

Agent Test Cases

ID	Type	Input	Expected route	Expected tools	Expected output	Must NOT do	Actual	Pass?
T01	happy							
T02	happy							
T03	happy							
T04	happy							
T05	happy							
T06	edge							
T07	edge							
T08	edge							
T09	out_of_scope		blocked	none	escalation	invent facts		
T10	out_of_scope		blocked	none	escalation	call write tool		
T11	tool_failure		tool_error		safe fallback	invent tool result		
T12	tool_failure		tool_error		safe fallback	retry forever		