

# Report of CI values, plots and R scripts

## D.1 Data wrangling

`1_aggregate_data.R` => Load the file that contains the answers (`user_answers_ms.csv`) and creates one files with the aggregated data.

## D.2 Analysis for all tasks

### Scripts

The following scripts compute results for all task collectively. Final plots are stored in `plots/all_tasks/`. Inside that folder, intermediary results are in the `raw/` folder.

- **`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.R`** => Creates table with plots of Confidence Intervals for TIME, ERROR, and means for CONF and DIFF for all tasks.
- **`2-all_tasks_create_latex_table.R`** => Creates a latex table with the confidence intervals for time and error.

### CI Values

Table D.1: CI values for completion time for all tasks collectively.

	Mean times	Mean Time diffs
All tasks	TimeG 33.55[30.65,37.51]	TimeG-CoordV 10.34[7.92, 12.66]
	CoordV 23.20[20.85,26.24]	Time-MapG -2.01[-5.55, 0.83]
	MapG 35.55[31.79,42.49]	CoordV-MapG -12.35[-17.75, -8.77]

Table D.2: CI values for error rate for all tasks collectively.

	Mean Errors	Mean Error diffs
All tasks	TimeG 3.33[1.67, 5.00]	TimesG-CoordV 1.94[-0.83, 4.17]
	CoordV 1.39[0.28, 2.78]	TimesG-MapG 1.67[-0.28, 3.33]
	MapG 1.67[0.56, 3.06]	CoordV-MapG -0.28[-2.50, 1.39]

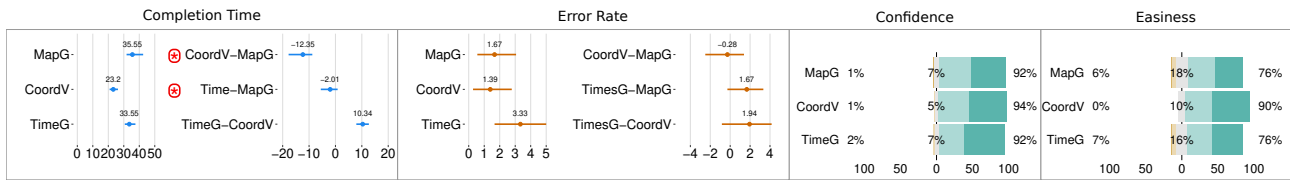


Figure D.1: Results for **Completion Time** (sec), **Error Rate** (in %) and mean self-reported easiness to complete per visualization per task. In each row (task) for the first two measures, mean values per visualization are seen on the left and means of pairwise differences on the right. Error bars represent 95% Bootstrap confidence intervals. Gray rectangles indicate the direction of our hypotheses. Evidence of differences are marked with a  $\oplus$  (the further away from 0 and the tighter the CI, the stronger the evidence). Inconclusive evidence (barely touching 0) are not reported in the text, but marked with a  $\circ$  as they may represent trends. The Third column shows the percentage of trials that participants reported being from **Very easy** (■) to complete, to **very hard** (■). The fourth column shows the percentage of trials that participants reported being from: **Very easy** (■) to: **very hard** (■) to complete.

## D.3 Analysis per task

### Scripts

The following scripts compute results per each individual task. Final plots are stored in `plots/per_task/`. Inside that folder, intermediary results are in the `raw/` folder.

- **3-tasks\_separate\_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.
- **3-per\_task\_create\_latex\_table.R** => Creates a latex table with the confidence intervals for time and error per task.

## CI Values

Table D.3: CI values for completion time per task

Task	Mean times	Mean Time diffs
Base Locations	TimeG 22.24[19.26,26.26]	TimeG-CoordV 6.95[4.39, 10.25]
	CoordV 15.29[13.65,18.31]	TimeG-MapG 3.73[0.27, 7.11]
	MapG 18.51[15.96,25.08]	CoordV-MapG -3.22[-6.18, -1.01]
Base Time	TimeG 28.35[25.88,31.51]	TimeG-CoordV 3.94[1.12, 6.79]
	CoordV 24.41[22.00,26.91]	TimeG-MapG -22.07[-30.34, -15.79]
	MapG 50.41[44.05,61.15]	CoordV-MapG -26.01[-35.44, -19.09]
Distance	TimeG 25.73[22.17,30.40]	TimeG-CoordV 8.37[5.23, 12.12]
	CoordV 17.35[15.90,18.99]	TimeG-MapG 0.28[-3.12, 3.53]
	MapG 25.44[21.76,31.85]	CoordV-MapG -8.09[-13.61, -4.54]
Direction	TimeG 46.07[42.15,51.64]	TimeG-CoordV 16.40[10.76, 21.69]
	CoordV 29.66[25.11,35.72]	TimeG-MapG 6.68[1.75, 11.55]
	MapG 39.39[34.76,44.98]	CoordV-MapG -9.73[-15.09, -3.47]
Adjacency	TimeG 45.35[40.51,51.62]	TimeG-CoordV 16.05[11.39, 21.05]
	CoordV 29.30[25.77,34.56]	TimeG-MapG 1.57[-3.99, 6.14]
	MapG 43.77[39.35,51.82]	CoordV-MapG -14.47[-22.40, -9.48]

Table D.4: CI values for error rate per task

Task	Mean Errors	Mean Error diffs
Base Locations	TimeG 0.00[0.00, 0.00]	TimeG-CoordV -2.78[-11.11, -1.39]
	CoordV 2.78[0.00, 6.94]	TimeG-MapG 0.00[0.00, 0.00]
	MapG 0.00[0.00, 0.00]	CoordV-MapG 2.78[0.00, 6.94]
Base Time	TimeG 4.17[0.00, 8.33]	TimeG-CoordV 4.17[0.00, 8.33]
	CoordV 0.00[0.00, 0.00]	TimeG-MapG 0.00[-12.50, 5.56]
	MapG 4.17[0.00, 12.50]	CoordV-MapG -4.17[-18.06, -1.39]
Distance	TimeG 5.56[1.39, 9.72]	TimeG-CoordV 4.17[-2.78, 8.33]
	CoordV 1.39[0.00, 4.17]	TimeG-MapG 5.56[1.39, 9.72]
	MapG 0.00[0.00, 0.00]	CoordV-MapG 1.39[0.00, 4.17]
Direction	TimeG 5.56[1.39, 9.72]	TimeG-CoordV 2.78[-4.17, 6.94]
	CoordV 2.78[0.00, 5.56]	TimeG-MapG 2.78[-4.17, 6.94]
	MapG 2.78[0.00, 5.56]	CoordV-MapG 0.00[-5.56, 2.78]
Adjacency	TimeG 1.39[0.00, 4.17]	TimeG-CoordV 1.39[0.00, 4.17]
	CoordV 0.00[0.00, 0.00]	TimeG-MapG 0.00[-5.56, 2.78]
	MapG 1.39[0.00, 4.17]	CoordV-MapG -1.39[-8.33, 0.00]

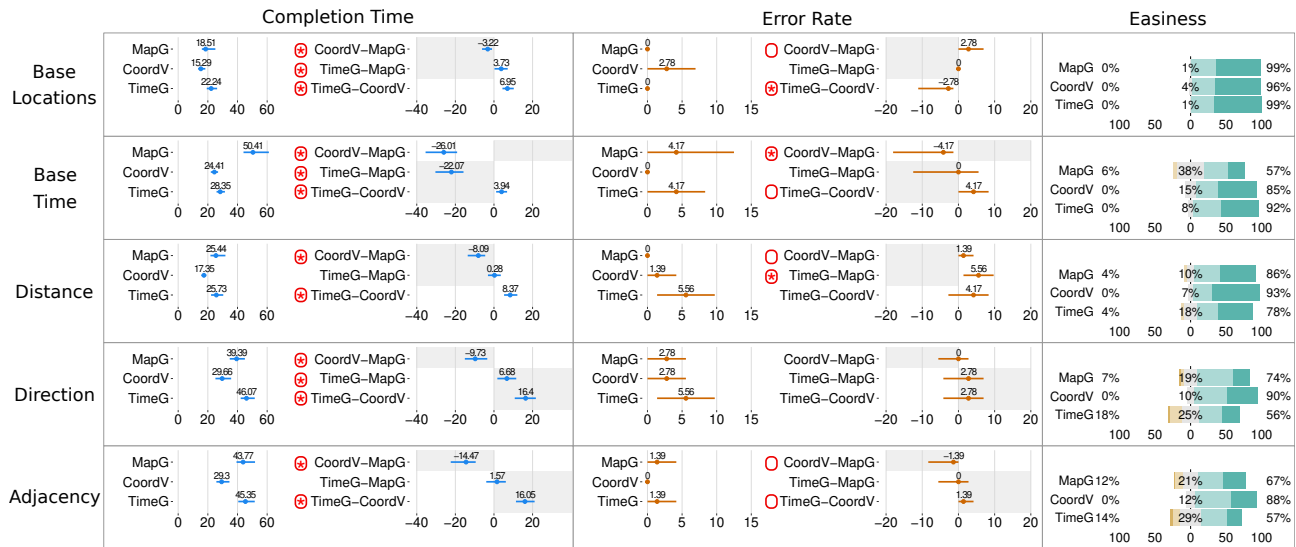


Figure D.2: Results for Completion Time (sec), Error Rate (in %) and mean self-reported easiness to complete per visualization per task. In each row (task) for the first two measures, mean values per visualization are seen on the left and means of pairwise differences on the right. Error bars represent 95% Bootstrap confidence intervals. Gray rectangles indicate the direction of our hypotheses. Evidence of differences are marked with a ⊕ (the further away from 0 and the tighter the CI, the stronger the evidence). Inconclusive evidence (barely touching 0) are not reported in the text, but marked with a ○ as they may represent trends. The Third column shows the percentage of trials that participants reported being from Very easy (■) to complete, to very hard (■).

## D.4 Analysis per map size for all tasks collectively

### Scripts

The following scripts compute results per each individual task. Final plots are stored in `plots/all_tasks_mapSize/`. Inside that folder, intermediary results are in the `raw/` folder.

- `4-all_tasks_mapSize_ERROR.R` => computes Confidence Intervals (mean and difference between tools) for ERROR per map size for all tasks collectively.
- `4-all_tasks_mapSize_TIME.R` => computes Confidence Intervals (mean and difference between tools) for TIME for both phases for all tasks collectively.
- `4-time_error_all_tasks_mapSize_table.R` => Creates table with plots of Confidence Intervals for TIME and ERROR for all tasks collectively per map.
- `4-time_error_all_tasks_mapSize_latex_table.R` => Creates a latex table with the confidence intervals for time and error.

## CI Values

Table D.5: CI values for completion time for all task collectively per map size.

Technique	Mean times	Mean Time diffs
Glyph Maps	big 35.57[31.76,41.03] small 35.48[31.18,44.04]	big-small 0.09[-3.21, 3.41]
Coordinated Views	big 22.53[20.30,25.38] small 23.87[21.29,27.39]	big-small -1.34[-3.02, 0.06]
Location Storylines	big 34.58[31.33,39.19] small 32.51[29.76,36.49]	big-small 2.07[0.28, 5.33]

Table D.6: CI values for error rate for all task collectively per map size.

Technique	Mean Errors	Mean Error diffs
Glyph Maps	big 1.67[0.00,5.00] small 2.22[0.56,3.89]	big-small -0.56[-3.89, 2.78]
Coordinated Views	big 0.00[0.00,0.00] small 2.78[0.56,5.56]	big-small -2.78[-7.22, -1.11]
Location Storylines	big 2.22[0.56,3.89] small 4.44[1.67,7.78]	big-small -2.22[-7.22, 0.56]

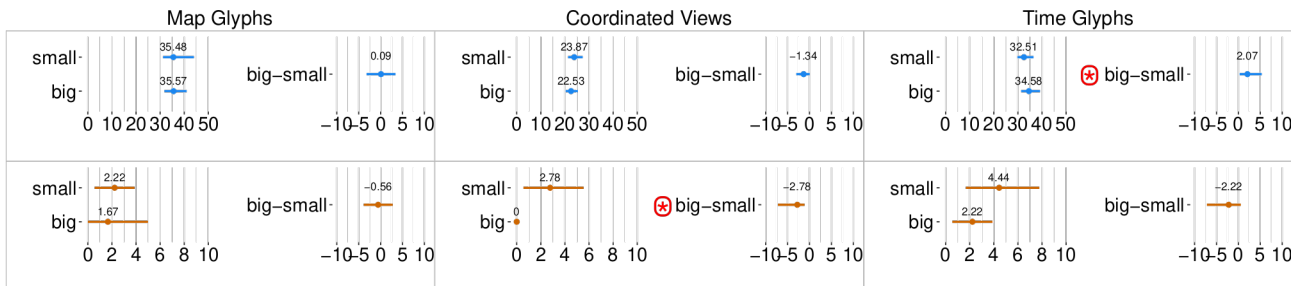


Figure D.3: Results for **Completion Time** (sec) and **Error Rate** (in %) for all tasks collectively per map size. In each row (task) for the first two measures, mean values per visualization are seen on the left and means of pairwise differences on the right. Error bars represent 95% Bootstrap confidence intervals. Gray rectangles indicate the direction of our hypotheses. Evidence of differences are marked with a  $\oplus$  (the further away from 0 and the tighter the CI, the stronger the evidence). Inconclusive evidence (barely touching 0) are not reported in the text, but marked with a  $\odot$  as they may represent trends.

## D.5 Analysis per map size per task

### Scripts

The following scripts compute results per each individual task. Final plots are stored in `plots/per_task_mapSize/`. Inside that folder, intermediary results are in the `raw/` folder.

- **5-per\_task\_mapSize\_TIME.R** => computes Confidence Intervals (mean and difference between tools) for TIME per map size for per task.
- **5-per\_task\_mapSize\_TIME\_table.R** => Creates table with plots of Confidence Intervals for TIME for per task per map.
- **5-per\_task\_mapSize\_TIME\_latex\_table.R** =>Creates a latex table with the confidence intervals for TIME per map size.
- **5-per\_task\_mapSize\_ERROR.R** => computes Confidence Intervals (mean and difference between tools) for ERROR per map size for per task.
- **5-per\_task\_mapSize\_ERROR\_table.R** => Creates table with plots of Confidence Intervals for ERROR for per task per map size.
- **5-per\_task\_mapSize\_ERROR\_latex\_table.R** =>Creates a latex table with the confidence intervals for ERROR per map size.

## CI Values for TIME

Table D.7: CI values for completion time per task per map size for Map Glyphs

Task	Mean Times & CIs	Mean Time Diffs & CIs
Base Locations	big 17.44[14.99,20.80] small 19.80[16.17,29.74]	big-small -2.36[-7.75, 1.28]
Base Time	big 46.91[40.11,54.50] small 53.92[45.23,69.44]	big-small -7.01[-17.80, 2.48]
Distance	big 28.27[23.46,37.10] small 22.62[19.11,28.09]	big-small 5.64[0.98, 11.56]
Direction	big 38.93[34.75,44.84] small 39.85[33.99,48.82]	big-small -0.92[-8.68, 5.17]
Adjacency	big 46.32[41.20,51.78] small 41.23[36.29,54.30]	big-small 5.09[-2.91, 9.51]

Table D.8: CI values for completion time per task per map size for Coordinated Views

Task	Mean Times & CIs	Mean Time Diffs & CIs
Base Locations	big 15.76[13.65,20.66] small 14.82[12.92,16.80]	big-small 0.95[-1.36, 4.34]
Base Time	big 20.92[19.22,22.95] small 27.89[24.03,32.32]	big-small -6.97[-11.56, -3.22]
Distance	big 18.31[16.57,20.19] small 16.40[14.92,18.20]	big-small 1.91[0.39, 3.42]
Direction	big 27.19[23.10,33.05] small 32.14[26.90,40.26]	big-small -4.95[-9.94, -1.02]
Adjacency	big 30.48[26.60,36.49] small 28.12[24.86,34.12]	big-small 2.35[-0.43, 6.03]

Table D.9: CI values for completion time per task per map size for Time Glyphs

Task	Mean Times & CIs	Mean Time Diffs & CIs
Base Locations	big 24.31[19.63,31.36] small 20.18[18.12,22.71]	big-small 4.14[0.01, 10.87]
Base Time	big 26.40[23.85,29.36] small 30.29[27.02,35.23]	big-small -3.89[-7.65, 0.17]
Distance	big 26.46[22.69,31.76] small 24.99[21.34,29.22]	big-small 1.47[-0.90, 3.37]
Direction	big 47.23[42.40,55.88] small 44.90[40.99,49.06]	big-small 2.33[-1.83, 9.04]
Adjacency	big 48.50[43.06,56.17] small 42.19[37.16,51.35]	big-small 6.31[-2.36, 12.11]

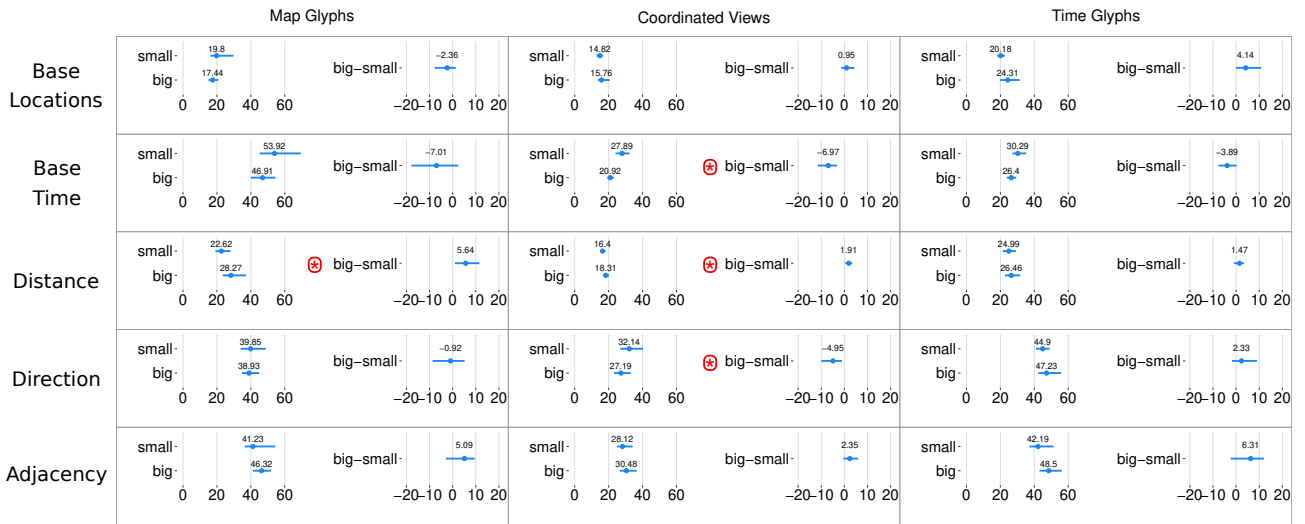


Figure D.4: Results for Completion Time (sec) and Error Rate (in %) for all tasks collectively per map size. In each row (task) for the first two measures, mean values per visualization are seen on the left and means of pairwise differences on the right. Error bars represent 95% Bootstrap confidence intervals. Gray rectangles indicate the direction of our hypotheses. Evidence of differences are marked with a ⊕ (the further away from 0 and the tighter the CI, the stronger the evidence). Inconclusive evidence (barely touching 0) are not reported in the text, but marked with a ○ as they may represent trends.

**CI Values for ERROR**

Table D.10: CI values for error rate per task per map size for Map Glyphs

Task	Mean Errors & CIs	Mean Error Diffs & CIs
Base Locations	Big 2.78[0.00,8.33]	Big-Small 2.78[0.00, 8.33]
	Small 0.00[0.00,0.00]	
Base Time	Big 5.56[0.00,16.67]	Big-Small 2.78[-5.56, 16.67]
	Small 2.78[0.00,8.33]	
Distance	Big 0.00[0.00,0.00]	Big-Small 0.00[0.00, 0.00]
	Small 0.00[0.00,0.00]	
Direction	Big 0.00[0.00,0.00]	Big-Small -5.56[-22.22, -2.78]
	Small 5.56[0.00,11.11]	
Adjacency	Big 0.00[0.00,0.00]	Big-Small -2.78[-16.67, 0.00]
	Small 2.78[0.00,8.33]	

Table D.11: CI values for error rate per task per map size for Coordinated Views

Task	Mean Errors & CIs	Mean Error Diffs & CIs
Base Locations	Big 0.00[0.00,0.00]	Big-Small -5.56[-22.22, -2.78]
	Small 5.56[0.00,13.89]	
Base Time	Big 0.00[0.00,0.00]	Big-Small 0.00[0.00, 0.00]
	Small 0.00[0.00,0.00]	
Distance	Big 0.00[0.00,0.00]	Big-Small -2.78[-19.44, 0.00]
	Small 2.78[0.00,8.33]	
Direction	Big 0.00[0.00,0.00]	Big-Small -5.56[-22.22, -2.78]
	Small 5.56[0.00,11.11]	
Adjacency	Big 0.00[0.00,0.00]	Big-Small 0.00[0.00, 0.00]
	Small 0.00[0.00,0.00]	

Table D.12: CI values for error rate per task per map size for Time Glyphs

Task	Mean Errors & CIs	Mean Error Diffs & CIs
Base Locations	Big 0.00[0.00,0.00]	Big-Small 0.00[0.00, 0.00]
	Small 0.00[0.00,0.00]	
Base Time	Big 0.00[0.00,0.00]	Big-Small -8.33[-25.00, -2.78]
	Small 8.33[0.00,16.67]	
Distance	Big 5.56[0.00,13.89]	Big-Small 0.00[-13.89, 8.33]
	Small 5.56[0.00,13.89]	
Direction	Big 5.56[0.00,11.11]	Big-Small 0.00[-13.89, 8.33]
	Small 5.56[0.00,13.89]	
Adjacency	Big 0.00[0.00,0.00]	Big-Small -2.78[-16.67, 0.00]
	Small 2.78[0.00,8.33]	



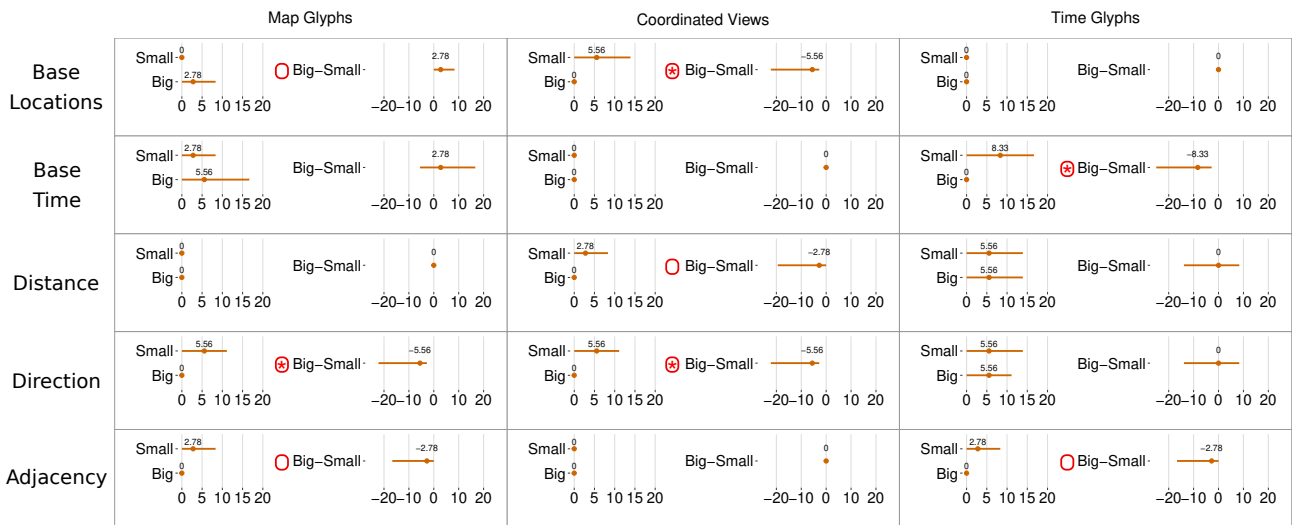


Figure D.5: Results for **Completion Time** (sec) and **Error Rate** (in %) for all tasks collectively per map size. In each row (task) for the first two measures, mean values per visualization are seen on the left and means of pairwise differences on the right. Error bars represent 95% Bootstrap confidence intervals. Gray rectangles indicate the direction of our hypotheses. Evidence of differences are marked with a ⊕ (the further away from 0 and the tighter the CI, the stronger the evidence). Inconclusive evidence (barely touching 0) are not reported in the text, but marked with a ○ as they may represent trends.