

| Q-Learning vs Param Policy - Detailed Comparison Table | | | | |
|--|------------------------|-----------------------|-------------|--------------|
| Metric | Q-Learning | Param Policy | Improvement | Winner |
| OVERALL PERFORMANCE | | | | |
| Average Reward | 264.26 (± 27.33) | 59.95 (± 12.38) | +340.8% | Q-Learning |
| Average Midterm Score (/10) | 7.14 (± 0.82) | 5.82 (± 0.48) | +22.5% | Q-Learning |
| Average LO Mastery | 0.707 (± 0.085) | 0.621 (± 0.056) | +13.9% | Q-Learning |
| Average Weak LO Count | 1.48 | 3.02 | -51.0% | Param Policy |
| CLUSTER medium | | | | |
| Reward (n=55) | 247.76 | 64.57 | +283.7% | Q-Learning |
| Midterm Score (/10) | 7.03 | 5.82 | +20.7% | Q-Learning |
| LO Mastery | 0.695 | 0.621 | +12.0% | Q-Learning |
| CLUSTER strong | | | | |
| Reward (n=28) | 264.68 | 68.42 | +286.8% | Q-Learning |
| Midterm Score (/10) | 8.18 | 6.53 | +25.3% | Q-Learning |
| LO Mastery | 0.817 | 0.704 | +16.1% | Q-Learning |
| CLUSTER weak | | | | |
| Reward (n=17) | 316.93 | 34.62 | +815.4% | Q-Learning |
| Midterm Score (/10) | 5.76 | 5.05 | +14.1% | Q-Learning |
| LO Mastery | 0.567 | 0.530 | +7.0% | Q-Learning |
| STATISTICAL TEST | | | | |
| T-statistic | 67.744 | | | |
| P-value | 0.000000 | | Significant | |