

Table S1 Chi-squared test for question difficulty distribution

Difficulty level	Correct	Incorrect	Total	P value
Observed				<0.001
1	71	43	114	
2	63	79	142	
3	48	101	149	
4	18	49	67	
5	4	16	20	
Total	204	288	492	
Expected				<0.001
1	47	67	114	
2	59	83	142	
3	62	87	149	
4	28	39	67	
5	8	12	20	
Total	204	288	492	

Observed versus expected values for difficulty level distribution by test. Expected values were calculated from a chi-squared test using the 41% overall correct ratio (204/492).