per_phase_ERROR_table.R -> Merge Confidence Intervals for ERROR per phase for each vis in a table

Final plots are stored in plots/4_per_task_learning/.
Inside that folder, intermediary results are in the raw/ folder.

4.2. CI Values

HyperStorylines

	Phase	Mean time	Comparison	Mean Differences
Find Relationships	Adv	45.84 [36.81, 57.46]	Adv-Explo	-22.85 [-33.46, -15.30]
	Explo	68.69 [56.59, 90.16]		
Characterize relationships	Adv	47.91 [41.73, 57.89]	Adv-Explo	-16.10 [-53.62, -2.09]
	Explo	64.01 [46.24, 95.66]		
Find similar entities	Adv	94.94 [74.79, 121.42]	Adv-Explo	-32.33 [-47.93, -15.45]
	Explo	127.27 [104.98, 156.98]		
Find massive events	Adv	79.94 [59.61, 92.05]	Adv-Explo	-23.95 [-76.49, -2.7]
	Explo	103.89 [76.1, 168.14]		

	Phase	Mean error	Comparison	Mean Differences
Find Relationships	Adv	0 [0, 0]	Adv-Explo	-2.78 [-13.89, 0]
	Explo	2.78 [0, 5.56]		
Characterize relationships	Adv	16.67 [0, 22.22]	Adv-Explo	8.33 [-16.67, 25]
	Explo	8.33 [0, 19.44]		

Find similar entities	Adv	11.11 [0, 22.22]	Adv-Explo	5.56 [-8.33, 16.67]
	Explo	5.56 [0, 11.11]		
Find massive events	Adv	16.67 [0, 22.22]	Adv-Explo	11.11 [-11.11, 25]
	Explo	5.56 [0, 11.11]		

PAOHvis

	Phase	Mean time	Comparison	Mean Differences
Find Relationships	Adv	69.8 [51.73, 91.39]	Adv-Explo	-6.22 [-26.56, 15.35]
	Explo	76.01 [67.71, 81.28]		
Characterize relationships	Adv	80.6 [61.91, 95.48]	Adv-Explo	26.03 [-11.67, 49.07]
	Explo	54.57 [35.3, 106.32]		
Find similar entities	Adv	97.68 [78.96, 121.43]	Adv-Explo	-3.72 [-22.25, 20.57]
	Explo	101.4 [85.02, 136.81]		
Find massive events	Adv	90.60 [74.46, 106.75]	Adv-Explo	-56.10 [-72.54, -28.47]
	Explo	146.71 [124.96, 170.52]		

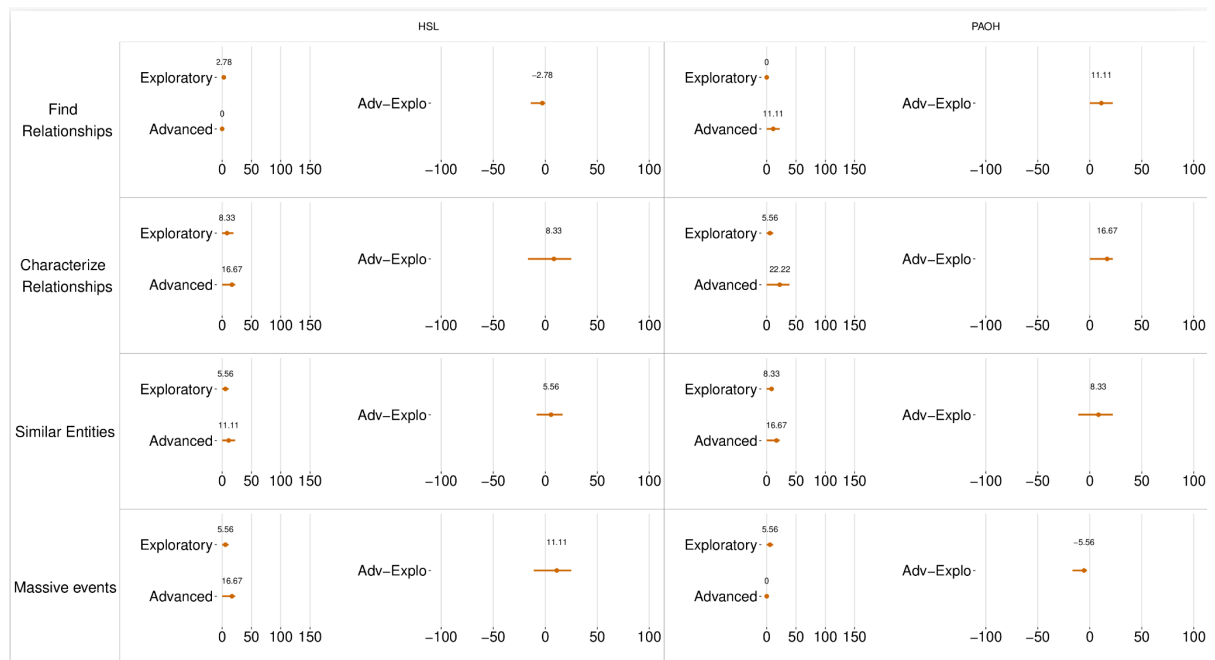
	Phase	Mean error	Comparison	Mean Differences
Find Relationships	Adv	11.11 [0, 22.22]	Adv-Explo	11.11 [0, 22.22]
	Explo	0 [0, 0]		
Characterize relationships	Adv	22.22 [0, 38.89]	Adv-Explo	16.67 [0, 22.22]
	Explo	55.56 [0, 11.11]		
Find similar entities	Adv	16.67 [0, 22.22]	Adv-Explo	8.33 [-11.11, 22.22]
	Explo	8.33 [0, 11.11]		
Find massive events	Adv	0 [0, 0]	Adv-Explo	-5.56 [-16.67, -2.78]
	Explo	5.56 [0, 11.11]		

4.3. Plots

Completion Time



Error Rate



5. Analysis of learning effect for all tasks collectively

5.1. Scripts

- 5-all_tasks_learning_ERROR.R -> compute Confidence Intervals for ERROR for all tasks collectively comparing between Exploratory and Advanced phase
- 5-all_tasks_learning_TIME.R -> compute Confidence Intervals for TIME for all tasks collectively comparing between Exploratory and Advanced phase
- 5-time_error_all_tasks_learning_table.R -> Creates a table with the results above

Results are stored in plots/5_all_tasks_learning/.

Inside that folder, intermediary results are in the raw/ folder.

5.2. CI Values

HyperStorylines

	Phase	Mean time	Comparison	Mean Differences
All Tasks	Adv Explo	67.16 [55.82, 7887] 92.44 [73.03, 124.45]	Adv-Explo	-25.29 [-46.87, -12.78]

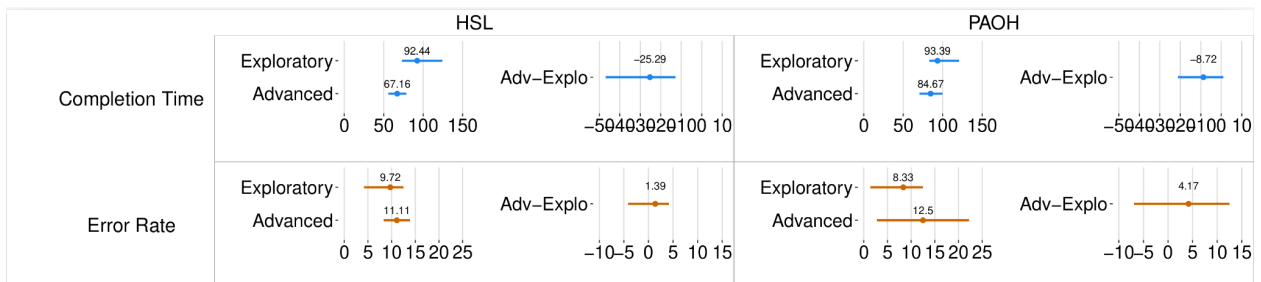
	Phase	Mean error	Comparison	Mean Differences
All Tasks	Adv Explo	11.11 [8.33, 13.89] 9.72 [4.17, 12.5]	Adv-Explo	1.39 [-4.17, 4.17]

PAOHvis

	Phase	Mean time	Comparison	Mean Differences
All Tasks	Adv	84.67 [70.57, 99.73]	Adv-Explo	-8.72 [-21.13, 0.98]
	Explo	93.39 [82.91, 120.85]		

	Phase	Mean error	Comparison	Mean Differences
All Tasks	Adv	12.5 [2.78, 22.22]	Adv-Explo	4.17 [-6.94, 12.5]
	Explo	8.33 [1.39, 12.5]		

5.3. Plots



6. Analysis per task complexity

6.1 Scripts

- 6-per_task_per_complexity_CI.R -> Computes Confidence Intervals for TIME and ERROR for the 3 levels of tasks complexity and creates a table with all the results

Results are stored in plots/6_per_complexity/.

Inside that folder, intermediary results are in the raw/ folder.

6.2. CI Values

People + Time

	Technique	Mean time	Tech Compared	Mean Differences
Find Relationships	PAOH	42, 68 [38.84, 46.55]	PAOH-HSL	-1.79 [-20.78, 13.25]
	HSL	44.48 [30.82, 61.10]		
Characterize relationships	PAOH	79.78 [59.83, 123.14]	PAOH-HSL	22.61[1.52 86.37]
	HSL	57.17 [49.10, 66.93]		
Find similar entities	PAOH	108.67 [76.28, 151.04]	PAOH-HSL	-33.39 [-117.2, 50.51]
	HSL	142.05 [101.58, 214.26]		
Find massive events	PAOH	127.32 [82.71, 166.15]	PAOH-HSL	26.14 [-35.30, 66.70]
	HSL	101.18 [78.09, 124.24]		

	Technique	Mean error	Tech Compared	Mean Differences
Find Relationships	PAOH	0 [0,0]	PAOH-HSL	0 [0,0]
	HSL	0 [0,0]		
Characterize relationships	PAOH	16.67 [0, 33.33]	PAOH-HSL	0 [-66.67, 33.33]
	HSL	16.67 [0, 33.33]		
Find similar entities	PAOH	33.33 [0, 66.67]	PAOH-HSL	33.33 [0, 66.66]
	HSL	0 [0,0]		
Find massive events	PAOH	0 [0,0]	PAOH-HSL	-50 [-100, -16.66]
	HSL	50 [0, 66.67]		

People + Locations

	Technique	Mean time	Tech Compared	Mean Differences
Find Relationships	PAOH	64.64 [48.16, 81.64]	PAOH-HSL	19.88 [-18.19, 37.77]
	HSL	44.76 [33.46, 60.56]		
Characterize relationships	PAOH	101.85 [69.93, 122.04]	PAOH-HSL	63.7 [31.97, 91.22]
	HSL	38.16 [30.6, 54.74]		
Find similar entities	PAOH	69.28 [49.02, 90.65]	PAOH-HSL	12.22 [-12.0, 26.72]
	HSL	57.06 [42.65, 73.03]		
Find massive events	PAOH	62.50 [53.47, 74.21]	PAOH-HSL	15.84 [4.56, 24.17]
	HSL	46.67 [32.54, 56.23]		

	Technique	Mean error	Tech Compared	Mean Differences
Find Relationships	PAOH	0 [0,0]	PAOH-HSL	16.66 [0,50]
	HSL	16.67 [0, 33.33]		
Characterize relationships	PAOH	33.33 [0, 66.67]	PAOH-HSL	33.33 [0, 66.67]
	HSL	0 [0,0]		
Find similar entities	PAOH	16.67 [0, 33.33]	PAOH-HSL	16.66 [0,50]
	HSL	0 [0,0]		
Find massive events	PAOH	0 [0,0]	PAOH-HSL	0 [0, 0]
	HSL	0 [0,0]		

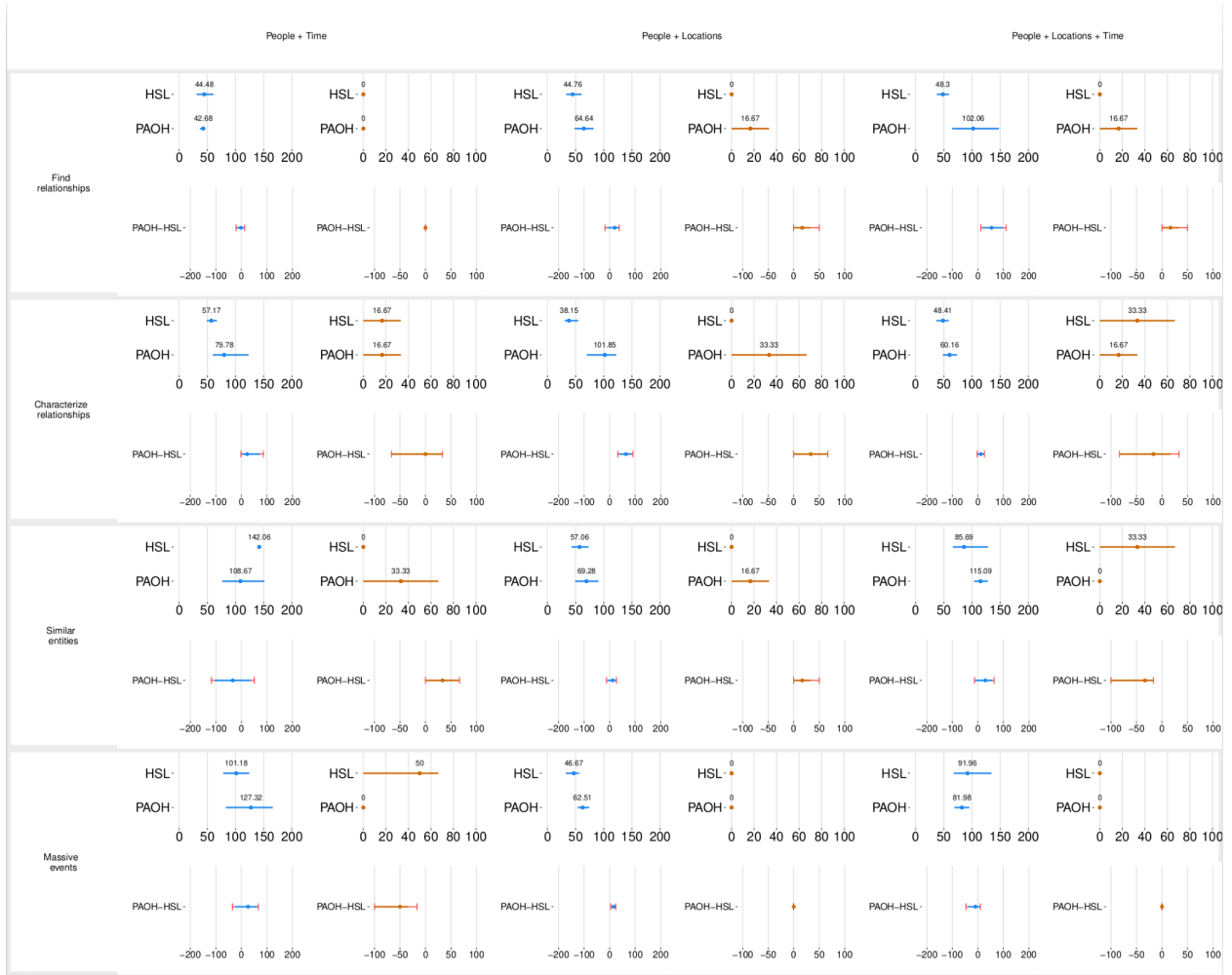
People + Locations + Time

	Technique	Mean time	Tech Compared	Mean Differences
Find Relationships	PAOH	102.06 [64.7, 147.67]	PAOH-HSL	53.76 [12.51, 110.09]
	HSL	48.3 [37.48, 59.10]		
Characterize relationships	PAOH	60.16 [48.21, 73.43]	PAOH-HSL	11.75 [-3.09, 25.88]
	HSL	48.41 [37.48, 58.67]		
Find similar entities	PAOH	115.09 [103.76, 127.93]	PAOH-HSL	29.4 [-12.77,63.30]
	HSL	85.69 [66.13, 128.44]		

Find massive events	PAOH	81.98 [68.59, 95.34]	PAOH-HSL	-9.97 [-46.85, 10.01]
	HSL	91.96 [67.28, 134.48]		

	Technique	Mean error	Tech Compared	Mean Differences
Find Relationships	PAOH	16.67 [0, 33.33]	PAOH-HSL	16.66 [0,50]
	HSL	0 [0,0]		
Characterize relationships	PAOH	16.67 [0, 33.33]	PAOH-HSL	-16.66 [-83.33, 33.33]
	HSL	33.33 [0, 66.67]		
Find similar entities	PAOH	0 [0,0]	PAOH-HSL	-33.33 [-100,-16.67]
	HSL	33.33 [0, 66.67]		
Find massive events	PAOH	0 [0,0]	PAOH-HSL	0 [0, 0]
	HSL	0 [0,0]		

6.3. Plots



7. Analysis of screen size

7.1. Scripts

- 7-screensize.R -> script that contains a set of subscripts that generates 3 plots per phase:
 - Means of all the metrics for all tasks collectively per participant
 - Means of ERROR per task per participant
 - Means of Completion Time per task per participant.

The phase to analyze is selected at the top of the file (user main for advanced phase, and training for exploratory phase).

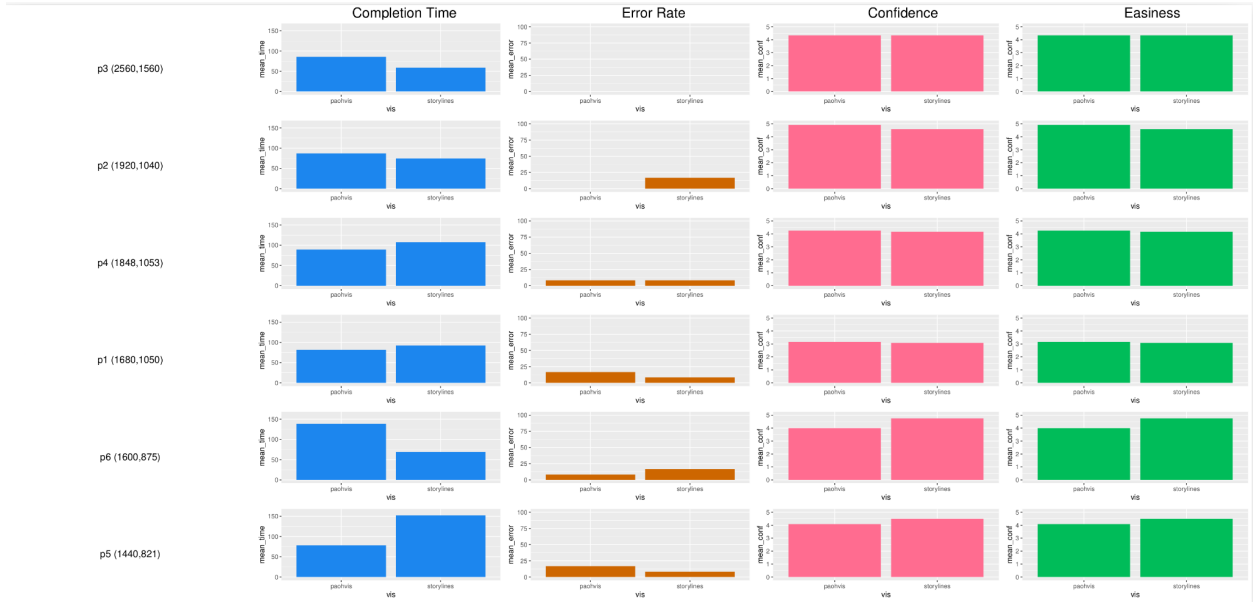
Final plots are stored in plots/7_screen_size/.

There is no CI values analysis for the analysis of the impact of screen size.

7.2. Plots

Exploratory Phase

All tasks:



Completion Time per task

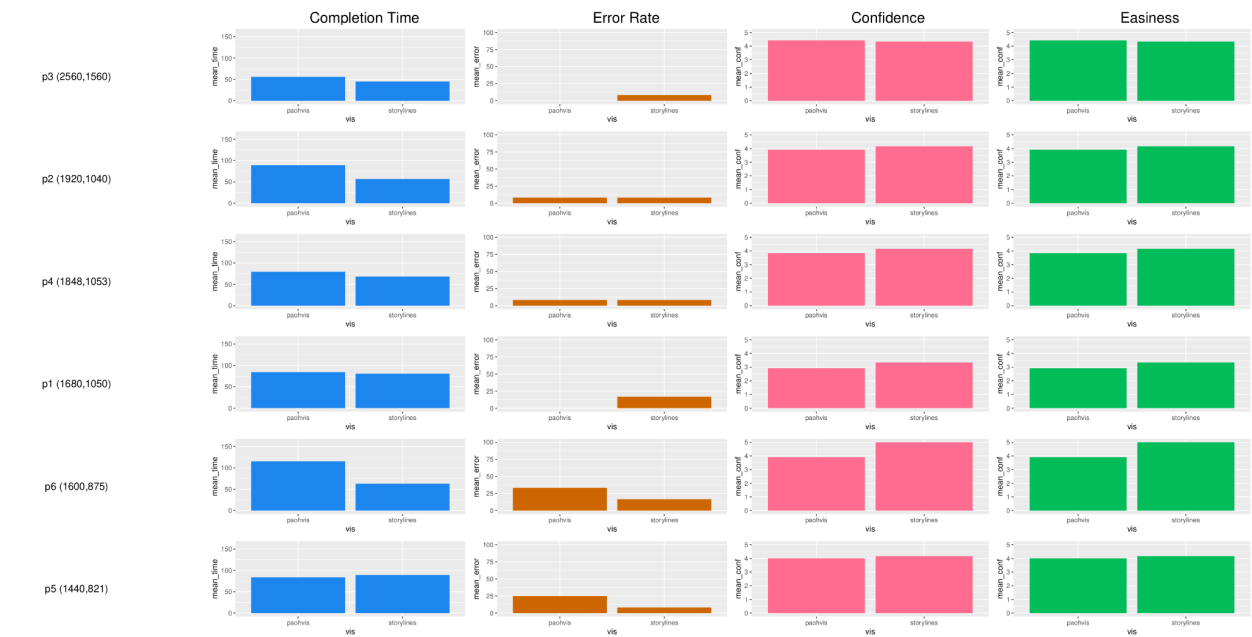


Error Rate per task



Advanced Phase

All tasks



Completion Time per task



Error Rate per task

