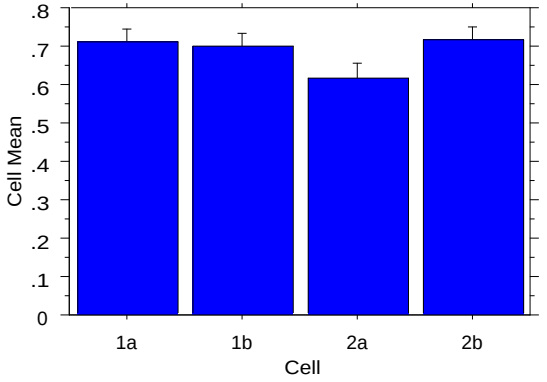
ANOVA Table for accuracy
Row exclusion: output_noLLM1 human data

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
CBType	3	1.210	.403	3.020	.0298	9.060	.708
Gender	1	.134	.134	1.005	.3168	1.005	.161
CBType * Gender	3	.406	.135	1.014	.3863	3.043	.267
Subject(Group)	373	49.816	.134				
problem type	1	55.402	55.402	396.480	<.0001	396.480	1.000
problem type * CBType	3	.504	.168	1.203	.3085	3.608	.313
problem type * Gender	1	.009	.009	.065	.7996	.065	.057
problem type * CBType * Gender	3	.390	.130	.930	.4261	2.791	.247
problem type * Subject(Group)	373	52.121	.140				

Means Table for accuracy
Effect: CBType
Row exclusion: output_noLLM1 human data

	Count	Mean	Std. Dev.	Std. Err.
1a	198	.712	.454	.032
1b	192	.698	.460	.033
2a	186	.618	.487	.036
2b	186	.715	.453	.033

Interaction Bar Plot for accuracy
Effect: CBType
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 human data



Means Table for accuracy
Effect: Gender
Row exclusion: output_noLLM1 human data

	Count	Mean	Std. Dev.	Std. Err.
Female	294	.667	.472	.028
Male	468	.699	.459	.021

Means Table for accuracy
Effect: CBType * Gender
Row exclusion: output_noLLM1 human data

	Count	Mean	Std. Dev.	Std. Err.
1a, Female	62	.726	.450	.057
1a, Male	136	.706	.457	.039
1b, Female	72	.708	.458	.054
1b, Male	120	.692	.464	.042
2a, Female	80	.588	.495	.055
2a, Male	106	.642	.482	.047
2b, Female	80	.662	.476	.053
2b, Male	106	.755	.432	.042

Means Table for accuracy
Effect: problem type
Row exclusion: output_noLLM1 human data

	Count	Mean	Std. Dev.	Std. Err.
standard	381	.407	.492	.025
control	381	.966	.182	.009

Means Table for accuracy
Effect: problem type * CBType
Row exclusion: output_noLLM1 human data

	Count	Mean	Std. Dev.	Std. Err.
1a, standard	99	.434	.498	.050
1a, control	99	.990	.101	.010
1b, standard	96	.448	.500	.051
1b, control	96	.948	.223	.023
2a, standard	93	.301	.461	.048
2a, control	93	.935	.247	.026
2b, standard	93	.441	.499	.052
2b, control	93	.989	.104	.011

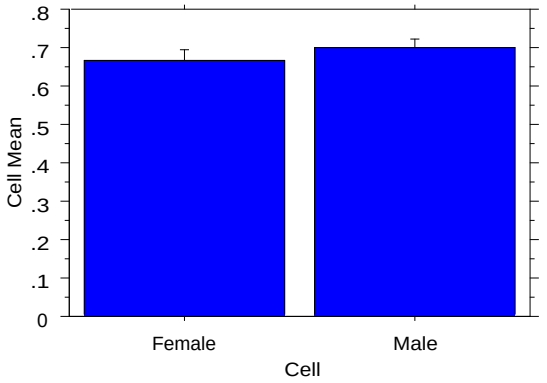
Means Table for accuracy
Effect: problem type * Gender
Row exclusion: output_noLLM1 human data

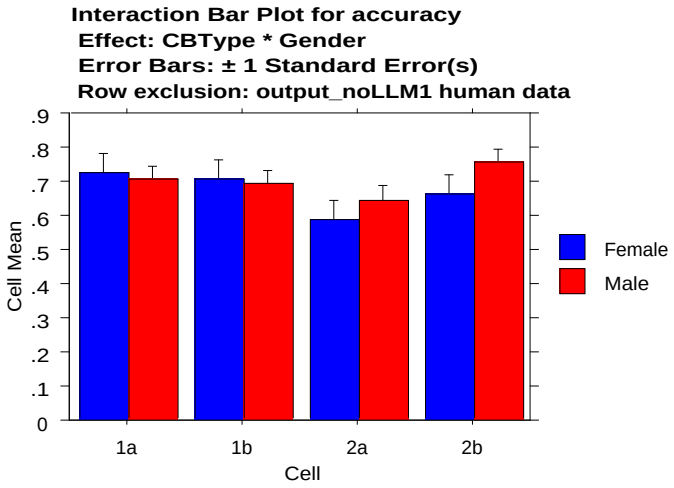
	Count	Mean	Std. Dev.	Std. Err.
Female, standard	147	.381	.487	.040
Female, control	147	.952	.214	.018
Male, standard	234	.423	.495	.032
Male, control	234	.974	.158	.010

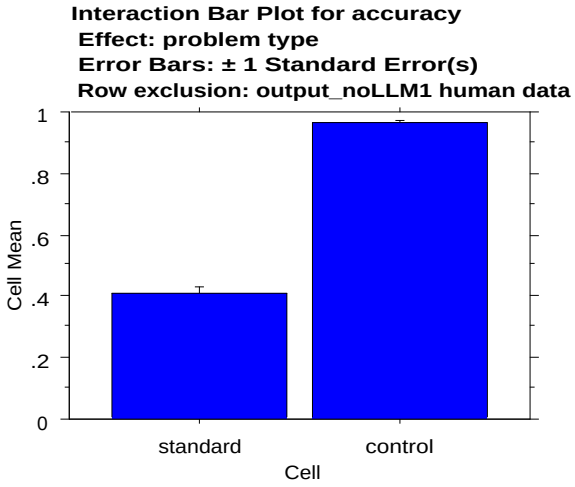
Means Table for accuracy
Effect: problem type * CBType * Gender
Row exclusion: output_noLLM1 human data

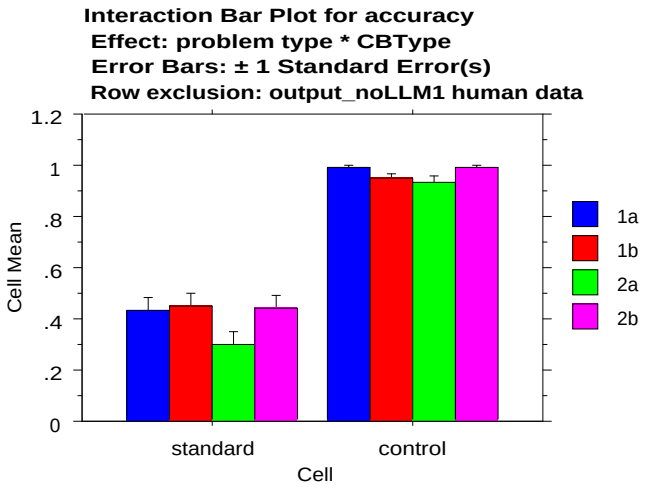
	Count	Mean	Std. Dev.	Std. Err.
1a, Female, standard	31	.484	.508	.091
1a, Female, control	31	.968	.180	.032
1a, Male, standard	68	.412	.496	.060
1a, Male, control	68	1.000	0.000	0.000
1b, Female, standard	36	.472	.506	.084
1b, Female, control	36	.944	.232	.039
1b, Male, standard	60	.433	.500	.065
1b, Male, control	60	.950	.220	.028
2a, Female, standard	40	.250	.439	.069
2a, Female, control	40	.925	.267	.042
2a, Male, standard	53	.340	.478	.066
2a, Male, control	53	.943	.233	.032
2b, Female, standard	40	.350	.483	.076
2b, Female, control	40	.975	.158	.025
2b, Male, standard	53	.509	.505	.069
2b, Male, control	53	1.000	0.000	0.000

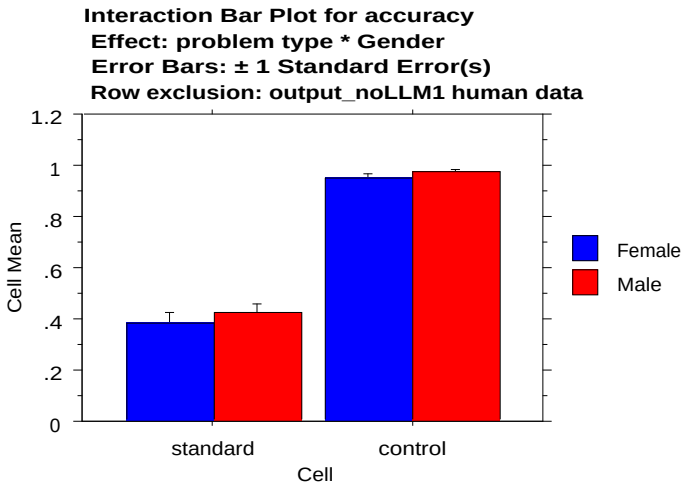
Interaction Bar Plot for accuracy
Effect: Gender
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 human data



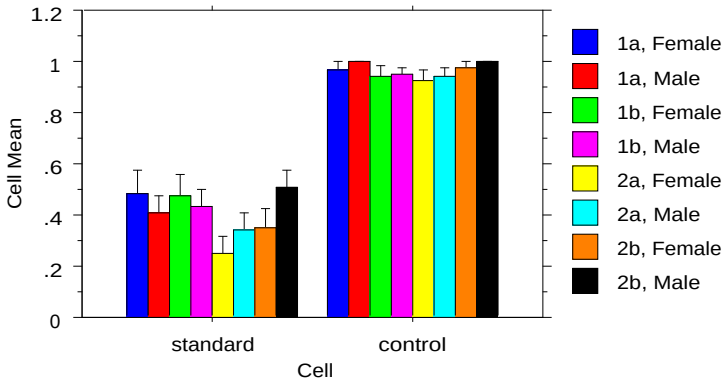

Interaction Bar Plot for accuracy
Effect: CBType * Gender
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 human data
Cell Mean
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
1a
1b
2a
2b
Cell
Female
Male


Interaction Bar Plot for accuracy
Effect: problem type
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 human data
Cell Mean
1
.8
.6
.4
.2
0
standard
control
Cell


Interaction Bar Plot for accuracy
Effect: problem type * CBType
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 human data
1.2
1
.8
.6
.4
.2
0
Cell Mean
standard
control
Cell
1a
1b
2a
2b


Interaction Bar Plot for accuracy
Effect: problem type * Gender
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 human data
1.2
1
.8
.6
.4
.2
0
Cell Mean
standard
control
Cell
Female
Male

Interaction Bar Plot for accuracy
Effect: problem type * CBType * Gender
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 human data



Fisher's PLSD for accuracy
Effect: CBType
Significance Level: 5 %
Row exclusion: output_noLLM1 human data

	Mean Diff.	Crit. Diff.	P-Value	
1a, 1b	.014	.073	.7014	
1a, 2a	.094	.073	.0123	S
1a, 2b	-.003	.073	.9374	
1b, 2a	.080	.074	.0348	S
1b, 2b	-.017	.074	.6488	
2a, 2b	-.097	.075	.0111	S

Fisher's PLSD for accuracy
Effect: Gender
Significance Level: 5 %
Row exclusion: output_noLLM1 human data

	Mean Diff.	Crit. Diff.	P-Value
Female, Male	-.032	.053	.2393

Fisher's PLSD for accuracy
Effect: problem type
Significance Level: 5 %
Row exclusion: output_noLLM1 human data

	Mean Diff.	Crit. Diff.	P-Value	
standard, control	-.559	.053	<.0001	S