Real Error Correlation Scores															70					
ALIGNScore	100	57	66	61	43	18	48	27	38	77	3	22	8	9	63	85	45			<b>-</b> 70
DSG-BLIP1	57	100	72	11	45	-11	69	59	85	77	29	54	-4	-12	77	71	20			
TIFA-BLIP1	66	72	100	26	44	34	71	69	64	83	25	69	14	-0	72	72	5		-	- 60
CLIPScore	61	11	26	100	7	32	23	2	-13	29	-14	-1	-1	-1	9	35	35			
DSG-Fuyu	43	45	44	7	100	18	28	39	45	59	13	46	-8	1	47	39	10		_	- 50
TIFA-Fuyu	18	-11	34	32	18	100	7	23	-9	25	-4	36	-18	3	15	-1	-13			00
DSG-InstructBLIP	48	69	71	23	28	7	100	67	66	65	30	62	25	-7	60	64	-6			
TIFA-InstructBLIP	27	59	69	2	39	23	67	100	73	68	45	82	4	13	48	42	-20		-	- 40
DSG-LLaVA (alt)	38	85	64	-13	45	-9	66	73	100	71	64	65	16	9	78	56	10			
TIFA-LLaVA (alt)	77	77	83	29	59	25	65	68	71	100	21	64	-1	-9	84	81	23		_	- 30
DSG-LLaVA	3	29	25	-14	13	-4	30	45	64	21	100	43	48	47	40	9	12			
TIFA-LLaVA	22	54	69	-1	46	36	62	82	65	64	43	100	1	3	55	37	-10			
LLMScore EC	8	-4	14	-1	-8	-18	25	4	16	-1	48	1	100	57	13	22	16		_	- 20
LLMScore Over	9	-12	-0	-1	1	3	-7	13	9	-9	47	3	57	100	3	-5	29			
DSG-MPlug	63	77	72	9	47	15	60	48	78	84	40	55	13	3	100	72	26		-	- 10
TIFA-MPlug	85	71	72	35	39	-1	64	42	56	81	9	37	22	-5	72	100	31			
VIEScore	45	20	5	35	10	-13	-6	-20	10	23	12	-10	16	29	26	31	100			0
	ALIGNScore	DSG-BLIP1	TIFA-BLIP1	CLIPScore	DSG-Fuyu	TIFA-Fuyu	DSG-InstructBLIP	TIFA-InstructBLIP	DSG-LLaVA (alt)	TIFA-LLaVA (alt)	DSG-LLaVA	TIFA-LLaVA	LLMScore EC	LLMScore Over	DSG-MPlug	TIFA-MPlug	VIEScore	•		- 0