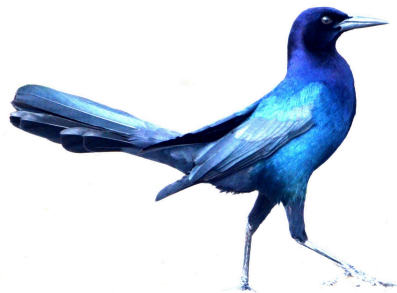
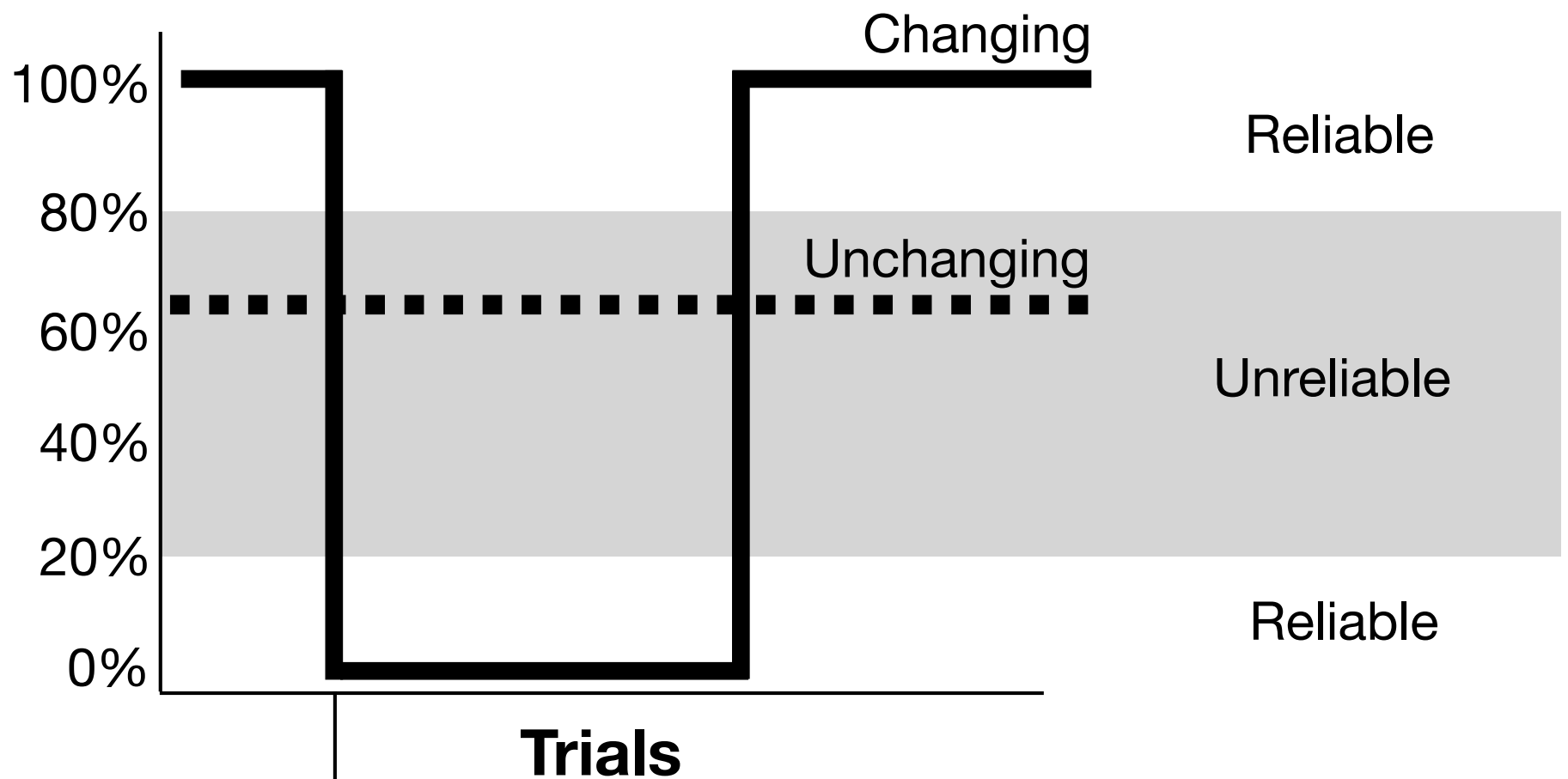


Probability option will be rewarded this trial



X

No reward

What to do in the trials after X?

- Reliable: switch away (large  $\phi$ )
- Changing: sometimes explore (small  $\lambda$ )
- Unreliable: select again (small  $\phi$ )
- Unchanging: always exploit (large  $\lambda$ )