

Table R-3: Experimental results of testing different emotion texts on models triggered by different emotion attacks

	\mathbf{X}_{sad}	$\mathbf{X}_{\text{angry}}$	$\mathbf{X}_{\text{isolated}}$
EmoBooth _{sad}	0.8456	0.6890	0.7213
EmoBooth _{angry}	0.6850	0.8028	0.6832
EmoBooth _{isolated}	0.7331	0.6922	0.8320