Code Statistics

This document summarizes the statistics of the four dimensions of codes in our corpus of 1150 questions. Definitions of the codes can be found in "codebook.xlsx" in the same folder.

First Dimension

The first dimension is for filtering out questions that are not clear pairwise comparison questions. Cohen's Kappa for this dimension is 0.72. It has three categories: explicit, implicit and unrelated. The first column in the following table corresponds to the eight scenarios for soliciting pairwise comparison questions from workers on MTurk.

	Explicit	Implicit	Unrelated	
Analogical1	91	75	12	178
Analogical2	114	29	2	145
Decision1	61	48	42	151
Decision2	66	55	24	145
Knowledge1	78	34	32	144
Knowledge2	62	32	34	128
Social1	28	88	25	141
Social2	49	26	43	118
	549	387	214	1150

Group Dimension

The group dimension has three categories: one to one, one to multiple and multiple to multiple. Cohen's Kappa for this dimension is 0.76.

	One to one	One to multiple	Multiple to Multiple	
Analogical1	91	0	0	91
Analogical2	105	8	1	114
Decision1	37	24	0	61
Decision2	65	1	0	66
Knowledge1	78	0	0	78
Knowledge2	37	0	25	62
Social1	1	26	1	28
Social2	2	21	26	49
	416	80	53	549

Attribute Dimension

The attribute dimension has two categories: attribute present and attribute absent. Cohen's Kappa for this dimension is 0.86.

	Attribute present	Attribute absent	
Analogical1	50	41	91
Analogical2	101	13	114
Decision1	50	11	61
Decision2	41	25	66
Knowledge1	67	11	78
Knowledge2	56	6	62
Social1	20	8	28
Social2	47	2	49
	432	117	549

Repetition Dimension

The repetition dimension has two categories: single and repeated. Cohen's Kappa for this dimension is 0.67.

	Single	Repeated	
Analogical1	89	2	91
Analogical2	78	36	114
Decision1	42	19	61
Decision2	32	34	66
Knowledge1	78	0	78
Knowledge2	61	1	62
Social1	28	0	28
Social2	48	1	49
	456	93	549