Neuron Activation Comparison: Buggy vs Non-Buggy Responses
Llama-3.1-8B-Instruct on "Which is bigger: 9.8 or 9.11?"

Layer	Buggy Response (Chat Template - "9.11 is bigger")	Non-Buggy Response (Simple Format - "9.8 is bigger")
7	Neuron 1978	-
11	-	Neuron 11862
13	Neuron 10352	Neuron 10352
14	Neuron 13315 (2x) Neuron 2451 Neuron 12639 ⚠	Neuron 13315 (2x) Neuron 12639 ⚠
15	Neuron 3136 (2x) Neuron 5076 Neuron 421	Neuron 3136 (2x) Neuron 5076 (2x) Neuron 421
28	Neuron 10823 Neuron 8818 Neuron 5336	Neuron 11450 Neuron 12900 Neuron 10823
29	Neuron 664 Neuron 12248 Neuron 1435	Neuron 12248 Neuron 10726 Neuron 2836
30	Neuron 840 Neuron 801 Neuron 14215 (multiple) Neuron 4581 Neuron 7303	Neuron 840 Neuron 801 Neuron 14215 (multiple) Neuron 4581
31	Neuron 13336 (12.0) Neuron 12004 (11.5) Neuron 9692 Neuron 2398 Neuron 12111	Neuron 13336 (14.8) ↑ Neuron 12004 (14.4) ↑ Neuron 9692 Neuron 2398 Neuron 9692

⚠ = Entangled neuron (fires in both buggy and correct responses)
↑ = Higher activation in correct response
Red borders = Layers with shared neurons between buggy/correct responses
Layers 2-15: Hijacker Circuit | Layers 28-31: Reasoning Circuit