Todd, P. M. (1997). Searching for the next best mate. In *Simulating social phenomena* (pp. 419-436). Springer, Berlin, Heidelberg.

Summary article about optimal stopping and the 37% rule. Resource preservation and speed are the cognitive processes that humans prioritize when solving these kinds of problems, and this paper explains how humans use simplified algorithms and heuristics to solve the problem more easily.
He studies three cases with their three simplified algorithms: shopping for tomatoes, applying for colleges, looking for a partner.
The one bounce rule is searching until a smaller number appears and picking the previous number, and is close to optimum. The gale shapely rule explains college ranking and then pairing up each spot with a student as a solution to this type of problem.
Time left is also used to make the decision to explore or exploit.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933-942.

This paper talks some more about opportunity cost as a factor humans consider when exploring, as well as how well they know the environment, and how likely the environment is to change. This is what we call the explore exploit dilemna in class. One solution is the gittins index, like in picking the most reward giving bandit in the bandit problem.

Juni, M., Gureckis, T., & Maloney, L. (2011). Don't stop 'til you get enough: adaptive information sampling in a visuomotor estimation task. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 33, No. 33).
This study looks at the explore exploit dilemma when the environment changes and how people perform across different environments. He gives the example of shopping for an item and how humans decide when to keep looking for a bargain or not based on the expected costs. Since these expected costs are hard to predict, people under sample. But when they did an experiment that gave people the opportunity cost of future actions and helped them keep past actions in memory, and a clearly sequential problem, people oversampled.
In another experiment they gave people a touch-target task (similar to darts) with an explicit cost for sampling.
Keeping the rewards statis didn't help with oversampling.

Juni, M. Z., Gureckis, T. M., & Maloney, L. T. (2016). Information sampling behavior with explicit sampling costs. *Decision*, *3*(3), 147.

Same touch-target experiment as above, but with penalties on final reward for one group. Similar results.