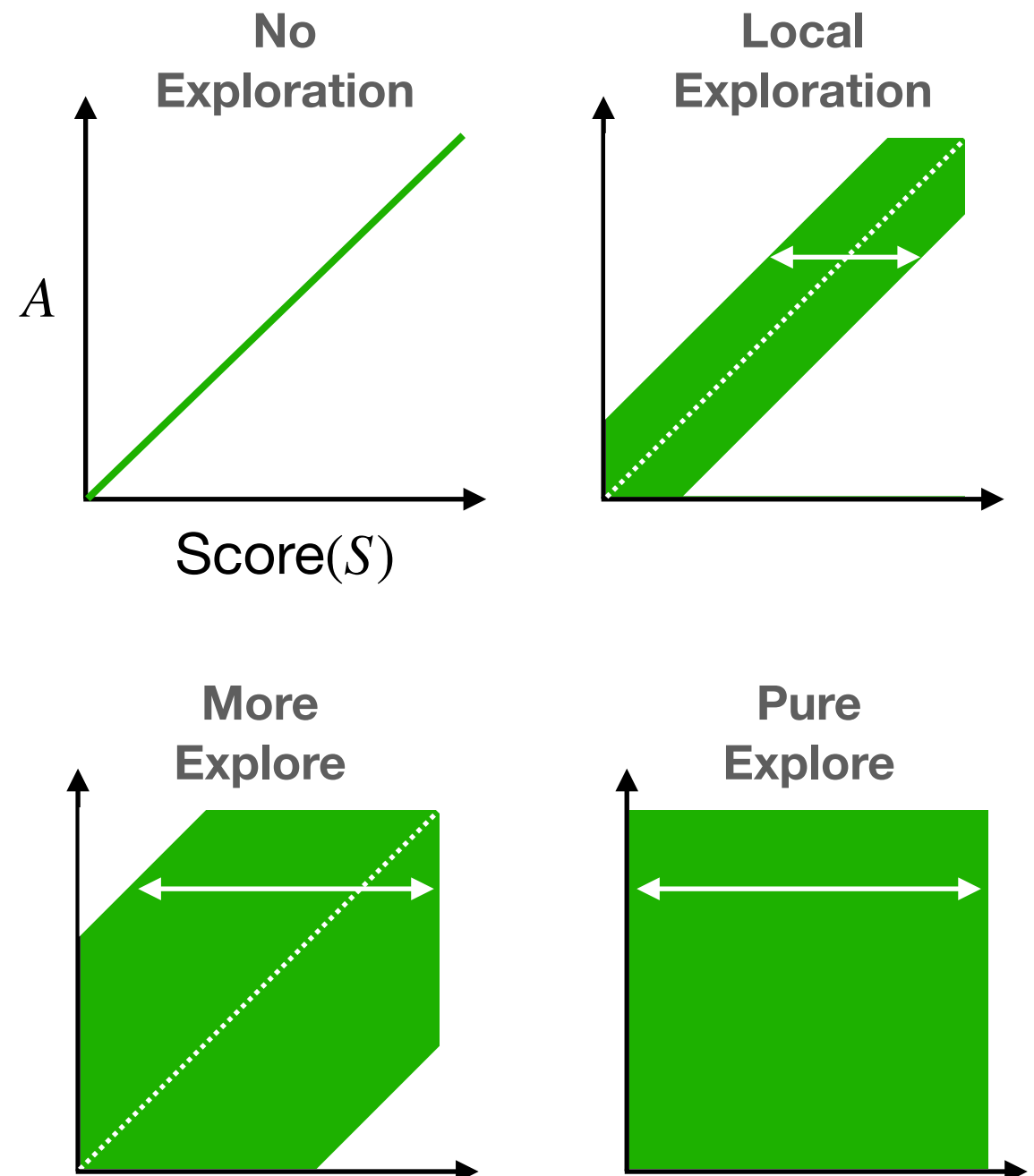
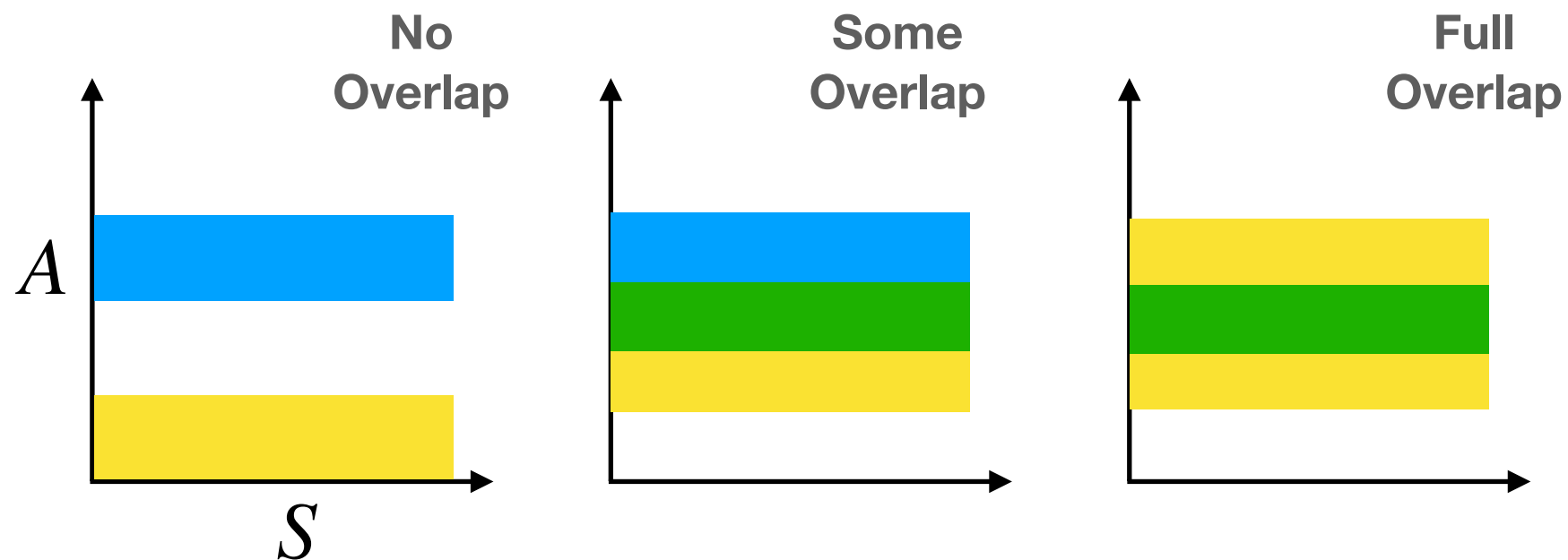


- Typically we only put highest scoring items in top positions.
- We can never know if lower scoring items might be useful for users in higher positions without exploration.
- Exploration bears costs: we have to sometimes take actions that are sub-optimal.
- Decouples A from S , allowing us to estimate effect of setting A : $E[Y | S, \text{do}(A)]$



■ $P(A | S) > 0$

Overlap in SpA ce



Data
Collection
Policy



Target
Policy

Both