ζ-R-max agent policy - last 500 steps -0.7 -3 -13 -13 -13 -13 -13 -13 -13 -0.6 -2.8 -12 -13 -13 -13 -13 -13 -0.7 -2.4 -9.1 -13 -13 -12 -13 -0.6 -1.7 -12 -13 -13 -2.2 -4.8 -12 -0.7 -0.7 -2 -3.8 -0.7 -0.8 -0.9 -1.4 -0.6 -0.8 -1.5 -1.7 -1.6 -0.8 -0.9 -0.6 -1.1 -1 -0.9 -0.9 -1.2 -12

2.5

3.0

3.5

1.5

2.0

alpha

0.5

1.5

3.0

3.5

0.1

0.5

1.0

m 2.5