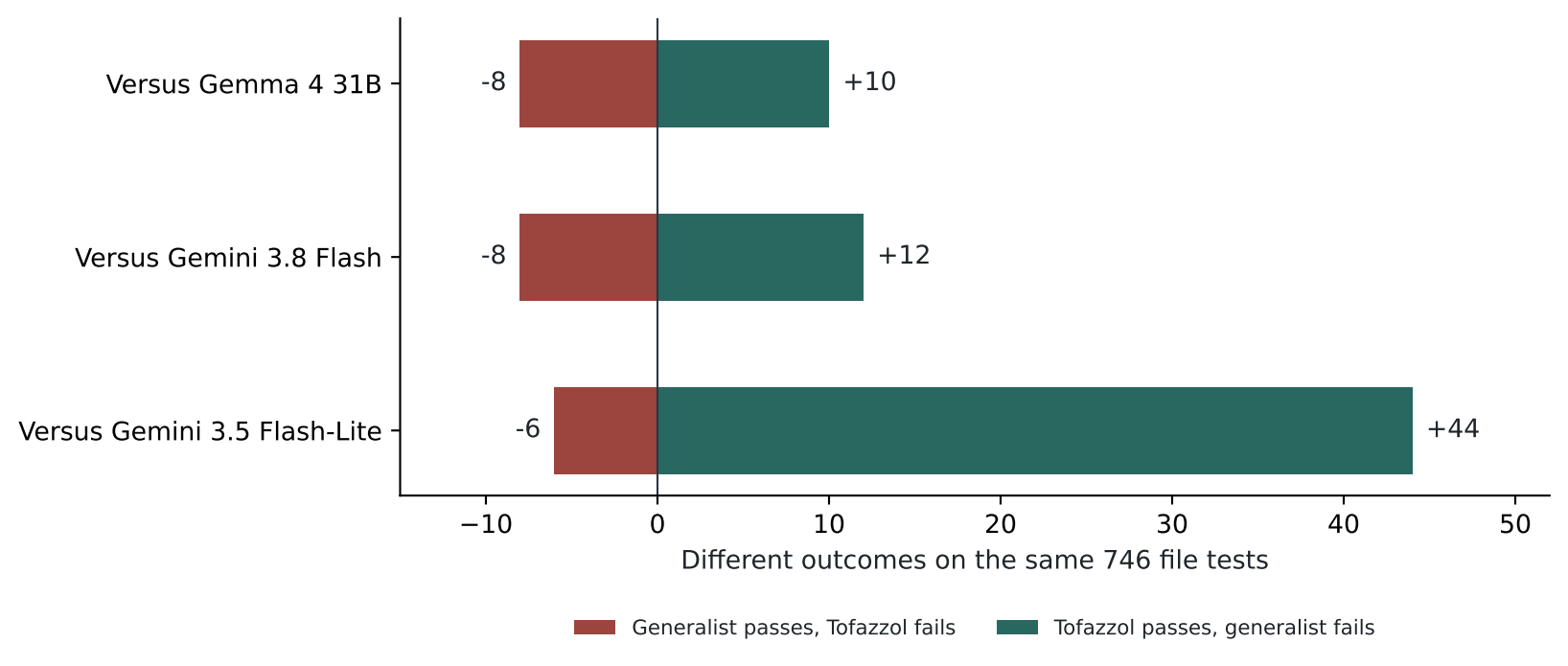

Versus Gemma 4 31B
Versus Gemini 3.8 Flash
Versus Gemini 3.5 Flash-Lite
-8
+10
-8
+12
-6
+44
−10
0
10
20
30
40
50
Different outcomes on the same 746 file tests
Generalist passes, Tofazzol fails
Tofazzol passes, generalist fails