

Answer sheet: Week 6 self-study, Same rewards, different learning rates

If your numbers differ, find the first trial where they differ.

Completed tables

Update rule: $V \leftarrow V + \alpha \cdot (r - V)$, $V_0 = 5$. "PE" = prediction error $r - V$ before the update.

World A: fixed mean, noisy payoffs (true mean 5)

Trial	r	V fast (0.8)	PE fast	V slow (0.1)	PE slow
1	6	5.80	+1.00	5.10	+1.00
2	4	4.36	-1.80	4.99	-1.10
3	7	6.47	+2.64	5.19	+2.01
4	3	3.69	-3.47	4.97	-2.19
5	5	4.74	+1.31	4.97	+0.03
6	6	5.75	+1.26	5.08	+1.03

World B: the mean changes (true mean 5, then 10 from trial 4)

Trial	r	V fast (0.8)	PE fast	V slow (0.1)	PE slow
1-3	5	5.00	0	5.00	0
4	10	9.00	+5.00	5.50	+5.00
5	10	9.80	+1.00	5.95	+4.50
6	10	9.96	+0.20	6.36	+4.05

(The slow learner at trial 6 in World B is 6.355 before rounding; small differences in the last digit can come from rounding at each step.)

Questions

1. Distance from the true value after trial 6, and which α is better.

- World A: fast $|5.75 - 5| = \mathbf{0.75}$; slow $|5.08 - 5| = \mathbf{0.08}$. The slow learner is better: it averages over many noisy payoffs, whereas the fast learner is pulled around by each one (its estimate swings between 3.69 and 6.47).
- World B: fast $|9.96 - 10| = \mathbf{0.04}$; slow $|6.36 - 10| = \mathbf{3.64}$. The fast learner is better: it discards the outdated past quickly, whereas the slow learner is still anchored to the old value.
- Message: no α is best everywhere. The best α depends on how much of the change in payoffs is real change and how much is noise.

2. Pattern of prediction errors.

- World A: errors mostly alternate in sign and are roughly symmetric around zero, so they roughly cancel: noise.
- World B, slow learner: errors are large and keep the same sign after trial 4 (+5.00, +4.50, +4.05). Persistent one-signed errors suggest that the estimate is systematically wrong: a change.
- World B, fast learner: the error is large once (+5.00) and then shrinks quickly because the learner has already caught up. One large error alone looks the same in both worlds (compare trial 4 in World B with trial 4 in World A for the fast learner: -3.47). What distinguishes the worlds is what comes next.
- So yes, a learner could use the pattern: persistent same-sign errors suggest change and call for a higher α ; alternating errors suggest noise and call for a lower α .

3. What else would an ideal learner estimate? How volatile the world is (how often and how much the true mean changes), and how noisy single payoffs are. Week 8 returns to this question (Behrens et al., 2007). Acceptable additional answer: the probability that a change has just occurred.

4. Assignment 2 fund (optional). The fund pays +15 (60%) or -10 (40%) each year, independently, with an unchanged expected value. That is World A (fixed mean, noisy payoffs), not World B. In finance terms, its returns are volatile. The large swings tempt a fast learner to chase noise. A good answer also notes that in a real market the underlying probabilities may themselves change, so the world can switch from A-like to B-like without warning.

Comment section: check questions and expected answers

- *What would you need to claim that the ACC sets the learning rate rather than reflects it?* An intervention: for example a lesion or stimulation of the region that changes the learning rate in the predicted direction (Week 3: only interventions license causal claims).

- *What does a group difference tell you about one person?* Little. Group averages differ, but individuals overlap heavily; the height analogy applies.
- *Why not read the anxiety result as a diagnosis?* It is a group-level difference in one aversive learning task; trait anxiety here is a normal-range personality measure.

Further reading

The comment in the exercise summarises three papers; none is required. If you read only one, read Behrens et al. (2007), which we return to in Week 8.