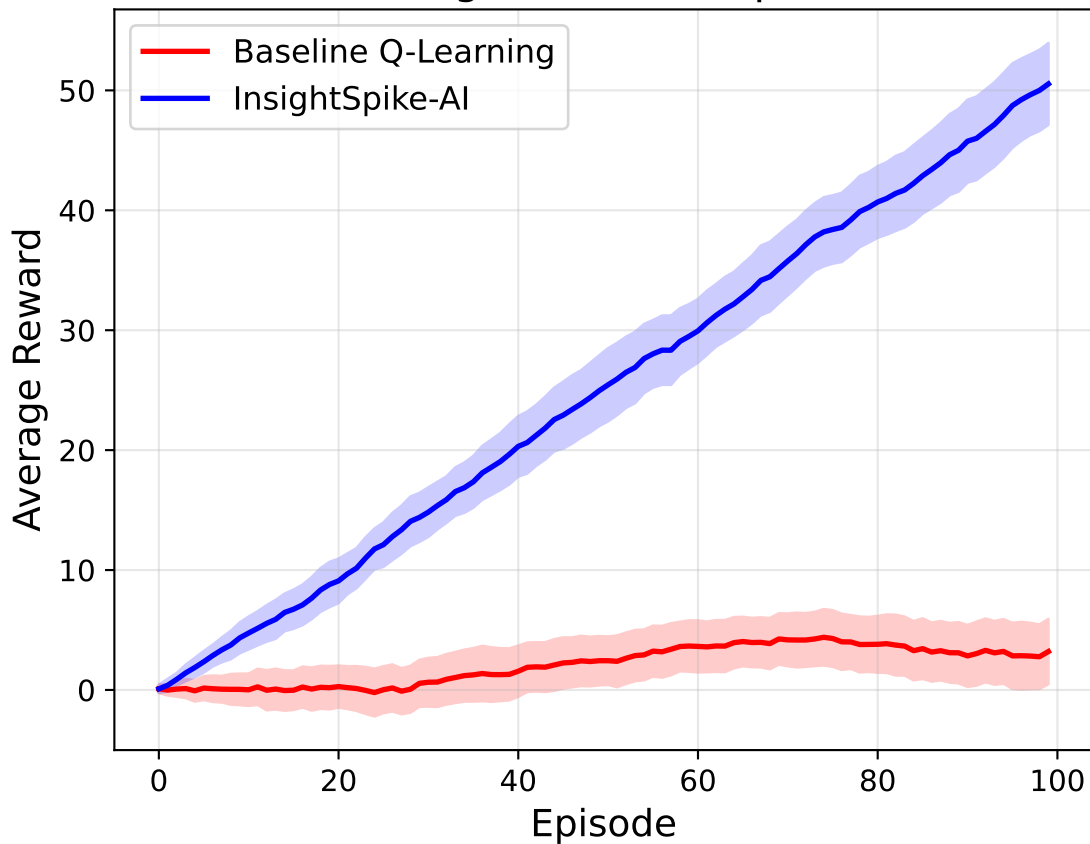
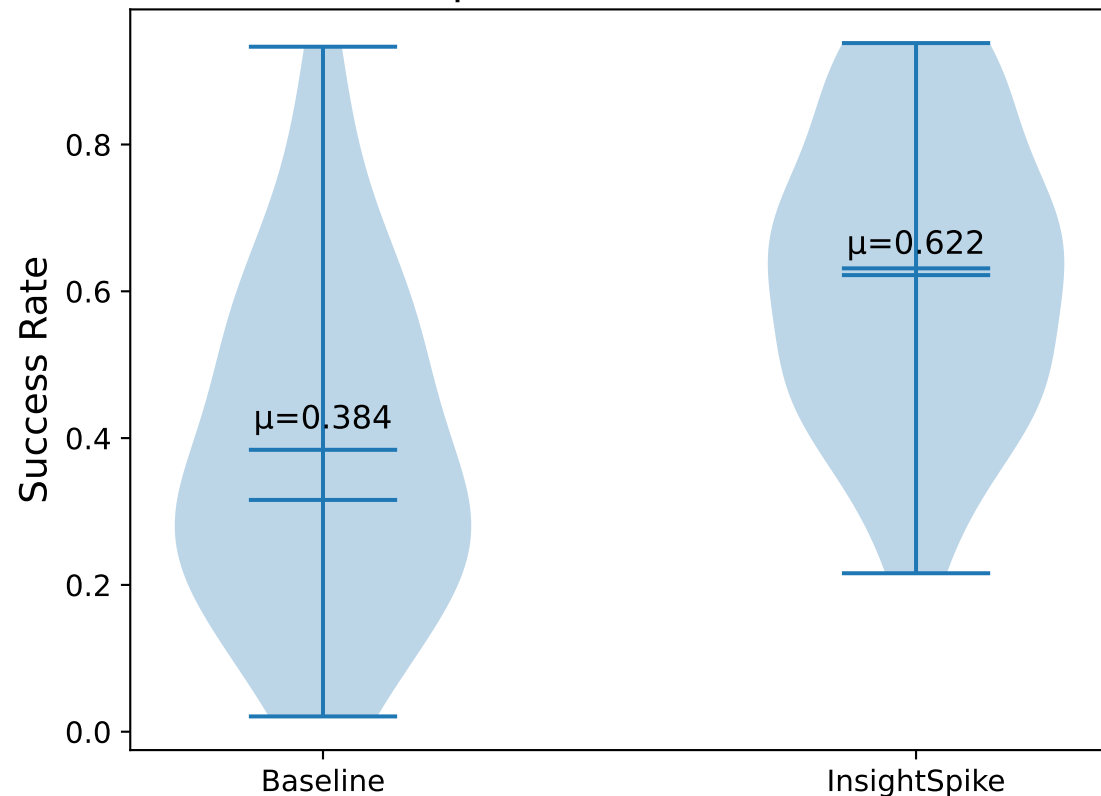


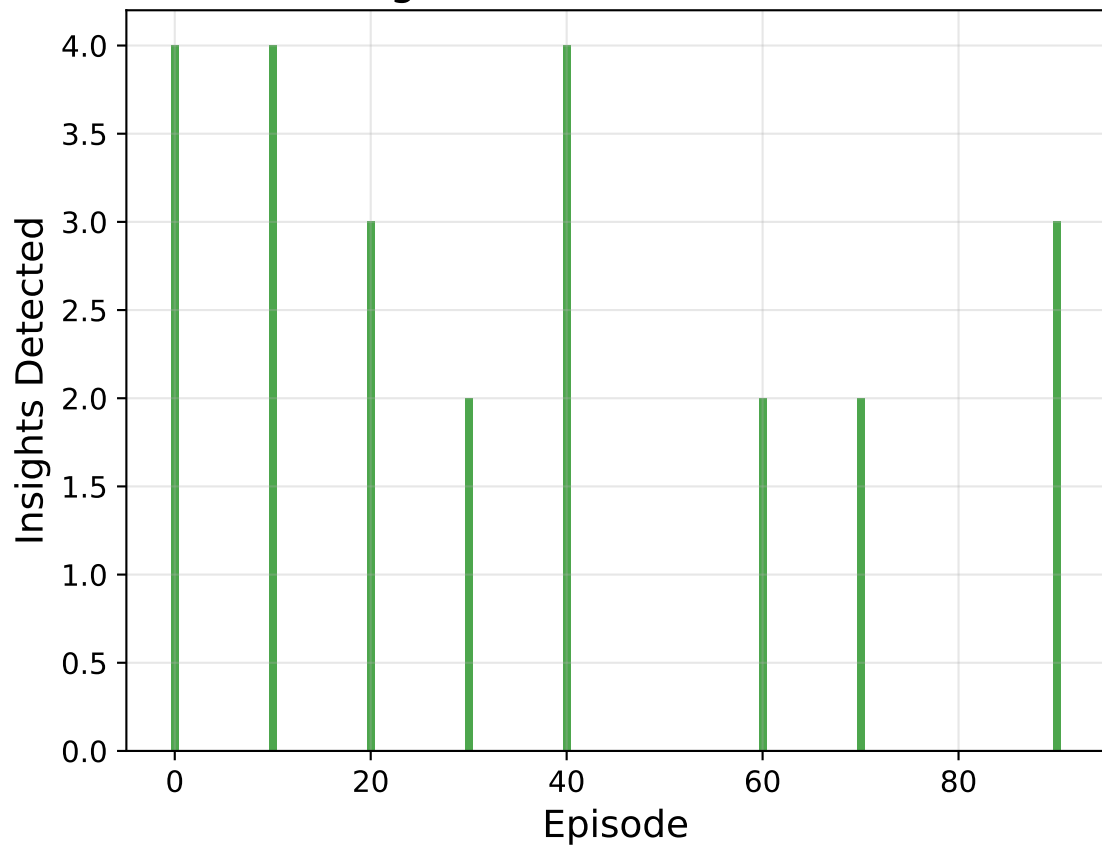
Learning Curves Comparison



Success Rate Distribution (p-value: 0.0001)



Insight Detection Over Time



Statistical Significance Tests

