University of Groningen

Unpacking the Exploration–Exploitation Tradeoff

Mehlhorn, Sabine; Newell, Ben R; Todd, Peter M; Lee, Michael D; Morgan, Kate; Braithwaite, Victoria A; Hausmann, Daniel; Fiedler, Klaus; Gonzalez, Cleotilde

Published in:
Decision

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Mehlhorn, S., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., & Gonzalez, C. (2015). Unpacking the Exploration–Exploitation Tradeoff: A Synthesis of Human and Animal Literatures. *Decision*, *2*(3), 191–215.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 05-06-2022

Decision
2015, Vol. 2, No. 3, 191–215

2325-9965/15/$12.00 http://dx.doi.org/10.1037/dec0000033

Unpacking the Exploration–Exploitation Tradeoff: A Synthesis of Human and Animal Literatures

Katja Mehlhorn
Carnegie Mellon University and
University of Groningen

Ben R. Newell
University of New South Wales

Peter M. Todd
Indiana University

Michael D. Lee
University of California

Kate Morgan
University of St. Andrews

Victoria A. Braithwaite
The Pennsylvania State University

Daniel Hausmann
University of Zürich

Klaus Fiedler
University of Heidelberg

Cleotilde Gonzalez
Carnegie Mellon University

Many decisions in the lives of animals and humans require a fine balance between the exploration of different options and the exploitation of their rewards. Do you buy the advertised car, or do you test drive different models? Do you continue feeding from the current patch of flowers, or do you fly off to another one? Do you marry your current partner, or try your luck with someone else? The balance required in these situations is commonly referred to as the exploration–exploitation tradeoff. It features prominently in a wide range of research traditions, including learning, foraging, and decision making literatures. Here, we integrate findings from these and other often-isolated literatures in order to gain a better understanding of the possible tradeoffs between exploration and exploitation, and we propose new theoretical insights that might guide future research. Specifically, we explore how potential tradeoffs depend on (a) the conceptualization of exploration and exploitation; (b) the influencing environmental, social, and individual factors; (c) the scale at which exploration and exploitation are considered; (d) the relationship and types of transitions between the 2 behaviors; and (e) the goals of the decision

---

Katja Mehlhorn, Department of Social and Decision Sciences, Carnegie Mellon University, and Department of Artificial Intelligence, University of Groningen; Ben R. Newell, School of Psychology, University of New South Wales; Peter M. Todd, Cognitive Science Program and Department of Psychological and Brain Sciences, Indiana University; Michael D. Lee, Department of Cognitive Sciences, University of California; Kate Morgan, School of Biology, University of St. Andrews; Victoria A. Braithwaite, Center for Brain, Behavior and Cognition, Department of Ecosystem Science and Management, The Pennsylvania State University; Daniel Hausmann, Department of Psychology, University of Zürich; Klaus Fiedler, Institute of Psychology, University of Heidelberg; Cleotilde Gonzalez, Department of Social and Decision Sciences, Carnegie Mellon University.

Katja Mehlhorn and Cleotilde Gonzalez were supported by the National Science Foundation Award Number 1154012 to Cleotilde Gonzalez. The ideas in this article originated from discussions in a workshop entitled Predicting Choice from Exploration, organized by Cleotilde Gonzalez and Katja Mehlhorn at the 9th Invitational Choice Symposium, Huis ter Duin, Noordwijk, The Netherlands, June 12–16, 2013.

Correspondence concerning this article should be addressed to Katja Mehlhorn, University of Groningen, Deptartment of Artificial Intelligence, Nijenborgh 9, 9747 AG Groningen, The Netherlands. E-mail: s.k.mehlhorn@rug.nl

maker. We conclude that exploration and exploitation are best conceptualized as points on a continuum, and that the extent to which an agent's behavior can be interpreted as exploratory or exploitative depends upon the level of abstraction at which it is considered.

Keywords: exploration–exploitation tradeoff, learning, foraging, decision making, decision theory

Consider the following scenarios. (a) You work for the Widget Corporation and you are paid according to how many functional widgets you can produce. You have access to two widget machines but can only use one at a time. On the first day, you know nothing about the machines so you pick one at random and start work. After 10 functional widgets, the machine produces a faulty one. What do you do? Do you tolerate the single faulty widget and persevere, or do you try your luck on the other machine? (b) You are a hummingbird feeding in a field of flowers. You pick one patch of flowers and begin to drink the nectar. How long should you remain at that patch before seeking another? Would you leave when all the flowers have been exhausted? What if the nectar in nearby patches has already been harvested? (c) You arrive in a city for a few days and have a range of restaurants to choose from. Do you try as many different restaurants as you can, or do you look for restaurants of a specific type? Toward the end of your visit, do you stop searching for new restaurants and revisit the ones that you enjoyed most? When do you make this switch in your strategy? (d) You are a college student on the dating market. Your goal might be to find a partner for life, or you might be more interested in dating as many people as you can. How would you approach those goals? Could you combine them into a perfect search strategy? What if the partner of your choice is not interested in you? (e) You finally decide to buy a new car. Do you search the Internet for information about different car companies, or do you trust your own experience and stick with your current company? Once you have chosen a dealership, how many cars do you look at? How long do you test drive a specific car before you decide to buy it?

Many approaches to the analysis of decision behavior would characterize these scenarios as representations of a tradeoff between *exploration* and *exploitation* (e.g., in reinforcement learning [RL] and neuroscience: Cohen, McClure, & Yu, 2007; in foraging: Cook, Franks, & Robinson, 2013; in binary risky choice: Gonzalez & Dutt, 2011; in organizational learning: Gupta, Smith, & Shalley, 2006; for a review, see Hills, Todd, Lazer, Redish, & Couzin, 2015). Remaining at an option—be it a machine, patch of flowers, restaurant, partner, or car—allows for exploitation, that is, making the most of where you are. A switch to another option, going somewhere else to see if you can get a better reward—fault-free widgets, more nectar, better food, a higher reproductive value, or a faster car—exemplifies exploration. Although these concepts seem quite simple on the surface (e.g., staying is exploitation; switching is exploration), the definitions, processes, and elements surrounding exploration and exploitation behavior are not simple at all. In fact, exploration–exploitation tradeoffs are considered one of the more fundamental challenges in our understanding of adaptive control and behavior (Cohen et al., 2007).

The theoretical analysis of exploration–exploitation tradeoffs is complicated in several ways. First, the concepts of exploration, exploitation, and a tradeoff between the two are used in a wide range of literatures, from animal behavior to human behavior and involving different terminologies, methodologies, and perspectives. The disparity and breadth of these concepts across such large and diverse literatures makes it difficult to synthesize existent knowledge into a coherent view. Researchers working in different areas operationalize exploratory and exploitative behavior in different ways, and this diversity motivates disagreements within conclusions about the essence of exploratory and exploitative behaviors. For example, researchers have debated assumptions about the exact elements that constitute exploratory and exploitative behavior, and about what defines a tradeoff between them. Second, exploration–exploitation tradeoffs may depend on a large number of environmental, individual, and social factors. The literature documenting these different factors is extensive, and there is little or no attempt to integrate their results (see Cohen et al., 2007; and Gupta et al., 2006, for laudable exceptions). Third, exploration and exploitation behaviors, and consequently, a potential tradeoff between the two, might not always be clearly identifiable. That is because these concepts can be described and understood on different spatial and temporal scales, as well as along different continua of behavior. Consequently, behaviors that might be understood as exploratory on one level of analysis might be seen as exploitative on another level, and even within a specific level of analysis, behaviors might have explorative and exploitative components.

Our goal is to provide an up-to-date synthesis of the exploration–exploitation literature by bringing together knowledge from different disciplines including human decision making, neuroscience, organizational learning, animal foraging, mate choice, and formal modeling approaches. To achieve this goal, we discuss the various challenges mentioned above and suggest a simple and straightforward framework for the theoretical analysis of potential tradeoffs between explorative and exploitative behaviors. Our synthesis illustrates the complexities surrounding these concepts and their tradeoffs. Our analyses highlight three elements needed in a unification of exploration and exploitation research: An exploration–exploitation continuum, different types of transitions between these two states, and the role of agents' goals in this process. These elements illustrate that the explore–exploit distinction often may not be a direct choice that an agent makes, but rather is an explanatory framework that researchers can apply to the agent's behavior to understand how to solves problems of the agent's interaction with the environment.

A Synthesis of Exploration–Exploitation Literatures

In an attempt to integrate a diverse and wide range of literatures, we organize our synthesis around three main themes: (a) concepts and definitions of exploration and exploitation; (b) environmental, individual, and social factors that influence exploration and exploitation; and (c) spatial and temporal scales that may influence how exploration and exploitation are conceptualized.

Concepts and Definitions of Exploration and Exploitation

Within the current literature, definitions of exploration and exploitation differ across at least three dimensions: *behavioral patterns* of the agent, *values and uncertainty* of the choice options, and *outcomes obtained* from a choice (cf. Todd, Hills, & Robbins, 2012). These three dimensions can be mapped to aspects of what a searching agent does, what the agent bases its decisions on, and what the agent gets out of the search. Focusing on these dimensions shapes the analysis and understanding of exploration and exploitation behavior in a particular situation, as well as any conclusions about their tradeoffs.

The agents' *behavioral patterns* are, perhaps not surprisingly, the most common dimension used to define exploration and exploitation in the animal foraging literature, where research often relies on the observation of behavior (Kramer & Weary, 1991; Nonacs, 2010). Behavioral patterns have also been considered, for example, in research on human information search (Hills, Todd, & Goldstone, 2010) and binary choice (Gonzalez & Dutt, 2011). In general, behavior is interpreted as exploration if it alternates between patches or options, is unfocused, and is variable over time. Behavior is interpreted as exploitation if it remains within a patch or option, is focused, and is stable over time. Therefore, the hummingbird remaining at its patch of flowers would be considered as exploiting the patch, while it would be considered as exploring if it alternates between patches (Nonacs, 2010). However, as we will discuss below, the distinction between exploration and exploitation is not always clear based on behavioral patterns alone. For example, the classification of a behavior as staying (exploit) rather than switching (explore) depends on the spatial and temporal scales of observation.

The *values* of choice options and the *uncertainty* associated with knowledge of those values are most prominently used to define exploration and exploitation in the RL literature. In some classic RL models, exploitation is defined as choosing the option that has the *higher sub-*

jective value and exploration is defined as choosing any other option at random (Sutton & Barto, 1998). For example, the prominent epsilon-greedy model mainly chooses the option with the greatest observed rate of reward (exploitation), but chooses an alternative at random (exploration) with some small probability of epsilon. The related epsilon-decreasing model allows the rate of exploration to change over time, but preserves the basic notions of expected value guiding exploitation and randomness guiding exploration. Other RL models have stressed the importance of *uncertainty* for defining exploration and exploitation. For example, in the restaurant scenario, exploration can be defined in terms of choosing an option with greater uncertainty, while exploitation is defined as opting for greater certainty (e.g., Lee, Zhang, Munro, & Steyvers, 2011). As pointed out in the neuroscience literature, the role of uncertainty might additionally be moderated by the agent's expectations (Aston-Jones & Cohen, 2005). For example, if you know that your machine periodically produces a faulty widget, one faulty widget will not stop you from using this machine (a situation that would be termed as *expected uncertainty* by Aston-Jones & Cohen, 2005). However, if the machine starts producing more and more faulty widgets (*unexpected uncertainty*), you might decide to give the other machine a try. Uncertainty and value also play a role in the animal literature, where exploration has been associated with choosing options with uncertain rewards and variable values, and exploitation has been associated with choosing options with known rewards and stable values (Krebs, Kacelnik, & Taylor, 1978).

Finally, exploration and exploitation have been discussed with respect to the *outcomes* that are *obtained* by the searching agent, which may include information, other types of resource rewards, or both. In many research areas, exploration is assumed to provide the agent with the opportunity for learning and obtaining information, while exploitation is assumed to provide explicit outcomes such as caloric or monetary rewards (neuroscience and RL: Cohen et al., 2007; foraging: Cook et al., 2013; decision making: Hills & Hertwig, 2010; organizational learning: March, 1991). For example, in "observe-or-bet" tasks (Navarro & Newell, 2014; Rakow, Newell, & Zougkou, 2010; Tversky & Edwards, 1966), participants can either obtain information (explore) or obtain monetary rewards (exploit). In each trial of these tasks, participants choose between (a) observing which of two lights comes on, and thereby gaining information about the underlying probabilities, and (b) betting on which light will come on, and thereby receiving rewards if they guess correctly. More importantly, if they choose to observe, participants receive no reward, and if they choose to bet, they receive no feedback as to which light comes on, thereby allowing the researchers to distinguish between "pure" exploration and exploitation with respect to the observed outcomes. However, the distinction between information and rewards is not always as straightforward as in this example. In most real-life situations, rewards tend to also provide the agent with information about the quality of the selected option (Gupta et al., 2006) and hence the distribution of rewards available. In some situations, agents might receive information about foregone payoffs in nonselected alternatives (Yechiam & Busemeyer, 2006); and even when no material rewards are obtained during exploration in a given situation, information search in itself can be rewarding if it delivers positive experiences (Denrell & Le Mens, 2011; Gonzalez & Dutt, 2012; Mehlhorn, Ben-Asher, Dutt, & Gonzalez, 2014).

The three dimensions we highlight here are by no means mutually exclusive. Most conceptualizations of exploration and exploitation are based on more than one of them. However, the respective contributions of these dimensions to a given conceptualization are not always clear and this can be especially problematic if the considered dimensions lead to opposing conclusions. An example of this can be found in the recent human decision making literature, where researchers have disagreed about the interpretation of exploration and exploitation behavior in a popular "sampling paradigm" of binary choice (Gonzalez & Dutt, 2011, 2012; Hills & Hertwig, 2010, 2012). In this paradigm, participants can sample outcomes without consequences from two choice options for as long as they want, before making a final consequential choice between the two (Hertwig, Barron, Weber, & Erev, 2004). Hills and Hertwig (2012) argue that this paradigm presents "pure exploration" during the sampling phase, because participants only receive information while it presents “pure exploitation” at the choice phase because they receive only rewards. In contrast, Gonzalez and Dutt (2012) argue for a gradual transition from exploration to exploitation *within* the sampling phase, because they find a decrease in alternation between the options over the course of sampling. One way of accounting for these two different perspectives on the data is that the first arises from a focus on *obtained outcomes* (information vs. rewards) as the defining dimension, while the second comes from a focus on *behavioral patterns* (which options are being sampled). But the complete picture must include consideration of how a particular search fits into a sequence of searches over time, where rewards from one search may become information to guide later searches and hence influence what is exploration versus exploitation in a particular behavioral pattern (a hierarchical view we return to in the Discussion).

In addition to affecting our understanding of exploratory and exploitative behaviors, assumptions about the underlying dimensions may also influence the nature of the considered tradeoff. For example, a distinction between exploration and exploitation based on *behavioral patterns* places the costs and benefits of staying at a resource versus switching to another one in the spotlight (Charnov, 1976), while a distinction based on *values and uncertainty of the choice options* may switch the focus to value or risk. For example, trading off the risk of exploring a new option with uncertain values against the safety of exploiting the current option with known values might require a balance between risk-seeking and risk-aversion tendencies, which has been studied in humans (Lee et al., 2011) and in animals (Tuttle, Wulfson, & Caraco, 1990). In terms of *obtained outcomes*, an exploration–exploitation tradeoff can be very similar to a tradeoff between speed and accuracy (in humans: Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; in animals: Chittka, Skorupski, & Raine, 2009), or accuracy and effort (Payne, Bettman, & Johnson, 1992). This is the case because longer exploration demands more time and effort, but generally leads to more information and, therefore, to better choices (but see Fiedler & Kareev, 2011; Gigerenzer & Gaissmaier, 2011; Hertwig & Pleskac, 2010; and Lee & Corlett, 2003, for a discussion of the less-is-more effect: situations where less information can be advantageous and, therefore, no tradeoff might exist).

Environmental, Individual, and Social Factors

The literature points us to at least three types of factors that influence exploration–exploitation behaviors and tradeoffs: *environmental*, *individual,* and *social* factors.

Environmental factors. A large number of *environmental factors* influence the way in which exploration and exploitation may or may not trade off. Table 1 provides an overview of such factors, together with exemplary references. In general, exploration is particularly relevant and useful when the resources in the currently selected option are depleted, thereby reducing the value of the current option, or when the agent has insufficient information about the state of alternative options, thereby increasing the value of information search. For example, if resources deplete faster than they replenish, agents may eventually be forced to switch from exploiting the current option to exploring new options (depending on depletion rate, switching costs, agent life span, etc.), while in environments where resources do not deplete, an agent could continue to exploit a selected option indefinitely (Charnov, 1976). At the same time, if agents have sufficient information about the alternatives, because they receive information about foregone payoffs, they might be tempted to reduce exploration (Steiner & Redish, 2014) or show an increased propensity to take risks (e.g., Yechiam & Busemeyer, 2006; Yechiam, Rakow, & Newell, 2015).

The usefulness of exploration further depends on its costs and benefits relative to the costs and benefits of exploitation (and the costs of switching between the two), as well as on the structure and distribution of resources and costs in the environment. For example, animals reduce their foraging efforts as predation risk increases (Verdolin, 2006). In addition, the number of available options and the distribution of their features can lead agents to explore more or less, or even possibly to defer their choice altogether (Fasolo, Hertwig, Huber, & Ludwig, 2009; Scheibehenne, Greifeneder, & Todd, 2010), depending on how they trade off the value of exploitation versus further exploration (Schwartz, 2004) and on their past history of

Table 1
Overview of Different Environmental Factors and Examples of Their Relevance for Exploration and Exploitation

Factor	Influence on exploration and exploitation
Depletion and replenishment of resources	Depletion of resources requires switching to new resources, while nondepleting and replenishing resources allow for continuing or resuming exploitation (Charnov, 1976).
	Even if patches do not deplete, exploration might be adaptive to increase information and reduce boredom (Cohen et al., 2007).
	The ratio of depletion relative to replenishment determines the rules foragers should use to decide when to leave a patch (Nonacs, 2010).
Available information about the options	Optimal stopping behavior and search/choice strategies are affected by the available information; ranging from situations with “pure” information/reward for the selected option as in observe-or-bet tasks (Rakow et al., 2010) or the sampling paradigm (Hertwig et al., 2004), over a mixture of information and rewards for the selected option as in multiarmed bandit problems (Gittins, 1979), to a combination of information and rewards for the selected option and information about foregone payoffs in not-selected options (Erev & Barron, 2005; Yechiam & Busemeyer, 2006).
	Foregone payoff information can induce regret and reduce future exploration in rats (Steiner & Redish, 2014).
	People who receive foregone payoff information show a higher propensity to take risks (Yechiam & Busemeyer, 2006; Yechiam, Rakow, & Newell, 2015).
Costs of information vs. value of reward	The usefulness of exploration and exploitation depends on the costs of exploration as well as the costs of switching between the two (Charnov, 1976).
	Exploration tends to decrease with increasing costs of search and decreasing values of possible reward (Fu & Gray, 2006; Gigerenzer et al., 2012; Hau et al., 2008).
	Chipmunks increase the time spent exploring alternatives as the quality of the currently exploited patch decreases (Kramer & Weary, 1991).
	Participants tend to switch to noncompensatory decision strategies, which involve less exploration, when costs of information search increase (Bröder, 2000; Newell & Shanks, 2003).
Structure of the environment	Structure can affect the type of exploration, with well-structured environments inviting more controlled, goal-directed exploration (Cohen et al., 2007; Lee et al., 2011).
	Different exploration and exploitation behavior is adaptive in patchy environments vs. distributed environments (Hills et al., 2013); e.g., bees make foraging decisions after only a few visits to nearby flowers, likely due to highly autocorrelated distribution of rewards that makes short exploration adaptive (Real, 1992).
	Agents explore more or less as a function of the number of available choice options and the distribution of their features (Fasolo et al., 2009; Scheibehenne et al., 2010).
Probability of “gains” vs. “losses”	People tend to overexplore in environments containing “rare disasters” and underexplore in environments containing “rare treasures” (Teodorescu & Erev, 2013).
	Foraging animals reduce their foraging efforts as predation risk increases in their environment (Verdolin, 2006); e.g., female fiddler crabs reduce the amount of mate search in response to increased predation risk (Booksmythe, Detto, & Backwell, 2008).
Stability and predictability of the environment	The need for exploration increases with reduced stability/predictability of the environment (neuroscience: Gold & Shadlen, 2007; reinforcement learning: Kaelbling et al., 1996; foraging: Kramer & Weary, 1991).
	However, if the environment is too dynamic/unpredictable, exploration can become dysfunctional, because obtained information is no longer valid (March, 1991; Todd & Miller, 1999).
	In stable environments, foragers can use simple fixed-time or fixed-number rules to decide when to leave a patch, while more sophisticated rules are needed in dynamic environments (Nonacs, 2010).

Table 1 (*continued*)

Factor	Influence on exploration and exploitation
Shape of underlying payoff distributions	The payoff distribution affects how much exploration is needed. For example, small samples can be advantageous for unimodal payoff distributions (Hertwig & Pleskac, 2010), while large samples are required to correctly identify a bimodal distribution. Animals seem to be sensitive to the shape of the distribution, as suggested by the fact that when a species is beginning to diverge, choosiness of individuals (i.e., exploration) increases; e.g., in butterflies (Friberg et al., 2008) and fish (Gabor & Ryam, 2001; Rundle & Schluter, 1998).
Horizon	Optimal stopping behavior and search/choice strategies are affected by the horizon of a choice problem (Kaelbling et al., 1996); with horizons ranging from finite (where the agent knows there will be *n* choice-episodes: Lee et al., 2011), to uncertain (where the agent knows there will be somewhere between *n* and *m* episodes: Seale & Rapoport, 2000), to infinite (where the agent knows there is a probability of any episode being the final one, but the actual number of episodes is neither known nor constrained to fall within any range: Gittins, 1979).
Internal (memory) vs. external (environment)	Similar exploration and patch-leaving rules are used in memory and in the environment (Fu & Pirolli, 2007; Hills, Todd, & Goldstone, 2008; Wolfe, 2012). Some characteristics of search behavior are similar for internal and external search tasks (Wilke et al., 2009).

exploration (e.g., with smaller or larger choice sets—Hills, Noguchi, & Gibbert, 2013). Most interestingly, exploration and exploitation do not seem to systematically vary much as a function of whether they occur in internal (i.e., memory) or external environments (e.g., Wilke, Hutchinson, Todd, & Czienskowski, 2009). This might be due to the exaptation of search strategies originally evolved from external search to internal search (Hills, 2006), and relatedly to the fact that the structure of information in memory typically reflects the structure of information in the environment (Anderson, 2007).

Individual factors. A variety of *individual factors* can affect an agent's explorative and exploitative behaviors and their tradeoffs. Table 2 summarizes this literature. For example, exploration tends to increase with an agent's cognitive capacity, aspiration level, physical strength, and reduced levels of dopamine. It tends to decrease with age, prior knowledge about the distribution of payoffs, and high current resource levels. While some of these factors are relatively dynamic and can change between different decision situations (e.g., experience with the task at hand and current energy level), others are more stable and persist across situations (e.g., ones' morphology and working memory capacity). An example of individual differences that have been proposed as more stable aspects of exploration/exploitation tendencies is the distinction between "maximizers" and "satisficers," with maximizers searching for more information and being more likely to defer choices altogether (Schwartz, 2004).

Social factors. *Social factors* of exploration/exploitation tradeoffs are often overlooked in the human decision making literature (for recent exceptions, see Goldstone, Ashpole, & Roberts, 2005; Phillips, Hertwig, Kareev, & Avrahami, 2014; Schulze, van Ravenzwaaij, & Newell, 2015). In natural settings, however, agents rarely act in isolation from others, and exploration–exploitation tradeoffs can be different when considered at the group level than when considered for the individual agent. For example, the organizational learning literature has suggested that it might be easier for an organization to simultaneously balance exploration and exploitation than for an individual (Gupta et al., 2006). Similarly, the investigation of socially foraging species, such as ants, suggests that groups of animals can minimize exploration–exploitation tradeoffs by having the individuals specialize in either exploration or exploitation and working together (Cook et al., 2013). For example, individuals of many bird species specialize as either "producers," who explore and find their own food, or as "scroungers," who join other birds that have already found food. Changes in the obtained

Table 2
Overview of Different Individual Factors and Examples of Their Relevance for Exploration and Exploitation

Factor	Influence on exploration and exploitation
Cognitive capacity	Exploration increases with working memory capacity (Hills & Pachur, 2012; Rakow et al., 2008), numeracy, and self-reported ability for rational thinking (Lejarraga, 2010). But *low* working memory capacity can be adaptive in dynamic environments where outcomes change over time (Brydges, Heathcote, & Braithwaite, 2008); when the decision maker relies on "simple" heuristics that require little information, such as the recognition heuristic (Schooler & Hertwig, 2005); and when contingencies between decision options need to be detected (Fiedler & Kareev, 2006; Gaissmaier, Schooler, & Rieskamp, 2006).
Aspiration levels	Higher individual aspiration levels, such as the *satisficing threshold* (Simon, 1990), the *decision threshold* (Newell & Lee, 2011), or the *desired level of confidence* (Hausmann & Läge, 2008), correspond to more exploration. Self-reported "maximizers" search for more information and have a higher likelihood to defer choice than "satisficers" (Parker, Bruine de Bruin, & Fischhoff, 2007; Schwartz et al., 2002). Mating aspiration levels and resulting exploration decrease with the number of available mates and one's own attraction on the mating market in fish (Borg, Forsgren, & Amundsen, 2006) and humans (Beckage, Todd, Penke, & Asendorpf, 2009; Todd & Miller, 1999).
Current state of the individual	Increased energy levels reduce the time animals allocate to foraging and increase the time they spent in vigilance to predators (Bachman, 1993; Kotler, Brown, & Bouskila, 2004). Negative energy budgets result in preferences for variable or unknown outcomes (i.e., exploration) and positive energy budgets result in preferences for stable and known outcomes (i.e., exploitation; Bacon et al., 2010; Caraco, 1981; Caraco et al., 1990). Depression can increase exploration (von Helversen, Wilke, Johnson, Schmid, & Klapp, 2011; Blanco, Otto, Maddox, Beevers, & Love, 2013).
Prior experience and knowledge about environmental and social characteristics of the situation	Prior expectations about risks and payoffs can affect preferences and duration of exploration (Denrell & March, 2001; Gonzalez & Dutt, 2011; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012). Laypeople explore more information than domain experts, because experts know what is relevant (Gigerenzer & Gaissmaier, 2011). Humans and animals are more likely to defer choice (i.e., neither explore nor exploit) when they have insufficient knowledge (Perry & Barron, 2013).
Morphology	Larger individual Atlantic salmon are more likely to explore than smaller individuals (Armstrong, Braithwaite, & Huntingford, 1997). Use of public information instead of risky personal patch exploration occurs in nine-spined sticklebacks, but not in three-spined sticklebacks; probably because nine-spined sticklebacks are at higher predation risk due to a morphological difference (Coolen, van Bergen, Day, & Laland, 2003). Sensory sensitivity of honey bees is related to foraging task specialization and resource exploitation (Riveros & Gronenberg, 2010).
Demographics	Humans may reduce exploration with increased age (Mata, Wilke, & Czienskowski, 2013) and show reduced exploration in girls compared to boys (Slovic, 1966). Humans show individual differences and a change in the discounting of future rewards across their lifespan (Green, Myerson, & Ostaszewski, 1999). There are sex differences in the amounts of exploration and exploitation for many animals; however, the sex doing the most exploring or exploiting is not consistent across species.

Table 2 (*continued*)

Factor	Influence on exploration and exploitation
Levels of neurotransmitters hormones	Dopaminergic activity levels are related to interindividual differences in exploratory behavior for animals and humans, with high dopamine associated with little exploration and low dopamine associated with higher levels of exploration (Hills, 2006). Interindividual variation in exploratory behavior in animals has been linked to testosterone levels, but the evidence is mixed (Kellam, Lucas, & Wingfield, 2006; Mutzel et al., 2011).

food types or changes in group composition can alter the exploratory behavior of individual birds (e.g., scroungers might act as producers during a temporary absence of other producers from the group) and thereby optimize food intake at the group level (Giraldeau & Lefebvre, 1986).

An overview of factors that affect exploration and exploitation in social settings is provided in Table 3. Two key factors are competition and social information. In general, competition increases the costs of exploration because it creates the risk that other agents might exploit the reward before you (e.g.,Goldstone et al., 2005; Todd, 2007). Consequently, agents may stop their exploration and switch to exploitation sooner in the presence of competitors. At the same time though, exploration might be especially useful in such competitive situations because it can help the agent to find options that competitors do not or cannot exploit (e.g., leading to dispersal to new resource patches). Availability of social information, in contrast, tends to decrease the costs of exploration because it allows the agent to learn from others' decisions and performances and thereby reduces the need to engage in risky exploratory behavior oneself (Valone, 2007). A third social factor becomes relevant if the searching agent must make mutual choices with others who are also exploring (e.g., Todd & Miller, 1999). This factor may be most obvious for mate choice, where an agent's ability to explore or exploit potential mates is often limited by the mate's willingness to engage in such an interaction, but it also plays a role in other situations of mutual search and

Table 3
Overview of Different Social Factors and Examples of Their Relevance for Exploration and Exploitation

Factor	Influence on exploration and exploitation
Competitiveness of the environment	Animals foraging in groups distribute themselves across patches based on patch profitability and competitive pressure (fish: Godin & Keenleyside, 1984; ducks: Harper, 1982). Exploration is less costly in noncompetitive environments than in competitive environments; e.g., during mate choice in fish (Lindstrom & Lehtonen, 2013) and humans (Todd, 2007; Goldstone, Ashpole, & Roberts, 2005). The presence of competitors reduces information search in the "sampling paradigm" (Phillips et al., 2014). Exploration might be especially useful in competitive environments because it can help the agent to find a "niche" where he is better adapted than his competitors (Cohen et al., 2007; Schulze et al., 2015).
Availability of social information	Honeybees share information about food-patch quality, thereby reducing others' need for exploration (Biesmeijer & de Vries, 2001). The use of public information can improve the estimation of patch quality and prevent the underutilization of resources (Valone, 1989). Observing others provides counterfactual information about strategies not yet pursued by oneself (Montague, King-Casas, & Cohen, 2006). Under certain circumstances, the use of public information can be suboptimal and misleading (Giraldeau, Valone, & Templeton, 2002).
Mutual exploration and exploitation	Optimal mate search strategies strongly differ as a function of whether or not mutual search is taken into account (Todd, 2007; Todd & Miller, 1999).

choice, such as job hunting or apartment search, where two parties must agree for the search to conclude successfully.

Interactions between factors. Not surprisingly, factors of the different types interact with one another in multiple ways. Often, this manifests in individual factors interacting with aspects of the environmental structure. For example, a low working memory capacity limits the agent's ability to keep track of obtained information and will therefore be disadvantageous in many exploration scenarios (e.g., Hills & Pachur, 2012): however, it can become adaptive in dynamic environments where rewards rapidly change and where previously learned information is no longer valid at a later point (Fiedler & Kareev, 2006). Another interaction can be seen when maximizers face environments with many options to choose from (e.g., in consumer settings). They tend to engage in longer exploration than satisficers but seem to be less happy with the (objectively better) outcomes, which can lead to further exploration (Schwartz, 2004).

Most importantly, it is not always clear to what extent interindividual differences in exploratory and exploitative behavior are caused by characteristics inherent to the decision maker, as opposed to environmental or social factors. For example, in the human decision making literature, it has been debated whether characteristic alternation patterns during information search are due to individual preferences of the decision maker (Hills & Hertwig, 2012) or due to the distribution of rewards in the environment (Gonzalez & Dutt, 2012).

Spatial and Temporal Scales of Exploration and Exploitation

A third important issue to understanding exploration and exploitation is represented by the spatial and temporal scale of observation. As we discuss in the subsequent section, what might be considered exploration at one level of observation could be considered exploitation at another level (cf. Cohen et al., 2007), and consequently, a tradeoff considered at one level might change or even disappear if considered at a different one. This conceptualization is consistent with the continuum idea that we develop in the Toward a Unifying Theory of Exploration–Exploitation Tradeoffs section and also reinforces the notion that the level of abstraction from which the 'tradeoff' is observed can often define whether an agent is interpreted as making a distinct choice between the two types of behavior.

Spatial scale. To illustrate the importance of spatial scale, consider the foraging hummingbird. According to foraging theory, "what the forager does from the point it enters a patch to when it leaves is its 'patch exploitation' behavior" (Nonacs, 2010, p. 683). But at what level do we define the "patch"? At a very broad level, a whole meadow full of different clumps of flowers might be considered a patch, and the bird would be considered to exploit as long as it stays in the meadow. At a much finer level, a single clump, or even a single flower, could constitute a patch. Remaining at the flower would then be considered exploitation, while switching between flowers could be seen as exploration. As this example shows, our understanding of a situation again depends on the considered conceptual dimensions of exploration and exploitation. Scale is particularly relevant for the classification of *behavioral patterns*. If we instead consider other dimensions, such as the *obtained outcomes*, even the exploitation of a single flower has an exploratory component, because in addition to caloric rewards, the bird obtains information relevant to this flower and others nearby. (Note that if the environment is not patchy but rather resources are divided uniformly, scale may matter less and the searcher may not switch between exploration and exploitation; Hills, Kalff, & Wiener, 2013.)

These ideas on spatial scale are further explored in Figure 1, where we illustrate three exemplary spatial levels for the scenario of human mate search. The broadest level, Level 3, consists of all groups of potential partners; Level 2 consists of a selected group of particularly interesting people; and Level 1 consists of one particular person. Depending on the conceptual dimensions being considered, both exploration and exploitation might occur at each level. For example, an agent can search for information about groups of people or about a single person (see Saad, Eba, & Sejean, 2009 for a formal model of this idea of nested mate search across and within potential partners). Similarly, the agent can "exploit" the benefits by dating different partners in a select group or

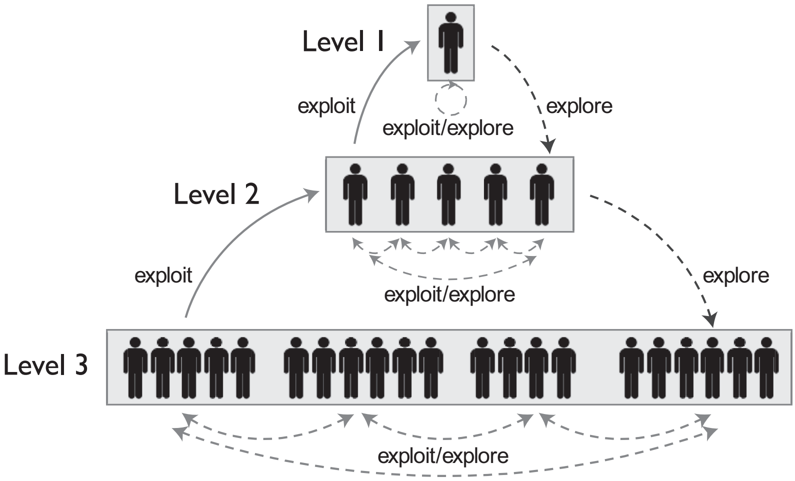

Figure 1. Illustration of human mate search at different levels. At the broadest level, one might consider groups of potential partners; while, at the narrowest level, one might consider one particular individual. Within each level, behavior can be explorative and exploitative.

by being married to one spouse. A differentiation between exploration and exploitation might occur in the transition between different levels. While "downward" transitions to larger groups mark a broadening focus that is often associated with exploration (e.g., getting divorced and starting to look for a new partner), "upward" transitions mark a narrowing focus often associated with exploitation (e.g., deciding to stop search and get engaged).

The relevance of the spatial scale for transitions between exploration and exploitation has been prominently demonstrated in the optimal foraging literature, where research investigates how agents make appropriate global decisions between locally depleting resource patches (Charnov, 1976; Stephens & Krebs, 1987). Such local-to-global transitions are not exclusive to foraging animals, but also occur in human memory search (Hills, Jones, & Todd, 2012). The role of spatial (and temporal) scales has also received attention in literature on habitat selection, which is concerned with how agents, such as grazing animals, select resources (Mayor, Schneider, Schaefer, & Mahoney, 2009). Some relevant conclusions here are that the scale of measurement affects the interpretation of results, that habitat selection (and thereby exploration–exploitation transitions) cannot simply be extrapolated across scales, and that different species select habitat at different scales (Mayor et al., 2009).

In combination with the considered conceptual dimensions, scale of measurement can affect our interpretation of exploration and exploitation situations. While some tradeoffs might change as a function of the considered scale as we discuss in more detail below, others might "scale" between different levels (e.g., explore between single partners vs. explore between groups of people). Moreover, scale becomes less relevant for the classification of some types of foraging behavior that are described as random walks with variable step size drawn from a scale-free distribution, known as Lévy walks or Lévy flights. These proposed Lévy foraging processes generate behavioral patterns that look the same on any scale, with no sensitivity to (or memory of) whether or not resources are actually found in a particular location and hence no clear distinction between exploration and exploitation. As such, they stand in contrast to area-restricted search mechanisms that switch gradually between exploitation and exploration as resources are found. The prevalence and adaptive performance of both mechanisms is under debate (Benhamou, 2007; Hills et al., 2013; Humphries & Sims, 2014).

Temporal scale. Much like for spatial scale, an agent's decisions can be described and understood on different temporal levels. At the finest, most local (short-term) level, only one single decision might be considered, while at the broadest, most global (long-term) level, all decisions over the agent's lifetime might be taken into account. Temporal scale is especially important for the classification of a behavior as exploratory or exploitative with respect to an agent's *behavioral patterns*, because such patterns can only be observed over time. However, the temporal scale can also be relevant on other conceptual dimensions. For example, while an agent might seem to choose randomly between options when observed on a short-term time scale, a longer-term scale might reveal that the agent follows an elaborate search strategy with respect to the options' *values* or *outcomes*. Temporal scale is also important in situations with a more gradual transition between exploration and exploitation, as the considered time window can determine whether a behavior within that window will be characterized as explorative, exploitative, or both.

An interesting demonstration of the relevance of temporal scale has been provided by Vul, Goodman, Griffiths, and Tenenbaum (2009). Using a Bayesian ideal observer model, they tested to what extent a suboptimal solution to a tradeoff at one level can be optimal at another level. They showed that exploring too little information to make the right choice in the current situation (i.e., suboptimal behavior on a local time scale) can maximize the number of exploitative decisions a person can make in the long run and thereby optimize behavior on a global scale.

The temporal scale is furthermore relevant for the agent's ability to learn. Over time, the information and rewards obtained in single decisions will affect the agent's expectations, behavior, and even the availability of resources in the future (Gold & Shadlen, 2007). A formal demonstration of the interrelatedness of choices through learning over time comes from Denrell and March (2001), whose "hot-stove effect" shows how previous negative experiences with an option can bias agents against selecting this option in the future. This connects to the broader issue of how agents learn to make appropriate transitions between exploration and exploitation. While people appear to be adept at learning this in some situations (Gupta et al., 2006; Sang, Todd, & Goldstone, 2011), they seem to miss relevant patterns in others (e.g., Hutchinson, Wilke, & Todd, 2008, whose participants did not learn the relevant structure of the patchy environment they were searching in).

Finally, temporal scale is related to the "horizon" of a choice sequence. In many situations, decision makers do not know how often they will make a similar choice in the future, thereby making it particularly difficult to weigh the costs and benefits of gathering information against obtaining rewards. In the RL literature, such situations are called "infinite horizon" problems (where the horizon may also be indefinite or unknown) in contrast to "finite horizon" problems, such as the restaurant choice scenario from our introduction (Kaelbling, Littman, & Moore, 1996). Uncertainty about the horizon has also been introduced in some behavioral search tasks, rendering them more ecologically realistic. For example, Seale and Rapoport (2000) introduced variability into the length of search in the secretary problem, and Hutchinson et al. (2008) kept participants from knowing how long they would be performing patch leaving search. As this literature shows, (optimal) exploration–exploitation behavior can differ considerably between horizons, thereby making the horizon an important factor for understanding exploration–exploitation tradeoffs (see also Table 1).

Scale and the avoidance of a tradeoff. As suggested by the literature reviewed so far, some tradeoffs might change when considered at different scales, while others remain consistent. But can tradeoffs also disappear at some scales? On a local scale, agents might be able to sidestep having to make an active tradeoff between exploration and exploitation by avoiding making a decision altogether in the current situation. Several decision-avoidance phenomena have been described in humans (Anderson, 2003) and animals (e.g., Aw, Vasconcelos, & Kacelnik, 2011; Hill, Hollis, & Wells, 2001; Perry & Barron, 2013). Perhaps the strongest argument for tradeoff avoidance is found in *choice deferral* in humans (Dhar, 1996), where decisions are postponed or refused altogether. A related phenomenon of *opting out* of a decision has been reported in the animal literature, where animals ranging from nonhuman primates to dolphins, dogs, rats, and honeybees have been shown to refuse a potentially consequential decision in light of insufficient information (Perry & Barron, 2013). On a narrow scale, such choice deferrals might indeed represent a form of tradeoff-avoidance (Luce, 1998) because the agent decides to neither explore nor exploit.

However, a tradeoff might reappear as soon as a broader scale is considered, for example, revealing that the choice was deferred in the current situation in order to search for more information elsewhere that can be used to guide decisions in later situations.

The *status quo bias* (Samuelson & Zeckhauser, 1988), *omission bias* (Ritov & Baron, 1992), and *default heuristic* (Johnson & Goldstein, 2003) refer to situations in which people prefer options that cause no change in the state of the world or which require no overt action on the decision maker's part. A related phenomenon from the foraging literature is *flower consistency*: Once bees have learned about one species of flower, they tend to stick with it, even if other types of flowers would be more rewarding (Hill et al., 2001; Hill, Wells, & Wells, 1997). Another related phenomenon in animals and humans is the *sunk cost effect* (also known as the Concorde fallacy), where an agent sticks to an option where it had previously invested resources even though it would be better to abandon that option and explore alternatives (in humans: Arkes & Ayton, 1999; in starlings: Aw et al., 2011). Similarly, *learned helplessness* can be seen as avoiding exploration in favor of the current unsatisfactory situation (Teodorescu & Erev, 2014). Finally, *inaction inertia* (Tykocinski & Ortmann, 2011) refers to the tendency to omit action when a similar and more attractive opportunity has been foregone (e.g., not buying a particular pair of shoes now because you had seen it earlier at a much-reduced price). Again, all of these phenomena might be interpreted as avoiding having to consider a tradeoff in the near term by adopting a choice (or no choice) already made, but on a longer scale, they can be interpreted as exploring what happens when the default (or no) choice is made or exploiting the outcome of that default (or no) choice.

Another way of escaping or at least reducing a tradeoff between exploration and exploitation might be to do both types of search at the same time. As described above, at a larger social scale, different individuals can specialize in either exploration or exploitation, and thereby at least reduce exploration–exploitation tradeoffs at the group level (Cook et al., 2013; Gupta et al., 2006). Whether or not tradeoffs can be avoided completely, and whether or not an individual human can similarly use parallel internal processes of exploration and exploitation remains to be determined (cf. Herzog & Hertwig, 2009).

Toward a Unifying Theory of Exploration–Exploitation Tradeoffs

The literature review and synthesis presented earlier illustrates the conceptual and measurement difficulties surrounding exploration, exploitation, and their tradeoffs. In the remainder of the paper, we identify and discuss elements that need to be considered in future theories and models to attempt a more comprehensive account and unification of the many facets of exploration and exploitation.

An Exploration–Exploitation Continuum

A main conclusion from the different conceptual dimensions discussed above is that exploration and exploitation do not necessarily represent qualitatively distinct behaviors that need to be traded off against each other, nor is it necessarily best to consider these behaviors as explicit choices. The interpretation of choice and behavior often depends on the scale of observation. While many conceptualizations of the tradeoff assume a binary distinction, recent reviews argue that exploration and exploitation might be better understood as end points of a continuum (Cohen et al., 2007; Gupta et al., 2006). This idea of a continuum seems to be in line with our synthesis of the conceptual dimensions. As shown in Figure 2, within each dimension, it is possible to consider situations at the end points of a continuum, where behavior can be described as "pure" exploration or "pure" exploitation. With respect to *behavioral patterns*, an agent on the extreme ends of the continuum might constantly switch between options (exploration) or might remain at one option over time (exploitation). For example, a hummingbird could frequently move between patches of flowers or stay within one patch. With respect to *values and uncertainty*, the agent might choose an option with unknown or low subjective values and high uncertainty (exploration), or an option with high subjective values and low uncertainty (exploitation). For example, a car buyer could look at a fancy car they had never heard about before or at the new version of their less exciting, but reliable previ-

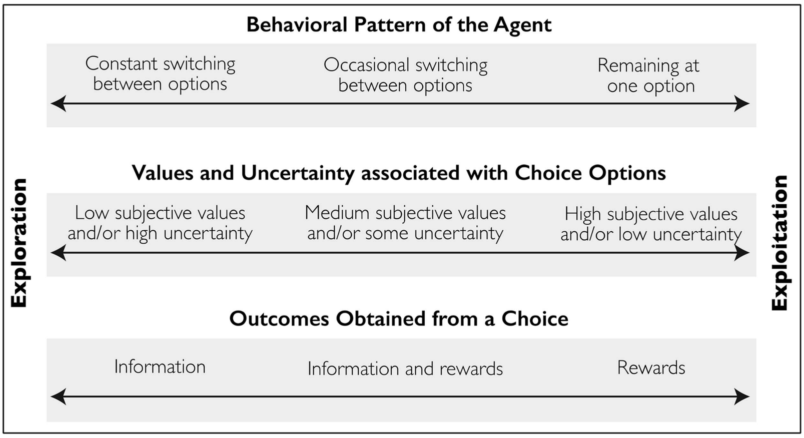

Figure 2. Illustration of an exploration–exploitation continuum for three conceptual dimensions: behavioral patterns of the agent, values and uncertainty related to the choice options, and obtained outcomes.

ous model. Finally, with respect to *obtained outcomes*, the agent might choose an option from which they obtain information but no actual rewards (exploration) or an option that gives rewards but no information (exploitation), as in the "observe-or-bet tasks" described above.

Although it is thus possible to think of situations at the end points of the exploration–exploitation continuum, many situations fall in between the extremes. The hummingbird might alternate more or less frequently between patches, the car buyer might look at cars with medium subjective values and uncertainty, and participants in an experiment might receive feedback information and rewards in a given trial. According to this continuum perspective, an agent can engage to some extent in both exploration and exploitation at each given point in time. Our ability to judge where a given action lies on this continuum could be improved by considering more than one conceptual dimension. Think again of the foraging hummingbird. When feeding from a flower, the bird obtains information about the flower's quality and its caloric rewards, thereby rendering exploration and exploitation virtually indistinguishable on the dimension of *obtained outcomes*. However, if it alternated frequently between patches of flowers or if it had a high uncertainty about a selected patch's quality, its behavior would likely be considered as exploratory (Bacon, Hurly, & Healy, 2010).

Transitions Between Exploration and Exploitation

A potential continuum between exploratory and exploitative behaviors is directly related to the question of how agents transition between these behaviors. Two kinds of transitions are most commonly addressed in the tradeoff literature, and research often focuses on tasks that emphasize one or the other (cf. Sang et al., 2011). The first is a transition from exploration to exploitation, such as ant colonies exploring options for a future nest site and then settling at a site (Pratt & Sumpter, 2006), travelers exploring alternative routes before deciding to take one (Fu & Gray, 2006), participants sampling from different options before deciding to choose one (Hertwig et al., 2004), or an employer interviewing potential candidates for a job as in the well-known secretary problem (also see Ferguson, 1989 for other types of optimal stopping problems). The second transition goes in the opposite direction, from exploitation to exploration. Examples include a scientist abandoning their current research topic and deciding to explore new ideas (Cohen et al., 2007), an animal such as a hummingbird or a chipmunk leaving its current patch of food to search elsewhere (Kramer & Weary, 1991), or a person giving up on retrieving information about a particular topic in favor of attempting to retrieve information about other topics, either from memory (Hills et al., 2012) or from the world, including online (Fu & Pirolli, 2007; Pirolli & Card, 1999). As we will discuss now, those two directions might require very different kinds of tradeoffs.

During an *exploration to exploitation tradeoff*, the agent faces the question of how long to continue exploration and thereby increase the amount of information obtained, and when to switch to exploitation and thereby increase the chance to obtain actual rewards. A variety of models have been developed for this tradeoff in many different contexts. They range from the idea of satisficing (Simon, 1990) and derived Bayesian satisficing models (Fu & Gray, 2006) to sequential heuristic models of optimal stopping (Seale & Rapoport, 2000), mutual mate choice (Todd & Miller, 1999), Bayesian optional stopping models (Edwards, 1965), Bayesian observer models (Vul et al., 2009), the accumulation of evidence to a threshold criterion (Busemeyer & Townsend, 1993; Ratcliff, 1978; Vickers, 1979), and models based on cognitive architectures for dynamic decision making (Gonzalez & Dutt, 2011; Gonzalez, Lerch, & Lebiere, 2003)

During an *exploitation to exploration tradeoff*, the agent faces the question of how long to continue exploiting a current option and thereby obtain its rewards, and when to switch to exploring alternatives and thereby increasing the chance to find potentially better options elsewhere. Perhaps the most prominent model of this tradeoff, the *marginal value theorem* (MVT; Charnov, 1976), comes from the foraging literature where research has long attempted to formalize optimal behavior (Stephens & Krebs, 1987). It proposes that foragers should abandon the exploitation of a food patch as soon as its expected rate of future reward falls below the expected rate of reward in the environment as a whole. It has been successfully applied not only to animal foraging (Pleasants, 1989), but also to human information search on the Internet (Pirolli & Card, 1999), in memory (Hills et al., 2012), and in visual displays (Wolfe, 2013).

A third major variant of the tradeoff has been discussed for situations without a clear direction for the transition. For example, March (1991) describes exploration–exploitation tradeoffs in the organizational context, which require a constant balance of both exploratory and exploitative tendencies. Perhaps the most prominent model of such a balance is provided by the Gittins Index (Gittins, 1979). It predicts whether a decision maker should exploit or explore for each point in time. The index is based on a formal analysis of multiarmed bandit problems and represents a useful benchmark for the analysis of a variety of behaviors, ranging from humans playing ultimatum games (Brenner & Vriend, 2006) to birds (great tits) sampling different food sources or committing to the most profitable one (Krebs et al., 1978).

Gradual transitions. The variations of the tradeoff discussed so far share the underlying assumption that the transition between exploration and exploitation is marked by a qualitative "switch" from one behavior to the other. However, as illustrated in our discussion of exploration–exploitation continua, a qualitative distinction between the behaviors is not always easy to determine or it might not even exist in some situations. Instead, the behaviors can occur concurrently or agents can gradually transition over time. Perhaps the most detailed investigation of such a gradual transition is provided in the foraging literature on area-restricted search, which has shown that foragers tend to shift gradually from exploitation, marked by small "steps" with high-angle turns when resources have been found, to exploration, marked by larger straighter steps as resources are no longer found (Bell, 1991; Hills, 2006).

Further evidence suggesting a gradual transition between exploration and exploitation comes from the neuroscience literature. Aston-Jones and Cohen (2005) have associated the transition with patterns of release of the neurotransmitter, norepinephrine, from the locus coreolus (LC), namely, spikes of release with exploitation (phasic mode of LC) and constant levels of release with exploration (tonic mode of LC). Finally, a gradual transition between exploration and exploitation is also assumed in many formal models, such as the RL models and the epsilon-decreasing rule discussed above (Sutton & Barto, 1998) or the instance-based learning model of binary choice (Gonzalez & Dutt, 2011).

Nonstraightforward transitions. Finally, the issue of transition is further complicated by the fact that even if a qualitative switch between exploration and exploitation can be determined, the two behaviors do not always follow one another in a straightforward manner. For example, there are situations where agents exploit an option without any prior exploration. This can be the case if a bird needs to consume calories as soon as possible because it is on a negative energy budget (Caraco et al., 1990), or if a female lizard initially chooses indifferently between mates and thereby increases its chance for offspring (Laloi, Eizaguirre, Fédérici, & Massot, 2011). Even though rewards in those examples are accompanied by obtained information, the agents' primary goal is to exploit the rewards in order to fulfill some basic need (food, mating) and, thus, even though the choice is made randomly, it can be considered as exploitative.

Another example of a nonstraightforward transition is presented by situations where agents continue to explore after having made an exploitative decision. For example, after the initial mating, the lizards just mentioned continue searching for potentially better mates; chipmunks devote some of their foraging time to finding and assessing alternative feeding sites even after having chosen a site (Kramer & Weary, 1991); and car buyers continue to read advertisements even after they buy a particular car (Engel, 1963; see also Gigerenzer, Dieckmann, & Gaissmaier, 2012, for more on searching for further information beyond what is necessary to make a quick decision). Although this kind of transition is rarely considered in the investigation of exploration–exploitation tradeoffs, there are several reasons why agents might engage in exploration beyond exploitation: It provides information about foregone payoffs for alternatives that were not chosen and, thereby, can increase one's chance to make better exploitative choices in the future, especially in dynamically changing environments (Kramer & Weary, 1991); it can help to reduce regret (Ritov & Baron, 1995) or cognitive dissonance (Festinger, 1957); and exploration itself can be rewarding (Denrell & Le Mens, 2011).

The Role of Agents' Goals

A final important conclusion from our synthesis of the literature points to the role of an agent's goals in the exploration–exploitation process. As we will show, goals cannot only affect whether behavior is considered as exploration and exploitation, but also how agents address a possible tradeoff between these behaviors and how a behavior is interpreted by an external observer.

An important distinction in the exploration–exploitation literature is made between random and goal-oriented exploration. Several theories, such as the RL models, assume that agents choose randomly between options during exploration (also see the explorative sampler model from Erev, Ert, & Yechiam, 2008). According to this assumption, seemingly random choices will be classified as exploration, while any goal-directed behavior will be classified as exploitation. The idea of random exploration is contrasted by the idea that exploration itself can also be goal-oriented and follow a higher-level strategy (cf. Cohen et al., 2007). Higher-level goals during exploration can be related to any of the conceptual dimensions discussed above. For example, agents can have goals based on specific search strategies or heuristics, which result in characteristic *behavioral patterns* (Gigerenzer & Goldstein, 1996; Hills & Hertwig, 2010). They can have goals based on experiences with or hypotheses about the *values* of explored options (Denrell & March, 2001; Mehlhorn et al., 2014; and Rakow, Demes, & Newell, 2008, respectively). Or they can have goals related to the maximization of *obtained reward*, resulting in extended information search (Hau, Pleskac, Kiefer, & Hertwig, 2008). Due to the similarities between goal-oriented exploration and exploitation, some researchers have suggested that goal-directed behaviors during exploration may rely on mechanisms similar to those required for exploitation (Cohen et al., 2007; Gonzalez & Dutt, 2011). However, the goals of exploration and exploitation are not always identical. In general, goals during exploration tend to be more focused on obtaining information and thereby reducing uncertainty, whereas goals during exploitation tend to be largely focused on obtaining rewards (Cohen et al., 2007; Gonzalez & Dutt, 2012; Gupta et al., 2006; Hills et al., 2010; Rakow & Newell, 2010; Sutton & Barto, 1998).

Goals are also relevant for exploration–exploitation decisions because they can affect how an agent addresses possible tradeoffs. For example, goals can influence how agents trade off the immediate rewards expected from exploitation against the deferred rewards expected from exploration, and how they trade off the costs of search versus the rewards of choice as studied in research on temporal discounting (in humans: Green, Myerson, & McFadden, 1997; neural correlates: Kable & Glimcher, 2007; in animals: Kagel, Green, & Caraco, 1986) and speed–accuracy tradeoffs (in humans: Bogacz et al., 2010; in animals: Chittka et al., 2009), respectively. An especially impressive example for the interaction between goals and exploration–exploitation tradeoffs comes from the investigation of risk sensitivity in animals (Bateson, 2002; Stephens, 1981): Animals tend to be risk prone when they are on a negative energy budget (i.e., when they are using more energy than they are consuming and so will eventually starve). The goal here is to survive and choosing the risky option (i.e., exploring) that offers a chance of survival. However, if this goes well and the animal returns to a positive energy budget, the goal changes from maximizing potential energy intake to finding options that are "good enough." Consequently, the animal tends to become risk averse and exploit safe options. Here, the agents' goals affect their choice, and the outcome of the choice then affects subsequent goals and behavior.

Finally, goals can affect an observer's *interpretation* of exploration–exploitation behavior, and thereby the conclusions that are drawn about the tradeoff. For example, peoples' tendency to overmatch and thus undermaximize in probability learning experiments, where one alternative stochastically dominates the other, is typically interpreted as irrational behavior. This interpretation is based on the assumption that the decision maker's goal is to exploit the option with the higher payoff (Newell & Rakow, 2007; Vulkan, 2000). However, in the context of probability-matching experiments where participants often have to make thousands of choices in which the optimal strategy is to always press the same key (i.e., the stochastically dominant option; Shanks, Tunney, & McCarthy, 2002), participants' goals might change, for example, to the alleviation of boredom. Consequently, a participant might occasionally choose the less likely, dominated option, thereby reducing payoff but satisfying his current intrinsic goal of reducing boredom (Goodnow, 1955; Newell, Koehler, James, Rakow, & van Ravenzwaaij, 2013). Here, the agent's goals affect behavioral patterns (occasional switching to the less likely option) because they mediate how the agent values the outcomes observed.

Conclusions

In this paper, we have provided a cross-disciplinary synthesis of a large and diverse knowledge base on exploration and exploitation. The synthesis shows that the idea of a single tradeoff between exploration and exploitation represents a stark oversimplification. Instead, as suggested by Cohen et al. (2007), there seems to be a family of tradeoffs spanning different concepts and scales and potentially being addressed by different kinds of mechanisms. Based on our synthesis, we identified essential elements that a unifying theory should be able to explain in order to provide a better understanding of this family of exploration–exploitation tradeoffs.

In short, a unifying theory should cover the different conceptual dimensions of exploration and exploitation, including their various interactions, potential contradictions, and implications on the focus of a tradeoff between the two behaviors. It should be able to account for the fact that exploration and exploitation can exist on a continuum where an agent can, to some extent, engage in both exploration and exploitation at the same time. This continuum perspective is especially interesting because it implies not only that exploration and exploitation are not necessarily mutually exclusive, but also that they might even represent *mutually enabling* conditions in certain situations. Most importantly, any considerations of their conceptualizations should take into account the potential effects of agents' goals on their valuation of costs and rewards in the most general sense and thereby on their approach to the tradeoff in any given situation.

A unifying theory should also account for the variety of transitions that can occur between explorative and exploitative modes of behavior. While most existing theories and models are targeted at a specific variation of such transitions, a comprehensive model should be able to account for bidirectional transitions, and for transitions without a clear directionality or without a clear qualitative switch. Here, a potentially fruitful line of investigation could lie in what we termed nonstraightforward transitions. Such transitions might open new perspectives on the tradeoff, as they stress the importance of factors such as regret avoidance that have not traditionally been a focus of exploration–exploitation research.

Perhaps the most challenging requirement lies in the fact that exploration and exploitation are ubiquitous at many levels of abstraction, both behaviorally and cognitively. They occur at multiple scales in space and time, at various social levels, and in external and internal environments. Our models must be able to capture the interactions between these different levels (Hills et al., 2010). For example, when shopping for food, a person can either exploit a current aisle in the supermarket or they can explore a new aisle. One level up, the person can "exploit" the supermarket they know or they can explore a new one. They can exploit shopping in supermarkets, a familiar experience, or they can try shopping at different sorts of stores or markets over time. Similarly at the cognitive level, the person can exploit the heuristics they usually employ for choosing aisles, stores, or store types, or they can explore whether a new heuristic will serve them better. On a metacognitive level, the person could think about whether the way they choose a heuristic should be focused on a previously exploited one or one that still needs to be explored. Furthermore, beyond individual cognition and at the social and organizational levels, one could consider how the person's desire to explore or exploit interacts with the desire to conform to social norms. For example, a social trend to consume organic and natural products may influence the decisions to exploit or explore supermarkets.

New research is already heading in promising directions to explore how agents deal with the different (or similar) tradeoffs that present themselves across these different hierarchical levels of space, time, and abstraction (cf. Cohen et al., 2007). Future research needs to not only explore how a hierarchical use of mechanisms at different scales can be implemented, but also to what extent such mechanisms are sensitive to the variety of environmental, social, and individual factors discussed in this review. We expect that the need to consider mechanisms across different levels of abstraction will require a broader approach than those most prevalent in the current literature. The development and evaluation of such a unifying model of exploration–exploitation tradeoffs is an important goal for future research.

References

Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin, 129,* 139–167. http://dx.doi.org/10.1037/0033-2909.129.1.139

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195324259.001.0001

Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin, 125,* 591–600. http://dx.doi.org/10.1037/0033-2909.125.5.591

Armstrong, J. D., Braithwaite, V. A., & Huntingford, F. A. (1997). Spatial strategies of wild Atlantic salmon parr: Exploration and settlement in unfamiliar areas. *Journal of Animal Ecology, 66,* 203–211. http://dx.doi.org/10.2307/6022

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience, 28,* 403–450. http://dx.doi.org/10.1146/annurev.neuro.28.061604.135709

Aw, J. M., Vasconcelos, M., & Kacelnik, A. (2011). How costs affect preferences: Experiments on state dependence, hedonic state and within-trial contrast in starlings. *Animal Behaviour, 81,* 1117–1128. http://dx.doi.org/10.1016/j.anbehav.2011.02.015

Bachman, G. C. (1993). The effect of body condition on the trade-off between vigilance and foraging in Belding's ground squirrels. *Animal Behaviour, 46,* 233–244. http://dx.doi.org/10.1006/anbe.1993.1185

Bacon, I. E., Hurly, T. A., & Healy, S. (2010). Both the past and the present affect risk-sensitive decisions of foraging rufous hummingbirds. *Behavioral Ecology, 21,* 626–632. http://dx.doi.org/10.1093/beheco/arq031

Bateson, M. (2002). Recent advances in our understanding of risk-sensitive foraging preferences. *The Proceedings of the Nutrition Society, 61,* 509–516. http://dx.doi.org/10.1079/PNS2002181

Beckage, N., Todd, P. M., Penke, L., & Asendorpf, J. (2009). Testing sequential patterns in human mate choice using speed dating. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (pp. 2365–2370). Austin, TX: Cognitive Science Society.

Bell, W. J. (1991). *Searching behaviour: The behavioural ecology of finding resources*. London, UK: Chapman and Hall.

Benhamou, S. (2007). How many animals really do the Lévy walk? *Ecology, 88,* 1962–1969. http://dx.doi.org/10.1890/06-1769.1

Biesmeijer, J. C., & de Vries, H. (2001). Exploration and exploitation of food sources by social insect colonies: A revision of the scout-recruit concept. *Behavioral Ecology and Sociobiology, 49,* 89–99. http://dx.doi.org/10.1007/s002650000289

Blanco, N. J., Otto, A. R., Maddox, W. T., Beevers, C. G., & Love, B. C. (2013). The influence of depression symptoms on exploratory decision-making. *Cognition, 129,* 563–568. http://dx.doi.org/10.1016/j.cognition.2013.08.018

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends in Neurosciences, 33,* 10–16. http://dx.doi.org/10.1016/j.tins.2009.09.002

Booksmythe, I., Detto, T., & Backwell, P. R. Y. (2008). Female fiddler crabs settle for less: The travel costs of mate choice. *Animal Behaviour, 76,* 1775–1781. http://dx.doi.org/10.1016/j.anbehav.2008.07.022

Borg, Å. A., Forsgren, E., & Amundsen, T. (2006). Seasonal change in female choice for male size in the two-spotted goby. *Animal Behaviour, 72,* 763–771. http://dx.doi.org/10.1016/j.anbehav.2005.11.025

Brenner, T., & Vriend, N. J. (2006). On the behavior of proposers in ultimatum games. *Journal of Economic Behavior & Organization, 61,* 617–631. http://dx.doi.org/10.1016/j.jebo.2004.07.014

Bröder, A. (2000). Assessing the empirical validity of the "take-the-best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1332–1346. http://dx.doi.org/10.1037/0278-7393.26.5.1332

Brydges, N. M., Heathcote, R. J. P., & Braithwaite, V. A. (2008). Habitat stability and predation pressure influence learning and memory in populations of three-spined sticklebacks. *Animal Behaviour, 75,* 935–942. http://dx.doi.org/10.1016/j.anbehav.2007.08.005

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459. http://dx.doi.org/10.1037/0033-295X.100.3.432

Caraco, T. (1981). Energy budgets, risk and foraging preferences in dark-eyed juncos (*Junco hyemalis*). *Behavioral Ecology and Sociobiology, 8,* 213–217. http://dx.doi.org/10.1007/BF00299833

Caraco, T., Blanckenhorn, W. U., Gregory, G. M., Newman, J. A., Recer, G. M., & Zwicker, S. M. (1990). Risk-sensitivity: Ambient temperature affects foraging choice. *Animal Behaviour, 39,* 338–345. http://dx.doi.org/10.1016/S0003-3472(05)80879-6

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology, 9,* 129–136. http://dx.doi.org/10.1016/0040-5809(76)90040-X

Chittka, L., Skorupski, P., & Raine, N. E. (2009). Speed-accuracy tradeoffs in animal decision making. *Trends in Ecology & Evolution, 24,* 400–407. http://dx.doi.org/10.1016/j.tree.2009.02.010

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 of London Series B, Biological Sciences, 362,* 933–942. http://dx.doi.org/10.1098/rstb.2007.2098

Cook, Z., Franks, D. W., & Robinson, E. J. H. (2013). Exploration versus exploitation in polydomous ant colonies. *Journal of Theoretical Biology, 323,* 49–56. http://dx.doi.org/10.1016/j.jtbi.2013.01.022

Coolen, I., van Bergen, Y., Day, R. L., & Laland, K. N. (2003). Species difference in adaptive use of public information in sticklebacks. *Proceedings Biological Sciences, 270,* 2413–2419. http://dx.doi.org/10.1098/rspb.2003.2525

Denrell, J., & Le Mens, G. (2011). Seeking positive experiences can produce illusory correlations. *Cognition, 119,* 313–324. http://dx.doi.org/10.1016/j.cognition.2011.01.007

Denrell, J., & March, J. G. (2001). Adaptation as information restriction: The hot stove effect. *Organization Science, 12,* 523–538. http://dx.doi.org/10.1287/orsc.12.5.523.10092

Dhar, R. (1996). The effect of decision strategy on deciding to defer choice. *Journal of Behavioral Decision Making, 9,* 265–281. http://dx.doi.org/10.1002/(SICI)1099-0771(199612)9:4<265::AID-BDM231>3.0.CO;2-4

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 2,* 312–329. http://dx.doi.org/10.1016/0022-2496(65)90007-6

Engel, J. F. (1963). Are automobile purchasers dissonant consumers? *Journal of Marketing, 27,* 55–58. http://dx.doi.org/10.2307/1248373

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112,* 912–931. http://dx.doi.org/10.1037/0033-295X.112.4.912

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making, 21,* 575–597. http://dx.doi.org/10.1002/bdm.602

Fasolo, B., Hertwig, R., Huber, M., & Ludwig, M. (2009). Size, entropy, and density: What is the difference that makes the difference between small and large real-world assortments? *Psychology & Marketing, 26,* 254–279. http://dx.doi.org/10.1002/mar.20272

Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical Science, 4,* 282–289. http://dx.doi.org/10.1214/ss/1177012493

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fiedler, K., & Kareev, Y. (2006). Does decision quality (always) increase with the size of information samples? Some vicissitudes in applying the law of large numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 883–903. http://dx.doi.org/10.1037/0278-7393.32.4.883

Fiedler, K., & Kareev, Y. (2011). Clarifying the advantage of small samples: As it relates to statistical Wisdom and Cahan's (2010). normative intuitions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1039–1043. http://dx.doi.org/10.1037/a0023259

Friberg, M., Vongvanich, N., Borg-Karlson, A.-K., Kemp, D. J., Merilaita, S., & Wiklund, C. (2008). Female mate choice determines reproductive isolation between sympatric butterflies. *Behavioral Ecology and Sociobiology, 62,* 873–886. http://dx.doi.org/10.1007/s00265-007-0511-2

Fu, W.-T., & Gray, W. D. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology, 52,* 195–242. http://dx.doi.org/10.1016/j.cogpsych.2005.08.002

Fu, W.-T., & Pirolli, P. (2007). SNIF-ACT: A cognitive model of user navigation on the World Wide Web. *Human-Computer Interaction, 22,* 355–412.

Gabor, C. R., & Ryam, M. J. (2001). Geographical variation in reproductive character displacement in mate choice by male sailfin mollies. *Proceedings Biological Sciences, 268,* 1063–1070. http://dx.doi.org/10.1098/rspb.2001.1626

Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 966–982. http://dx.doi.org/10.1037/0278-7393.32.5.966

Gigerenzer, G., Dieckmann, A., & Gaissmaier, W. (2012). Efficient cognition through limited search. In P. M. Todd & G. Gigerenzer, & the ABC Research Group. (Eds.), *Ecological rationality: Intelligence in the world* (pp. 242–273). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195315448.003.0075

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62,* 451–482. http://dx.doi.org/10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669. http://dx.doi.org/10.1037/0033-295X.103.4.650

Giraldeau, L.-A., & Lefebvre, L. (1986). Exchangeable producer and scrounger roles in a captive flock of feral pigeons: A case for the skill pool effect. *Animal Behaviour, 34,* 797–803. http://dx.doi.org/10.1016/S0003-3472(86)80064-1

Giraldeau, L. A., Valone, T. J., & Templeton, J. J. (2002). Potential disadvantages of using socially acquired information. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 357,* 1559–1566. http://dx.doi.org/10.1098/rstb.2002.1065

Gittins, J. C. (1979). Bandit processes and dynamic allocation indexes. *Journal of the Royal Statistical Society Series B, Methodological, 41,* 148–177.

Godin, J.-G. J., & Keenleyside, M. H. (1984). Foraging on patchily distributed prey by a cichlid fish (*Teleostei, Cichlidae*): A test of the ideal free distribution theory. *Animal Behaviour, 32,* 120–131. http://dx.doi.org/10.1016/S0003-3472(84)80330-9

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535–574. http://dx.doi.org/10.1146/annurev.neuro.29.051605.113038

Goldstone, R. L., Ashpole, B. C., & Roberts, M. E. (2005). Knowledge of resources and competitors in human foraging. *Psychonomic Bulletin & Review, 12,* 81–87. http://dx.doi.org/10.3758/BF03196350

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118,* 523–551. http://dx.doi.org/10.1037/a0024558

Gonzalez, C., & Dutt, V. (2012). Refuting data aggregation arguments and how the instance-based learning model stands criticism: A reply to Hills and Hertwig (2012). *Psychological Review, 119,* 893–898. http://dx.doi.org/10.1037/a0029445

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27,* 591–635. http://dx.doi.org/10.1207/s15516709cog2704_2

Goodnow, J. J. (1955). Determinants of choice-distribution in two-choice situations. *The American Journal of Psychology, 68,* 106–116. http://dx.doi.org/10.2307/1418393

Green, L., Myerson, J., & McFadden, E. (1997). Rate of temporal discounting decreases with amount of reward. *Memory & Cognition, 25,* 715–723. http://dx.doi.org/10.3758/BF03211314

Green, L., Myerson, J., & Ostaszewski, P. (1999). Discounting of delayed rewards across the life span: Age differences in individual discounting functions. *Behavioural Processes, 46,* 89–96. http://dx.doi.org/10.1016/S0376-6357(99)00021-2

Gupta, A. K., Smith, K. G., & Shalley, C. E. (2006). The interplay between exploration and exploitation. *Academy of Management Journal, 49,* 693–706. http://dx.doi.org/10.5465/AMJ.2006.22083026

Harper, D. (1982). Competitive foraging in mallards: "Ideal free" ducks. *Animal Behaviour, 30,* 575–584. http://dx.doi.org/10.1016/S0003-3472(82)80071-7

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making, 21,* 493–518. http://dx.doi.org/10.1002/bdm.598

Hausmann, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment and Decision Making, 3,* 229–243.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15,* 534–539. http://dx.doi.org/10.1111/j.0956-7976.2004.00715.x

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition, 115,* 225–237. http://dx.doi.org/10.1016/j.cognition.2009.12.009

Herzog, S. M., & Hertwig, R. (2009). The wisdom of many in one mind: Improving individual judgments with dialectical bootstrapping. *Psychological Science, 20,* 231–237. http://dx.doi.org/10.1111/j.1467-9280.2009.02271.x

Hill, P. S. M., Hollis, J., & Wells, H. (2001). Foraging decisions in nectarivores: Unexpected interactions between flower constancy and energetic rewards. *Animal Behaviour, 62,* 729–737. http://dx.doi.org/10.1006/anbe.2001.1775

Hill, P. S. M., Wells, P. H., & Wells, H. (1997). Spontaneous flower constancy and learning in honey bees as a function of colour. *Animal Behaviour, 54,* 615–627. http://dx.doi.org/10.1006/anbe.1996.0467

Hills, T. T. (2006). Animal foraging and the evolution of goal-directed cognition. *Cognitive Science, 30,* 3–41. http://dx.doi.org/10.1207/s15516709cog0000_50

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience. Do our patterns of sampling foreshadow our decisions? *Psychological Science, 21,* 1787–1792. http://dx.doi.org/10.1177/0956797610387443

Hills, T. T., & Hertwig, R. (2012). Two distinct exploratory behaviors in decisions from experience: Comment on Gonzalez and Dutt (2011). *Psychological Review, 119,* 888–892. http://dx.doi.org/10.1037/a0028004

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review, 119,* 431–440. http://dx.doi.org/10.1037/a0027373

Hills, T. T., Kalff, C., & Wiener, J. M. (2013). Adaptive Lévy processes and area-restricted search in human foraging. *PLoS ONE, 8,* e60488. http://dx.doi.org/10.1371/journal.pone.0060488

Hills, T. T., Noguchi, T., & Gibbert, M. (2013). Information overload or search-amplified risk? Set size and order effects on decisions from experience. *Psychonomic Bulletin & Review, 20,* 1023–1031. http://dx.doi.org/10.3758/s13423-013-0422-3

Hills, T. T., & Pachur, T. (2012). Dynamic search and working memory in social recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 218–228. http://dx.doi.org/10.1037/a0025161

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2008). Search in external and internal spaces: Evidence for generalized cognitive search processes. *Psychological Science, 19,* 802–808. http://dx.doi.org/10.1111/j.1467-9280.2008.02160.x

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2010). The central executive as a search process: Priming exploration and exploitation across domains. *Journal of Experimental Psychology: General, 139,* 590–609. http://dx.doi.org/10.1037/a0020666

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D., & the Cognitive Search Research Group. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences, 19,* 46–54. http://dx.doi.org/10.1016/j.tics.2014.10.004

Humphries, N. E., & Sims, D. W. (2014). Optimal foraging strategies: Lévy walks balance searching and patch exploitation under a very broad range of conditions. *Journal of Theoretical Biology, 358,* 179–193. http://dx.doi.org/10.1016/j.jtbi.2014.05.032

Hutchinson, J. M. C., Wilke, A., & Todd, P. M. (2008). Patch leaving in humans: Can a generalist adapt its rules to dispersal of items across patches? *Animal Behaviour, 75,* 1331–1349. http://dx.doi.org/10.1016/j.anbehav.2007.09.006

Johnson, E. J., & Goldstein, D. (2003). Medicine. Do defaults save lives? *Science, 302,* 1338–1339. http://dx.doi.org/10.1126/science.1091721

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience, 10,* 1625–1633. http://dx.doi.org/10.1038/nn2007

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research, 4,* 237–285.

Kagel, J. H., Green, L., & Caraco, T. (1986). When foragers discount the future: Constraint or adaptation. *Animal Behaviour, 34,* 271–283. http://dx.doi.org/10.1016/0003-3472(86)90032-1

Kellam, J. S., Lucas, J. R., & Wingfield, J. C. (2006). The role of testosterone in male downy woodpeckers in winter home range use, mate interactions and female foraging behaviour. *Animal Behaviour, 71,* 695–707. http://dx.doi.org/10.1016/j.anbehav.2005.07.012

Kotler, B. P., Brown, J. S., & Bouskila, A. (2004). Apprehension and time allocation in gerbils: The effects of predatory risk and energetic state. *Ecology, 85,* 917–922. http://dx.doi.org/10.1890/03-3002

Kramer, D. L., & Weary, D. M. (1991). Exploration versus exploitation: A field study of time allocation to environmental tracking by foraging chipmunks. *Animal Behaviour, 41,* 443–449. http://dx.doi.org/10.1016/S0003-3472(05)80846-2

Krebs, J. R., Kacelnik, A., & Taylor, P. (1978). Test of optimal sampling by foraging great tits. *Nature, 275,* 27–31. http://dx.doi.org/10.1038/275027a0

Laloi, D., Eizaguirre, C., Fédérici, P., & Massot, M. (2011). Female choice for heterozygous mates changes along successive matings in a lizard. *Behavioural Processes, 88,* 149–154. http://dx.doi.org/10.1016/j.beproc.2011.08.011

Lee, M. D., & Corlett, E. Y. (2003). Sequential sampling models of human text classification. *Cognitive Science, 27,* 159–193. http://dx.doi.org/10.1207/s15516709cog2702_2

Lee, M. D., Zhang, S., Munro, M., & Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research, 12,* 164–174. http://dx.doi.org/10.1016/j.cogsys.2010.07.007

Lejarraga, T. (2010). When experience is better than description: Time delays and complexity. *Journal of Behavioral Decision Making, 23,* 100–116. http://dx.doi.org/10.1002/bdm.666

Lindstrom, K., & Lehtonen, T. K. (2013). Mate sampling and choosiness in the sand goby. *Proceedings of the Royal Society B: Biological Sciences, 280,* 20130983. http://dx.doi.org/10.1093/beheco/7.2.213

Luce, M. F. (1998). Choosing to avoid: Coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research, 24,* 409–433. http://dx.doi.org/10.1086/209518

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2,* 71–87. http://dx.doi.org/10.1287/orsc.2.1.71

Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: Is there a reduction in exploration with aging? *Frontiers in Neuroscience, 7,* 53. http://dx.doi.org/10.3389/fnins.2013.00053

Mayor, S. J., Schneider, D. C., Schaefer, J. A., & Mahoney, S. P. (2009). Habitat selection at multiple scales. *Ecoscience, 16,* 238–247. http://dx.doi.org/10.2980/16-2-3238

Mehlhorn, K., Ben-Asher, N., Dutt, V., & Gonzalez, C. (2014). Observed variability and values matter: Toward a better understanding of information search and decisions from experience. *Journal of Behavioral Decision Making, 27,* 328–339. http://dx.doi.org/10.1002/bdm.1809

Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience, 29,* 417–448. http://dx.doi.org/10.1146/annurev.neuro.29.051605.112903

Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *The Journal of Neuroscience, 32,* 2335–2343. http://dx.doi.org/10.1523/JNEUROSCI.4156-11.2012

Mutzel, A., Kempenaers, B., Laucht, S., Dingemanse, N. J., & Dale, J. (2011). Circulating testosterone levels do not affect exploration in house sparrows: Observational and experimental tests. *Animal Behaviour, 81,* 731–739. http://dx.doi.org/10.1016/j.anbehav.2011.01.001

Navarro, D. J., & Newell, B. R. (2014). Information versus reward in a changing world. In P. Bellow, M. Guarani, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1054–1059). Austin, TX: Cognitive Science Society.

Newell, B. R., Koehler, D. J., James, G., Rakow, T., & van Ravenzwaaij, D. (2013). Probability matching in risky choice: The interplay of feedback and strategy availability. *Memory & Cognition, 41,* 329–338. http://dx.doi.org/10.3758/s13421-012-0268-3

Newell, B. R., & Lee, M. D. (2011). The right tool for the job? Comparing an evidence accumulation and a naive strategy selection model of decision making. *Journal of Behavioral Decision Making, 24,* 456–481. http://dx.doi.org/10.1002/bdm.703

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin & Review, 14,* 1133–1139. http://dx.doi.org/10.3758/BF03193102

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing “one-reason” decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 53–65. http://dx.doi.org/10.1037/0278-7393.29.1.53

Nonacs, P. (2010). Patch exploitation. In M. D. Breed & J. Moore (Eds.), *Encyclopedia of animal behavior* (pp. 683–690). Oxford, UK: Academic Press.

http://dx.doi.org/10.1016/B978-0-08-045337-8.00212-6

Parker, A. M., Bruine de Bruin, W., & Fischhoff, B. (2007). Maximizers versus satisficers: Decision-making styles, competence, and outcomes. *Judgment and Decision Making, 2,* 342–350.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43,* 87–131. http://dx.doi.org/10.1146/annurev.ps.43.020192.000511

Perry, C. J., & Barron, A. B. (2013). Honey bees selectively avoid difficult choices. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 110,* 19155–19159. http://dx.doi.org/10.1073/pnas.1314571110

Phillips, N. D., Hertwig, R., Kareev, Y., & Avrahami, J. (2014). Rivals in the dark: How competition influences search in decisions under uncertainty. *Cognition, 133,* 104–119. http://dx.doi.org/10.1016/j.cognition.2014.06.006

Pirolli, P., & Card, S. K. (1999). Information foraging. *Psychological Review, 106,* 643–675. http://dx.doi.org/10.1037/0033-295X.106.4.643

Pleasants, J. M. (1989). Optimal foraging by nectarivores: A test of the marginal-value theorem. *American Naturalist, 134,* 51–71. http://dx.doi.org/10.1086/284965

Pratt, S. C., & Sumpter, D. J. T. (2006). A tunable algorithm for collective decision-making. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 103,* 15906–15910. http://dx.doi.org/10.1073/pnas.0604801103

Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes, 106,* 168–179. http://dx.doi.org/10.1016/j.obhdp.2008.02.001

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making, 23,* 1–14. http://dx.doi.org/10.1002/bdm.681

Rakow, T., Newell, B. R., & Zougkou, K. (2010). The role of working memory in information acquisition and decision making: Lessons from the binary prediction task. *The Quarterly Journal of Experimental Psychology, 63,* 1335–1360. http://dx.doi.org/10.1080/17470210903357945

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. http://dx.doi.org/10.1037/0033-295X.85.2.59

Real, L. A. (1992). Information processing and the evolutionary ecology of cognitive architecture. *American Naturalist, 140,* S108–S145. http://dx.doi.org/10.1086/285399

Ritov, I., & Baron, J. (1992). Status-quo and omission biases. *Journal of Risk and Uncertainty, 5,* 49–61. http://dx.doi.org/10.1007/BF00208786

Ritov, I., & Baron, J. (1995). Outcome knowledge, regret, and omission bias. *Organizational Behavior and Human Decision Processes, 64,* 119–127. http://dx.doi.org/10.1006/obhd.1995.1094

Riveros, A. J., & Gronenberg, W. (2010). Sensory allometry, foraging task specialization and resource exploitation in honeybees. *Behavioral Ecology and Sociobiology, 64,* 955–966. http://dx.doi.org/10.1007/s00265-010-0911-6

Rundle, H. D., & Schluter, D. (1998). Reinforcement of stickleback mate preferences: Sympatry breeds contempt. *Evolution; International Journal of Organic Evolution, 52,* 200–208. http://dx.doi.org/10.2307/2410935

Saad, G., Eba, A., & Sejean, R. (2009). Sex differences when searching for a mate: A process-tracing approach. *Journal of Behavioral Decision Making, 22,* 171–190. http://dx.doi.org/10.1002/bdm.620

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1,* 7–59. http://dx.doi.org/10.1007/BF00055564

Sang, K., Todd, P. M., & Goldstone, R. L. (2011). Learning near-optimal search in a minimal explore/exploit task. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 2800–2805). Austin, TX: Cognitive Science Society.

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research, 37,* 409–425. http://dx.doi.org/10.1086/651235

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112,* 610–628. http://dx.doi.org/10.1037/0033-295X.112.3.610

Schulze, C., van Ravenzwaaij, D., & Newell, B. R. (2015). Of matchers and maximizers: How competition shapes choice under risk and uncertainty. *Cognitive Psychology, 78,* 78–98. http://dx.doi.org/10.1016/j.cogpsych.2015.03.002

Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York, NY: HarperCollins.

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology, 83,* 1178–1197. http://dx.doi.org/10.1037/0022-3514.83.5.1178

Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of Behavioral Decision Making, 13,* 391–411. http://dx.doi.org/10.1002/1099-0771(200010/12)13:4<391::AID-BDM359>3.0.CO;2-I

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making, 15,* 233–250. http://dx.doi.org/10.1002/bdm.413

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41,* 1–20. http://dx.doi.org/10.1146/annurev.ps.41.020190.000245

Slovic, P. (1966). Risk-taking in children: Age and sex differences. *Child Development, 37,* 169–176. http://dx.doi.org/10.2307/1126437

Steiner, A. P., & Redish, A. D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nature Neuroscience, 17,* 995–1002. http://dx.doi.org/10.1038/nn.3740

Stephens, D. W. (1981). The logic of risk-sensitive foraging preferences. *Animal Behaviour, 29,* 628–629. http://dx.doi.org/10.1016/S0003-3472(81)80128-5

Stephens, D. W., & Krebs, J. R. (1987). *Foraging theory*. Princeton, NJ: Princeton University Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Teodorescu, K., & Erev, I. (2014). On the decision to explore new alternatives: The coexistence of under- and over-exploration. *Journal of Behavioral Decision Making, 27,* 109–123. http://dx.doi.org/10.1002/bdm.1785

Teodorescu, K., & Erev, I. (2014). Learned helplessness and learned prevalence: Exploring the causal relations among perceived controllability, reward prevalence, and exploration. *Psychological Science, 25,* 1861–1869. http://dx.doi.org/10.1177/0956797614543022

Todd, P. M. (2007). Coevolved cognitive mechanisms in mate search: Making decisions in a decision-shaped world. In J. P. Forgas, M. G. Haselton, & W. Von Hippe (Eds.), *Evolution and the social mind: Evolutionary psychology and social cognition* (pp. 145–159). New York, NY: Psychology Press.

Todd, P. M., Hills, T. T., & Robbins, T. (2012). *Cognitive search: Evolution, algorithms, and the brain*. Cambridge, MA: MIT Press.

Todd, P. M., & Miller, G. F. (1999). From pride and prejudice to persuasion: Satisficing in mate search. In G. Gigerenzer & P. M. Todd, & the ABC Research Group. (Eds.), *Simple heuristics that make us smart* (pp. 287–308). New York, NY: Oxford University Press.

Tuttle, E. M., Wulfson, L., & Caraco, T. (1990). Risk-aversion, relative abundance of resources and foraging preference. *Behavioral Ecology and Sociobiology, 26,* 165–171. http://dx.doi.org/10.1007/BF00172083

Tversky, A., & Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology, 71,* 680–683. http://dx.doi.org/10.1037/h0023123

Tykocinski, O. E., & Ortmann, A. (2011). The lingering effects of our past experiences: The sunk-cost fallacy and the inaction-inertia effect. *Social and Personality Psychology Compass, 5,* 653–664. http://dx.doi.org/10.1111/j.1751-9004.2011.00369.x

Valone, T. J. (1989). Group foraging, public information, and patch estimation. *Oikos, 56,* 357–363. http://dx.doi.org/10.2307/3565621

Valone, T. J. (2007). From eavesdropping on performance to copying the behavior of others: A review of public information use. *Behavioral Ecology and Sociobiology, 62,* 1–14. http://dx.doi.org/10.1007/s00265-007-0439-6

Verdolin, J. L. (2006). Meta-analysis of foraging and predation risk trade-offs in terrestrial systems. *Behavioral Ecology and Sociobiology, 60,* 457–464. http://dx.doi.org/10.1007/s00265-006-0172-6

Vickers, D. (1979). *Decision processes in visual perception*. New York, NY: Academic Press.

von Helversen, B., Wilke, A., Johnson, T., Schmid, G., & Klapp, B. (2011). Performance benefits of depression: Sequential decision making in a healthy sample and a clinically depressed sample. *Journal of Abnormal Psychology, 120,* 962–968. http://dx.doi.org/10.1037/a0023238

Vul, E., Goodman, N. D., Griffiths, T. L., & Tenenbaum, J. B. (2009). One and done? Optimal decisions from very few samples. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society* (pp. 148–153). Austin, TX: Cognitive Science Society.

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys, 14,* 101–118. http://dx.doi.org/10.1111/1467-6419.00106

Wilke, A., Hutchinson, J. M. C., Todd, P. M., & Czienskowski, U. (2009). Fishing for the right words: Decision rules for human foraging behavior in internal search tasks. *Cognitive Science, 33,* 497–529. http://dx.doi.org/10.1111/j.1551-6709.2009.01020.x

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science, 23,* 698–703. http://dx.doi.org/10.1177/0956797612443968

Wolfe, J. M. (2013). When is it time to move to the next raspberry bush? Foraging rules in human visual search. *Journal of Vision, 13,* 10. http://dx.doi.org/10.1167/13.3.10

Yechiam, E., & Busemeyer, J. R. (2006). The effect of foregone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making, 19,* 1–16. http://dx.doi.org/10.1002/bdm.509

Yechiam, E., Rakow, T., & Newell, B. R. (2015). Super-underweighting of rare events with repeated descriptive summaries. *Journal of Behavioral Decision Making, 28,* 67–75. http://dx.doi.org/10.1002/bdm.1829

Received January 31, 2014
Revision received January 30, 2015
Accepted February 2, 2015 ■