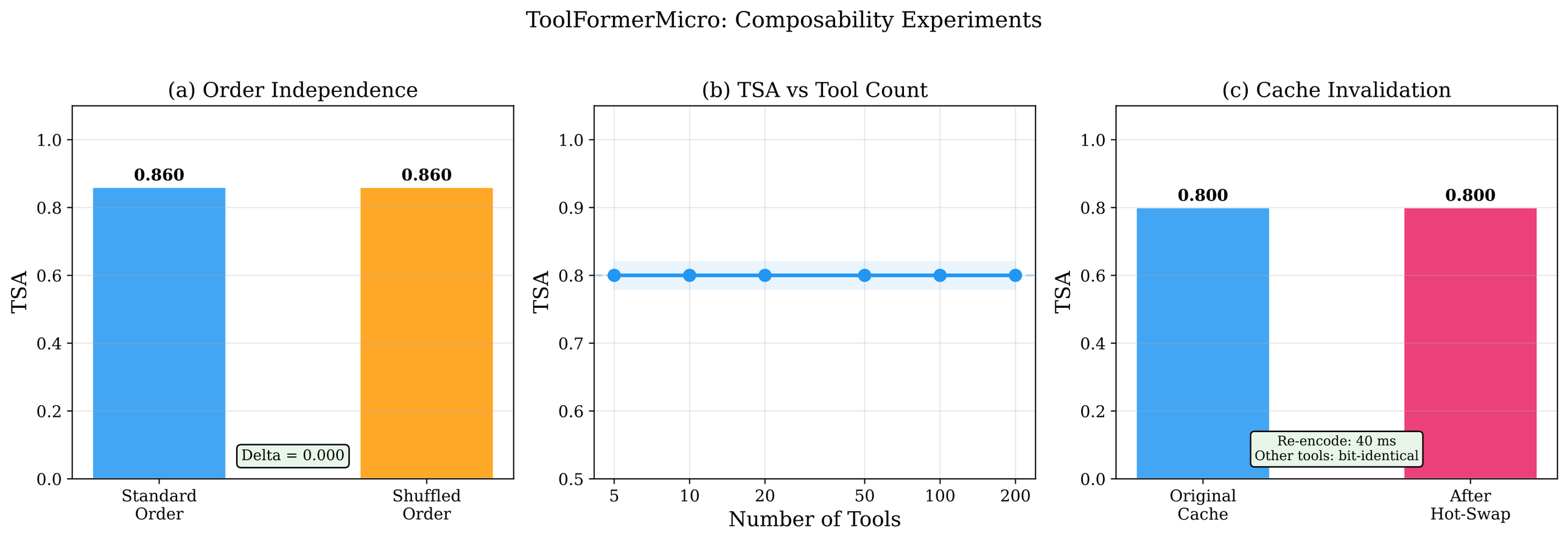
ToolFormerMicro: Composability Experiments
(a) Order Independence
TSA
1.0
0.8
0.6
0.4
0.2
0.0
0.860
0.860
Delta = 0.000
Standard Order
Shuffled Order
(b) TSA vs Tool Count
TSA
1.0
0.9
0.8
0.7
0.6
0.5
5
10
20
50
100
200
Number of Tools
(c) Cache Invalidation
TSA
1.0
0.8
0.6
0.4
0.2
0.0
0.800
0.800
Re-encode: 40 ms
Other tools: bit-identical
Original Cache
After Hot-Swap