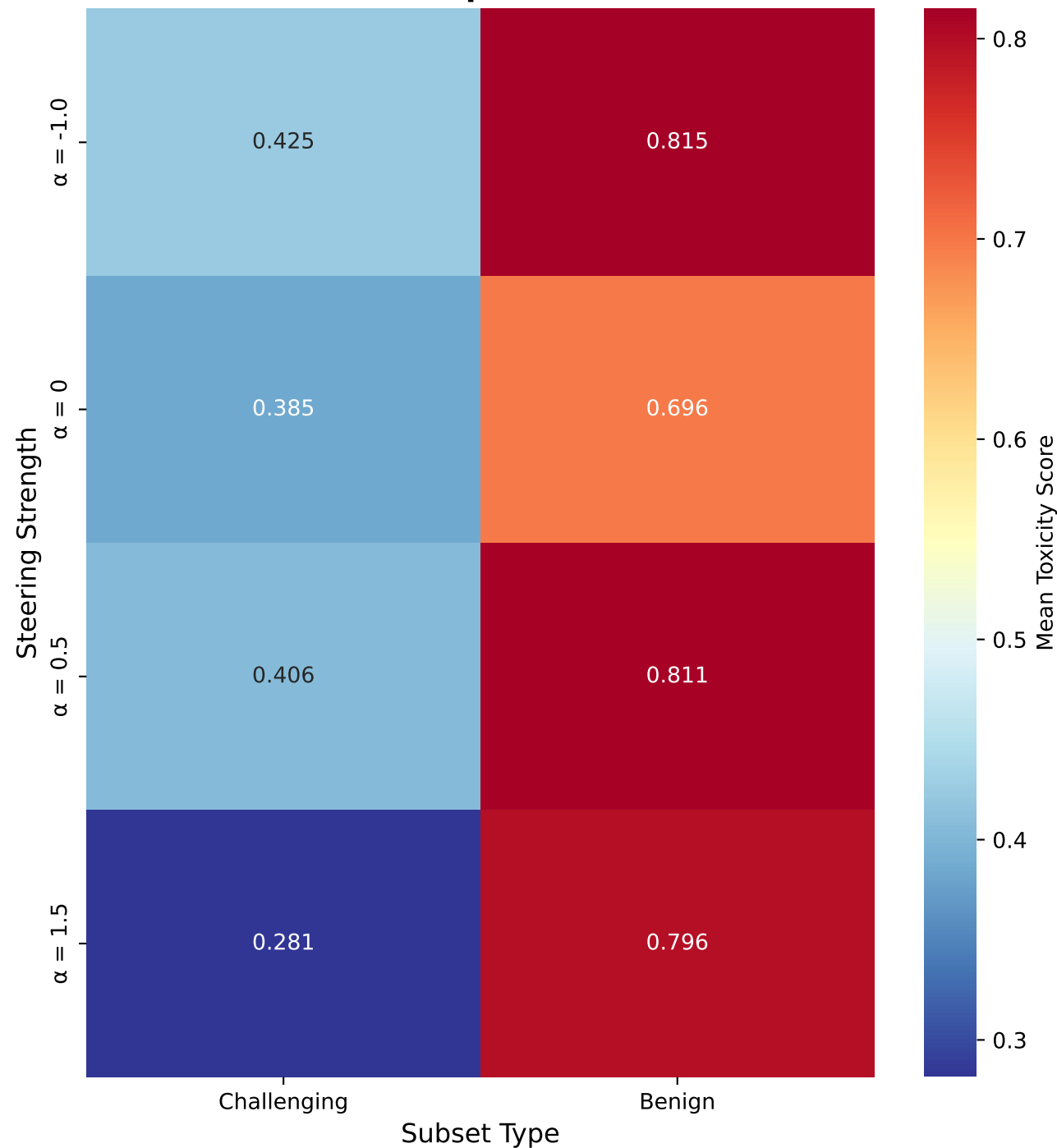


**Toxicity Scores Heatmap  
Across Alpha Values**



**Toxicity Improvement from Baseline  
(% Change)**

