Table 2. Description of data labelling structure for BERT fine-tuning, and inter-coder reliabilities

Underlying concept captured	Classifier reference	Key concept captured as explained by key labelling instruction for fine-tuning (answer options in parentheses)	Inter-coder reliabilities (Krippendorff's alpha)	
			Round 1	Round 2
Interrogative	1A	Does this interrogative request an answer? (yes/no)	0.66	0.71
	1B	Is this a declarative/imperative interrogative? (NA/no/yes)	0.95	NA
Selection-size-specification	2A	Is this an interrogative that expects a yes or no answer? (yes/no)	0.87	NA
	2B	Does it explicitly present a series of options? (yes/no)	0.67	0.69
	2C	How many options does it present? (0/1/2/any other integer/ undefined)	0.74	NA
Presupposition	3A	Do answers to this interrogative require some other fact/opinion already being true? (yes/no)	0.78	NA
Description	4A	Does this interrogative ask for a description (yes/no) or an opinion (opinion)?	0.60	0.70

Note. Inter-coder reliabilities were assessed using Krippendorff's alpha, because it accommodates multiple annotators, handles missing data, and supports more than two labelling categories. Alpha values of ≥ 0.80 were interpreted as strong agreement, while values between 0.60 and 0.79 indicated moderate agreement. Round 1 reports Krippendorff's alpha for three annotators, and Round 2 for two annotators. Solely those with lower inter-coder reliabilities were carried forward to the second labelling round.