While many recent studies have successfully used reinforcement learning (RL) frameworks to explain large portions of variance within neurobiological and decision-making datasets, the relatability of such models to the true mechanisms and dynamics underlying human learning, cognition, and behavior is arguably still quite limited—in part due to the exclusion of well-defined mechanisms controlling the dynamics of sensory-model updating (particularly during exploratory behavior) and sensory-model extraction (for use of exploitative behavior) processes. In an attempt to mend this gap, the current study investigates the diameter of the pupil as a potential signature of both ongoing sensory-model updating and sensory-model extraction processes. With the use of a hybrid Q-learning model, these hypothesized correlates are found to account for discrepancies in pupil diameter between model-based and model-free learning strategies during exploratory and exploitative behavior, and simultaneously frame human learning experience as a dynamic interplay between sensory-model updating and recollection processes.