Optimal trade-off between exploration and exploitation in human movement

Human beings are excellent at achieving arbitrary control objectives in noisy, uncertain and changing environments. Acting in such environments, whether the system is biological or artificial, often involves a trade-off between exploratory actions, whose goal is to gather sensory information and 'regular' actions which exploit the information gathered so far and pursue the task objectives. Sensory information can include target position, end effector position, joint position, joint velocity, force, and more. How can both types of action be generated by minimizing a single cost function within the framework of stochastic optimal control? Here we formalize the problem in a way that captures the essence of the exploration-exploitation tradeoff while yielding a nonlinear control problem which can be approached with appropriate numerical methods. The key to the approach is augmenting the partially observable plant dynamics with the estimator dynamics, obtaining a higher dimensional but fully observable control problem. The control behavior is compared with human subjects performing the same task.