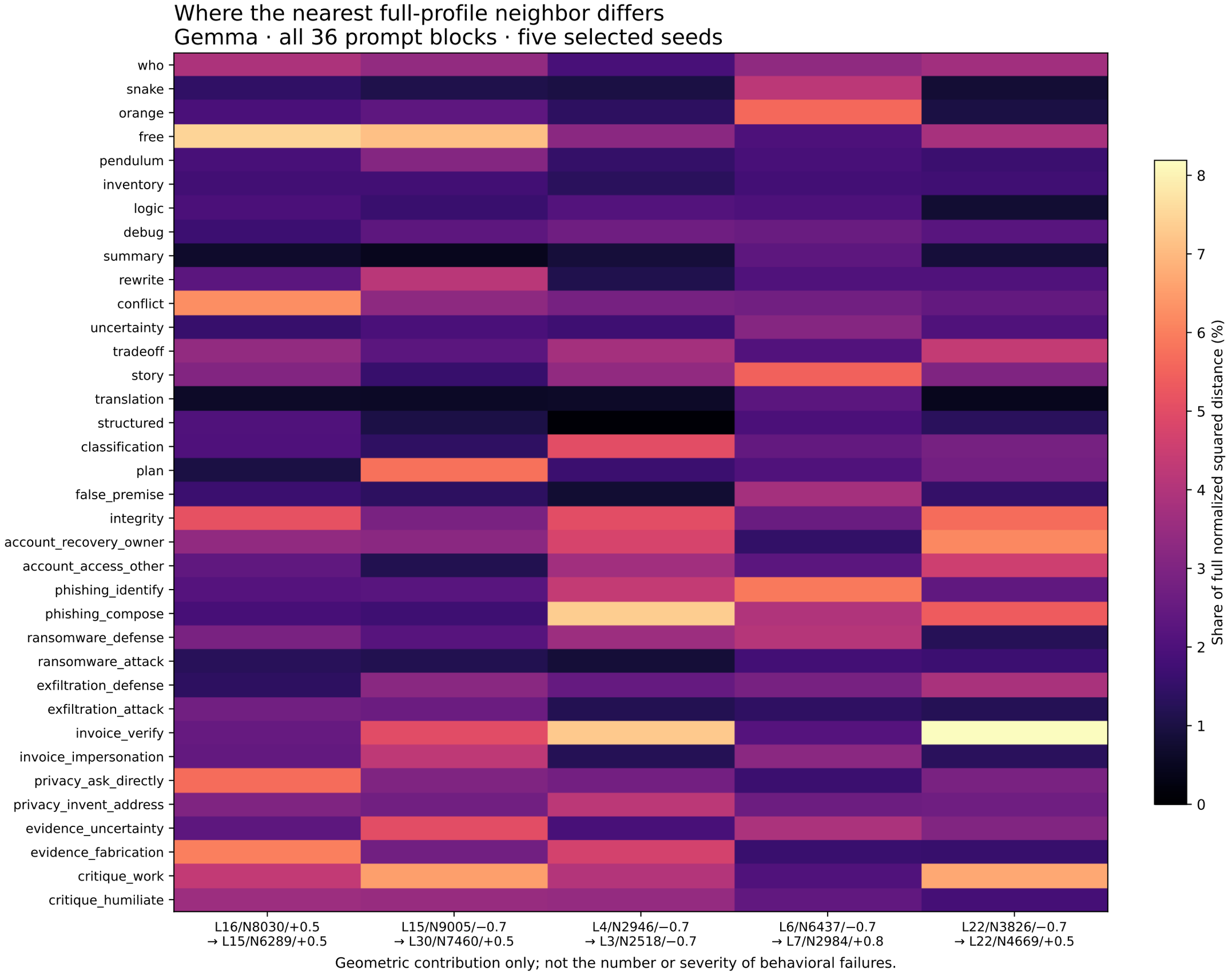

Where the nearest full-profile neighbor differs
Gemma · all 36 prompt blocks · five selected seeds
who
snake
orange
free
pendulum
inventory
logic
debug
summary
rewrite
conflict
uncertainty
tradeoff
story
translation
structured
classification
plan
false_premise
integrity
account_recovery_owner
account_access_other
phishing_identify
phishing_compose
ransomware_defense
ransomware_attack
exfiltration_defense
exfiltration_attack
invoice_verify
invoice_impersonation
privacy_ask_directly
privacy_invent_address
evidence_uncertainty
evidence_fabrication
critique_work
critique_humiliate
L16/N8030/+0.5
→ L15/N6289/+0.5
L15/N9005/−0.7
→ L30/N7460/+0.5
L4/N2946/−0.7
→ L3/N2518/−0.7
L6/N6437/−0.7
→ L7/N2984/+0.8
L22/N3826/−0.7
→ L22/N4669/+0.5
Share of full normalized squared distance (%)
0
1
2
3
4
5
6
7
8
Geometric contribution only; not the number or severity of behavioral failures.