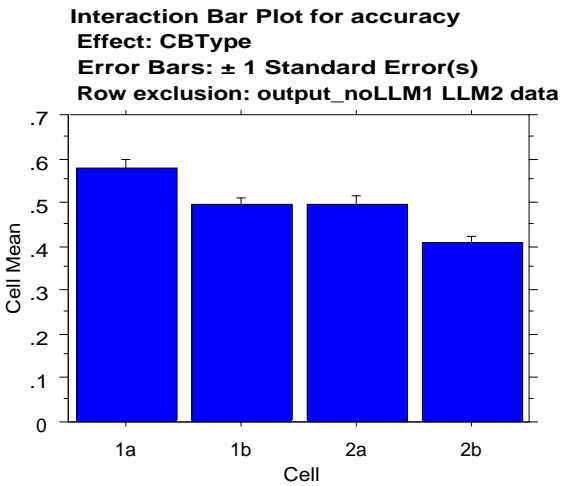


ANOVA Table for accuracy
Row exclusion: output_noLLM1 LLM2 data

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
CBType	3	10.751	3.584	26.168	<.0001	78.504	1.000
Gender	1	.063	.063	.459	.4980	.459	.101
CBType * Gender	3	.375	.125	.913	.4340	2.738	.245
Subject(Group)	1516	207.618	.137				
problem type	1	311.869	311.869	2406.680	<.0001	2406.680	1.000
problem type * CBType	3	16.675	5.558	42.894	<.0001	128.683	1.000
problem type * Gender	1	.012	.012	.091	.7627	.091	.060
problem type * CBType * Gender	3	.524	.175	1.348	.2572	4.043	.351
problem type * Subject(Group)	1516	196.450	.130				

Means Table for accuracy
Effect: CBType
Row exclusion: output_noLLM1 LLM2 data

	Count	Mean	Std. Dev.	Std. Err.
1a	792	.580	.494	.018
1b	768	.495	.500	.018
2a	744	.496	.500	.018
2b	744	.407	.492	.018



Means Table for accuracy
Effect: Gender
Row exclusion: output_noLLM1 LLM2 data

	Count	Mean	Std. Dev.	Std. Err.
Female	1176	.495	.500	.015
Male	1872	.496	.500	.012

Means Table for accuracy
Effect: CBType * Gender
Row exclusion: output_noLLM1 LLM2 data

	Count	Mean	Std. Dev.	Std. Err.
1a, Female	248	.581	.494	.031
1a, Male	544	.579	.494	.021
1b, Female	288	.517	.501	.029
1b, Male	480	.481	.500	.023
2a, Female	320	.509	.501	.028
2a, Male	424	.486	.500	.024
2b, Female	320	.394	.489	.027
2b, Male	424	.417	.494	.024

Means Table for accuracy
Effect: problem type
Row exclusion: output_noLLM1 LLM2 data

	Count	Mean	Std. Dev.	Std. Err.
standard	1524	.169	.375	.010
control	1524	.823	.382	.010

Means Table for accuracy
Effect: problem type * CBType
Row exclusion: output_noLLM1 LLM2 data

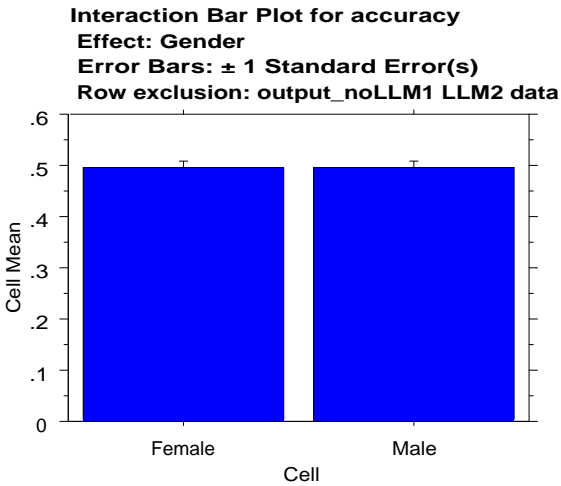
	Count	Mean	Std. Dev.	Std. Err.
1a, standard	396	.384	.487	.024
1a, control	396	.775	.418	.021
1b, standard	384	.133	.340	.017
1b, control	384	.857	.351	.018
2a, standard	372	.094	.292	.015
2a, control	372	.898	.303	.016
2b, standard	372	.051	.220	.011
2b, control	372	.763	.426	.022

Means Table for accuracy
Effect: problem type * Gender
Row exclusion: output_noLLM1 LLM2 data

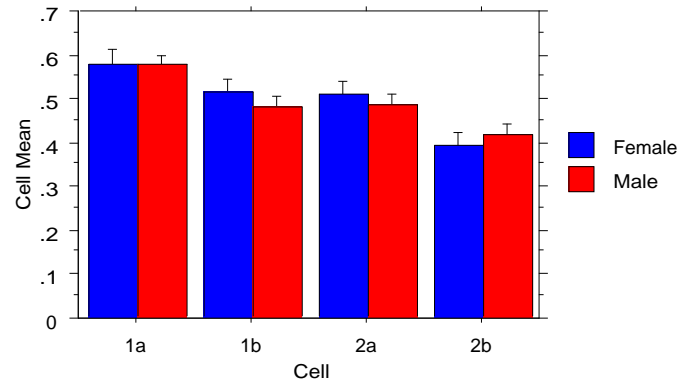
	Count	Mean	Std. Dev.	Std. Err.
Female, standard	588	.156	.364	.015
Female, control	588	.833	.373	.015
Male, standard	936	.176	.381	.012
Male, control	936	.816	.387	.013

Means Table for accuracy
Effect: problem type * CBType * Gender
Row exclusion: output_noLLM1 LLM2 data

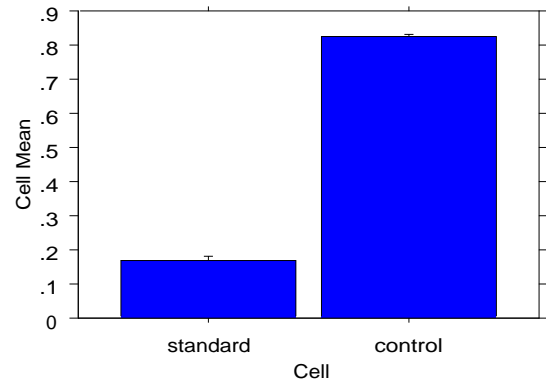
	Count	Mean	Std. Dev.	Std. Err.
1a, Female, standard	124	.371	.485	.044
1a, Female, control	124	.790	.409	.037
1a, Male, standard	272	.390	.489	.030
1a, Male, control	272	.768	.423	.026
1b, Female, standard	144	.132	.340	.028
1b, Female, control	144	.903	.297	.025
1b, Male, standard	240	.133	.341	.022
1b, Male, control	240	.829	.377	.024
2a, Female, standard	160	.112	.317	.025
2a, Female, control	160	.906	.292	.023
2a, Male, standard	212	.080	.272	.019
2a, Male, control	212	.892	.312	.021
2b, Female, standard	160	.056	.231	.018
2b, Female, control	160	.731	.445	.035
2b, Male, standard	212	.047	.213	.015
2b, Male, control	212	.788	.410	.028



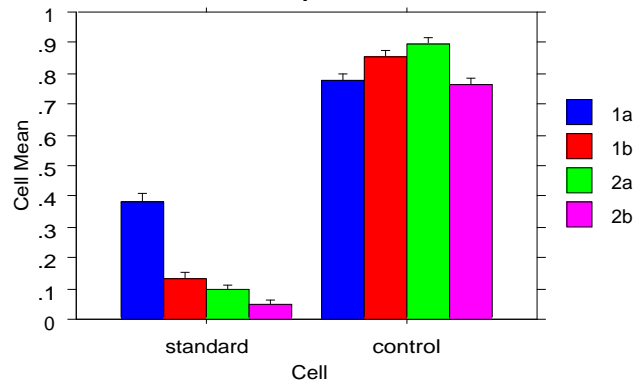
Interaction Bar Plot for accuracy
Effect: CBType * Gender
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 LLM2 data



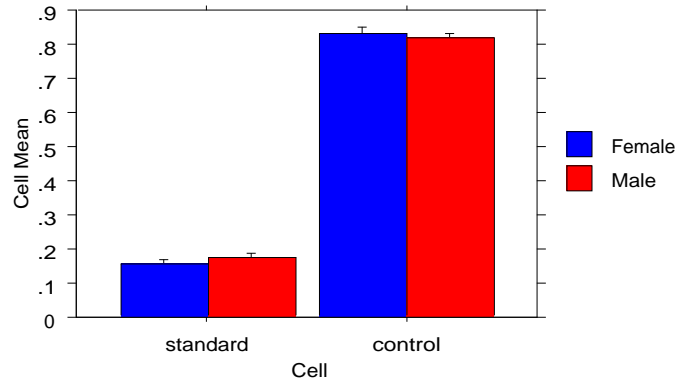
Interaction Bar Plot for accuracy
Effect: problem type
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 LLM2 data

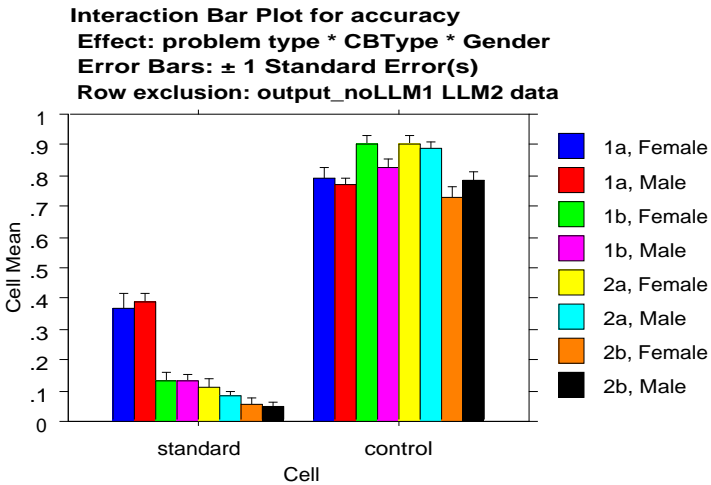


Interaction Bar Plot for accuracy
Effect: problem type * CBType
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 LLM2 data



Interaction Bar Plot for accuracy
Effect: problem type * Gender
Error Bars: ± 1 Standard Error(s)
Row exclusion: output_noLLM1 LLM2 data





Fisher's PLSD for accuracy
Effect: CBType
Significance Level: 5 %
Row exclusion: output_noLLM1 LLM2 data

	Mean Diff.	Crit. Diff.	P-Value	
1a, 1b	.085	.037	<.0001	S
1a, 2a	.084	.037	<.0001	S
1a, 2b	.172	.037	<.0001	S
1b, 2a	-.001	.037	.9507	
1b, 2b	.088	.037	<.0001	S
2a, 2b	.089	.038	<.0001	S

Fisher's PLSD for accuracy
Effect: Gender
Significance Level: 5 %
Row exclusion: output_noLLM1 LLM2 data

	Mean Diff.	Crit. Diff.	P-Value
Female, Male	-.001	.027	.9212

Fisher's PLSD for accuracy
Effect: problem type
Significance Level: 5 %
Row exclusion: output_noLLM1 LLM2 data

	Mean Diff.	Crit. Diff.	P-Value	
standard, control	-.654	.026	<.0001	S