CHAPTER 13

c0013 # The Development of Goal-Directed Decision-Making

Hillary A. Raab[1], Catherine A. Hartley[1,2]

[AU1] [1]Department of Psychology, New York University; [2]Center for Neural Science, New York University

s0010 ## INTRODUCTION

p0010 In recent years, an extensive literature stemming from the study of learning in animal models has begun to elucidate the mechanisms that underpin goal–directed behavior (Dolan & Dayan, 2013). This conceptual framework distinguishes two types of "instrumental" learning, or ways in which actions can be facilitated based on their resulting rewards or punishments. A "goal-directed" action is a behavior driven by an expectation that it is likely to bring about a desired outcome (Dickinson, 1985). Goal–directed decisions leverage causal knowledge of the potential consequences of actions to flexibly pursue a current goal. Such deliberate action selection is distinguished from "habitual" behavior, in which an action is reflexively elicited by the cues or contexts associated with its prior successful performance. Whereas goal–directed actions are selected based on expectations of their consequent outcomes, habits are thought to stem from the formation of stimulus—response associations, reinforced by reward and automatically elicited by their antecedent stimuli.

p0015 This distinction between two types of instrumental action has its origins in a historic scientific debate. The dominant behaviorist account of instrumental behavior in the early 20th century proposed that instrumental action reflected an assembly of stimulus—response associations, stamped in through reinforcement, and reflexively elicited by environmental cues or contexts. In the mid–20th century Edward Tolman advocated for an alternative view, proposing instead that animals and humans form mental models of their environments (cognitive models or "maps") that can be flexibly consulted and recruited to pursue a current goal (Tolman, 1948). In the ensuing decades, a large interdisciplinary empirical literature has provided support for this conceptualization of goal-directed learning. Behavioral tasks have been designed to test for the key features of goal–directed behavior and examine the experimental factors that influence the balance between goal–directed versus habitual action (Dickinson, 1985). Convergent computational, psychological, and neuroscientific literature have begun to characterize the diverse algorithmic, cognitive, and neurobiological processes that enable goal–directed action (Dolan & Dayan, 2013; Doll, Simon, & Daw, 2012).

1

p0020   To date, the vast majority of this empirical literature has examined goal-directed behavior in adult humans and animals. In contrast, the typical development of goal-directed decision-making and the neurocognitive processes that underlie its developmental trajectory have not been widely studied. In this chapter, we will review empirical findings that illustrate marked changes in goal-directed behavior over the course of development from infancy to young adulthood. We will discuss developmental changes in the component cognitive processes and underlying neural circuits that may contribute to these shifts in behavior, highlighting the current gaps in our understanding. We begin by describing tasks designed to dissociate goal-directed behavior from habitual action. We then present findings from developmental studies employing these tasks, which reveal age-related changes in goal-directed decision-making. While we focus primarily on data exploring human development, we also integrate relevant developmental studies in animal models. We briefly outline a provisional model of the neurocircuitry implicated in goal-directed decision-making from studies in adult humans and animals and give an overview of the dynamic changes that occur within these circuits from childhood to adulthood. We then discuss what is known about the development of the component cognitive processes involved in the construction and use of the cognitive models that underpin goal-directed action, as well as the neural correlates of these developmental changes.

s0015 ASSAYS OF GOAL-DIRECTED BEHAVIOR

p0025 The goal-directedness of instrumental behavior can be assessed in experimental paradigms gauging the degree to which a decision to take an action is governed by knowledge of its likely causal consequences (i.e., action—outcome contingencies) as well as the current desirability of these expected outcomes (i.e., outcome value). Behavior that stems from the formation of habitlike stimulus—response associations is insensitive to manipulations of either of these properties, allowing for the dissociation between goal-directed and habitual behavior.

p0030   One such experimental manipulation is outcome revaluation. In these tasks, an animal learns actions that can effectively bring about the delivery of rewards. For example, an animal might learn that pressing one lever will yield a food pellet, while pressing a second lever will deliver water from a spout. In the next stage of the study, the value of one outcome is altered. This change in value may be brought about through changes in the animal's motivational state (e.g., altering the animal's degree of hunger or thirst) or through manipulations that alter the intrinsic value of the outcome (e.g., pairing a food with pharmacologically induced illness or revealing a new rewarding use for an object). The effect of this manipulation on performance of the instrumental action is then tested. Importantly, this test is performed in extinction (i.e., when no further outcomes are delivered), which ensures that instrumental behavior can only be informed by outcome

knowledge learned during initial training, as well as by any change in representation of the current value of the outcome. To the extent that action is driven by consideration of the current desirability of its likely outcome, a decrease in outcome value should result in attenuation of the instrumental response, whereas an increase in value should yield a corresponding increase in performance of the action. In contrast, habitual actions, which are thought to involve no consideration of the likely outcome, are insensitive to such changes in value. For example, an animal that continues to press a lever associated with food delivery after having eaten to satiety would reveal its actions to be habitual, whereas a cessation of lever pressing would reflect a goal-directed evaluation.

p0035   A second assay of the goal-directedness of action involves altering the learned contingency between an action and a desired outcome. A behavior is only an effective means of bringing about a desired outcome if the probability of obtaining a reinforcer when a specific action is performed is greater than the probability of reinforcement when that action is not taken. Such action–outcome contingencies can be degraded through the provision of noncontingent reward (i.e., delivering a food reinforcer in the absence of any instrumental action). If an action is habitual, the lack of consideration of action–outcome relationships will reduce sensitivity to the causal ineffectiveness of the action. Thus, a behavior that persists when its action–outcome contingency has been degraded is considered to be habitual.

p0040   Two classes of reinforcement learning algorithms have been proposed to approximate the neural computations underlying goal-directed and habitual behavior and to reproduce their key behavioral properties (Daw, Niv, & Dayan, 2005). "Model-based" algorithms select actions via a flexible but computationally and representationally intensive "tree-search" process of evaluating potential state transitions and outcomes to determine the action most likely to yield reward. In contrast, "model-free" algorithms recruit trial-and-error feedback to update a stored action value associated with a stimulus, allowing the most highly valued action to be readily elicited when the stimulus is encountered.

p0045   A sequential decision-making task (the "two-step task") was designed to dissociate these two learning processes (Daw, Gershman, Seymour, Dayan, & Dolan, 2011) and has recently been adapted to a child-friendly format (Decker, Otto, Daw, & Hartley, 2016). On each trial of the task (Fig. 13.1A), participants make a first-stage choice between two stimuli (spaceships), which is followed by a probabilistic transition to one of two second-stage states (a red or purple planet). In stage two, participants choose between two second-stage choice options (red or purple aliens), each of which is associated with a probability of yielding reward (space treasure). Reward probabilities change slowly and independently, encouraging participants to explore second-stage options throughout the task. Importantly, the probabilistic transition between first- and second-stage states (i.e., each spaceship commonly (70%) goes to one planet, and rarely (30%) to the other) creates a task structure that enables the distinction between goal-directed (model-based) and habitual (model-free) choices. Whereas a goal-directed chooser uses a cognitive
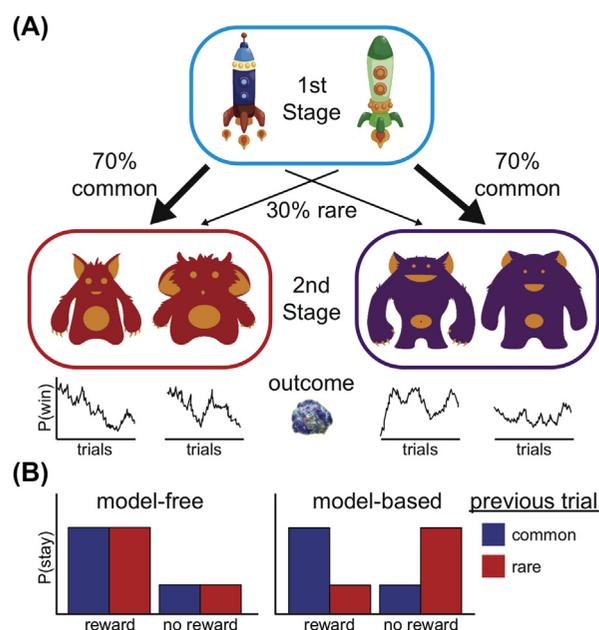
f0010 **Figure 13.1** (A) The transition structure of the child-friendly "two-step task." Participants make a first-stage choice between two stimuli (spaceships), which is followed by a probabilistic transition to one of two second-stage states (a red or purple planet). They then choose between two second-stage choice options (red or purple aliens), each of which is associated with a slowly varying probability of receiving reward (space treasure). (B) The probability of repeating a first-stage choice for an idealized model-free (left) or model-based (right) chooser is shown as a function of previous transition type (common or rare) and outcome (rewarded or unrewarded). *(From Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners tracking the developmental emergence of model-based reinforcement learning. Psychological Science, 27, 848—858.)*

model of the transitions and rewards in the task to select actions, a habitual model–free chooser, who does not recruit such a cognitive model, simply repeats previously rewarded actions. Thus, the influence of the previous trial on the first–stage choice of the next trial depends on one's learning strategy (Fig. 13.1B). A habitual, model–free learner is likely to repeat a previously rewarded first–stage choice regardless of the transition type that led to the reward (i.e., a win–stay effect; Fig. 13.1B, left). In contrast, a goal–directed chooser takes into account the state transition structure, reflected by an interactive effect of transition type (common vs. rare) and reward on "stay" decisions (Fig. 13.1B, right). For example, a habitual chooser is more likely to repeat a first–stage choice following a rare transition that led to a reward, whereas a goal–directed chooser is more likely to switch, choosing instead the spaceship that is most likely to lead to the previously rewarded planet.

p0050   Whereas tests of outcome revaluation and contingency degradation rely on a small number of trials following the task manipulation to assess goal–directedness, the

two-step task can derive estimates of the degree to which learning is model-based or model-free using all trials of the task. Moreover, this paradigm lends itself to trial-by-trial computational modeling of the distinct model-based and model-free algorithms for action evaluation. These features of the task make it particularly useful for human neuro-imaging analyses investigating the brain mechanisms underlying each learning strategy (Daw et al., 2011), as well as studies examining the effects of cognitive or affective manipulations on learning strategies (e.g., Otto, Gershman, Markman, & Daw, 2013; Otto, Raio, Chiang, Phelps, & Daw, 2013).

p0055   Each of these experimental assays of goal-directed action has been used in recent studies in adult humans to better understand the principles governing goal-directed and habitual learning and action selection. Below, we present findings from the few studies that have leveraged these tasks to begin to characterize changes in goal-directed decision-making across development.

s0020 DEVELOPMENT OF GOAL-DIRECTED INSTRUMENTAL ACTION

p0060 Recent studies have used experimental assays including outcome revaluation, contingency degradation, and sequential reinforcement learning tasks in both humans and animals at differing ages to better characterize developmental changes in the learning and reliance upon goal-directed versus habitual action.

p0065   In one study examining the sensitivity of learned actions to outcome devaluation (Kenward, Folke, Holmberg, Johansson, & Gredebäck, 2009), young children (aged 14, 19, or 24 months) learned to press a button to release an object from a box. In a second stage of the experiment, the children learned how to use an object as part of an enjoyable game (either a ball could be placed into a series of chutes, or a wooden block could be used to activate a music box). For only one of the two groups of children, the object used in the play demonstration was the same object that the child previously learned to release from the box. Thus, learning that the object could be used for play constituted an increase in its value. In a test phase, the objects used in the play demonstration were removed so that children's responses could be assessed in extinction, and children's button presses on the box from the first stage were measured as an index of goal-directed action. The group of 24-month-old children, for whom the game involved the object that they had learned to obtain through a button press, exhibited a shorter latency to press the button and a greater number of button presses compared to 24-month-old children in the control group, who had not learned any valuable use for the object they had learned to release from the box. In contrast, 14- and 19-month-old children showed no effect of the outcome revaluation on their button press behavior, suggesting that their actions were not influenced by outcome expectancies.

p0070   Another outcome revaluation study tested children ranging in age from 18 to 48 months (Klossek, Russell, & Dickinson, 2008). The children learned that by making

distinct responses on a touch screen, they could view an animated video clip from one of two cartoon series. After the responses were acquired, one video outcome was devalued by presenting children with four repeated viewings of each clip from that cartoon series. During the subsequent test phase, children were able to make responses on the screen, but no video clips were presented as a result of their actions (i.e., responses were tested in extinction). Children aged approximately 3 years and older reduced their performance of the action associated with the cartoon series that had been played repeatedly, relative to the nondevalued series, reflecting consideration of the decreased desirability of the extensively viewed cartoon series when making choices. In contrast, younger children showed no differentiation in their responses to the two cartoon series, suggesting that their actions were insensitive to this change in outcome value, and instead were driven by the previously learned reward associations.

p0075   Collectively, these studies suggest that the propensity to prospectively consider the likely outcome of an action and its current value when making choices increases with age. While in the first study we discussed (Kenward et al., 2009), sensitivity to changes in outcome value was only evident in 24-month-old children, in the second study (Klossek et al., 2008) this signature of goal-directedness only emerged in children at approximately 3 years of age. One factor that may contribute to these age differences in the emergence of goal-directed choice is task complexity. The cartoon task (Klossek et al., 2008), in which outcome value sensitivity became evident at a later age, required children to consider two potential actions that differed in value following devaluation, rather than a single action associated with a revalued outcome, as in the first study (Kenward et al., 2009). These results suggest that for tasks in which greater cognitive demand is required to bring to mind the outcomes of potential actions, goal-directed behavior may only be evident at later developmental time points.

p0080   A recent study examining sensitivity to outcome devaluation and contingency degradation in adolescent and adult rodents provides additional support for this proposal, suggesting that age-related increases in goal-directed behavior extend beyond adolescence (Naneix, Marchand, Scala, Pape, & Coutureau, 2012). Adolescent and adult rodents were trained to lever-press to obtain a food reward. Rodents in both age groups exhibited sensitivity to reward devaluation, decreasing their lever pressing after having unrestricted access to the food reward. However, when the action–outcome contingency was degraded through noncontingent delivery of the food reward, lever pressing in adult, but not adolescent, rodents decreased. When these adolescent rodents subsequently underwent the same procedure upon reaching adulthood, they exhibited sensitivity to contingency degradation at this later developmental time point. These findings suggest that there may be a more protracted developmental trajectory of goal-directed learning for behaviors that rely on greater cognitive demands. Another implication of this study is that sensitivity to outcome revaluation is dissociable from sensitivity to contingency degradation, and that outcome value sensitivity may be

evident at earlier developmental time points than sensitivity to changes in the causal efficacy of an action.

p0085   A recent study in humans corroborates this evidence of continued increases in goal-directed evaluation into adulthood (Decker et al., 2016). Children, adolescents, and adults, aged 8 to 25 years, completed a child–friendly adaptation of the "two–step task," a sequential reinforcement learning task designed by Daw et al. (2011) to disentangle goal–directed and habitual learning strategies (see Fig. 13.1 above for task details). In this task, participants can pursue reward either by simply repeating previously rewarded actions (a strategy reflected by a main effect of previous reward on first-stage choices) or by using knowledge of the transition structure of the task to select an action most likely to lead to a goal (a strategy reflected by an interaction effect of previous reward and transition type on first-stage choices).

p0090   Whereas participants across ages were equally likely to demonstrate a main effect of reward from the previous trial on first–stage choices (the behavioral signature of model–free learning), the interaction effect, indicating use of the previous trial's transition structure and outcome to pursue reward (the model–based behavioral signature), increased with age (Fig. 13.2A and B). A computational analysis was conducted to determine whether participants' trial–by–trial choices in the task were better captured by a model–based evaluation algorithm, which selects actions via a forward search through a mental model of actions and outcomes, or a model–free algorithm that recruits trial–and–error feedback to efficiently update a cached value associated with an action, but maintains no representation of the outcomes themselves. Corroborating the behavioral finding that only considers the effect of the previous trial on subsequent choices, the computational analysis that incorporates learning from the full history of task trials showed that the tendency to engage in model–based, but not model–free, computations of value increased with age.
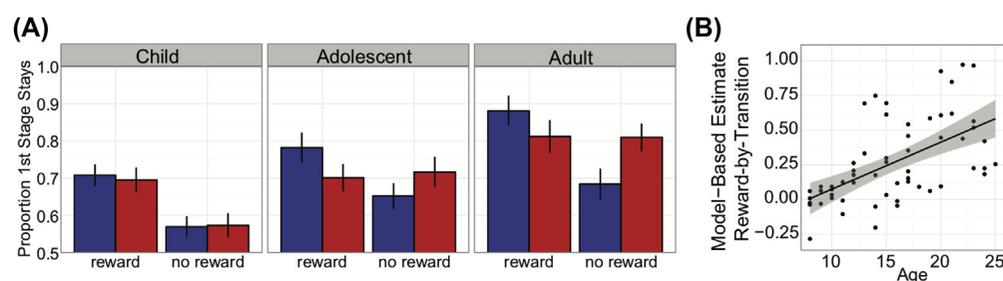
f0015 **Figure 13.2** (A) Evidence of model-free learning (significant main effect of reward) is present from childhood onward, whereas model-based learning (significant reward × transition interaction effect) is only evident in adolescents and adults. (B) Model-based behavior increases with age from childhood into adulthood ($P < .001$). *(From Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners tracking the developmental emergence of model-based reinforcement learning.* Psychological Science, 27, 848–858.)

p0095   Collectively, these studies suggest that across a diverse set of tasks, the learning of motivated behaviors in individuals at younger ages is more likely to result in habitlike behaviors than goal-directed actions that are deliberately engaged to achieve a desired outcome. Based on a large literature in adult humans and animals, instrumental learning is proposed to initially begin as goal-directed but typically become habitual over time through extensive training of the instrumental response (Graybiel, 2008; Yin & Knowlton, 2006). The developmental studies reviewed above suggest that ontogenetically, this sequential progression from goal-directed to habitual behavior may be reversed. Behavioral responses may initially be acquired through the formation of stimulus—response associations that are reinforced by rewarding outcomes. Through repetition of these actions, causal understanding of action—outcome contingencies may be acquired and ultimately used to enable the selection of actions likely to achieve one's present goals. The developmental time point at which an individual shifts toward employing a goal-directed strategy may depend on both the intrinsic complexity of the task at hand and the development of the myriad cognitive processes involved in the formation and recruitment of a mental model of that task. In the following sections, we discuss developmental changes in the neural and cognitive processes that support goal-directed decision-making.

s0025   NEUROCIRCUITRY UNDERPINNING GOAL-DIRECTED BEHAVIOR AND ITS DEVELOPMENT

p0100   A provisional model of the neural substrates of goal-directed action has emerged from a convergent body of research in adult animals and humans. Below, we outline the neural circuitry implicated in goal-directed learning in adulthood and discuss how these circuits change over the course of development from childhood to adulthood.

p0105   Goal-directed behavior involves selecting and performing an action based on the current value of its outcome. The striatum, a subcortical region of the brain, is centrally implicated in the evaluation and selection of actions (Balleine & O'Doherty, 2010). Dopaminergic input to the striatum is proposed to encode a computational reward prediction error signal, reflecting the degree to which an experienced outcome is better or worse than expected. This signal can support a feedback-driven learning process through which action values, which differ according to one's learning strategy, are estimated (McClure, Berns, & Montague, 2003; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Schultz, Dayan, & Montague, 1997). The prediction error signals that support model-free and model-based learning depend on action values and thus vary by learning strategy. Correlates of both model-free and model-based prediction errors can be observed in the ventral striatum and are associated respectively with a greater tendency toward habitual or goal-directed choice behavior (Daw et al., 2011). Distinct subregions of the dorsal striatum underpin the selection of goal-directed and habitual actions (Balleine & O'Doherty, 2010; Balleine, Delgado, & Hikosaka, 2007; Yin & Knowlton, 2006).

The caudate nucleus (dorsomedial striatum) is implicated in the learning of contingent action—outcome associations that are central to goal-directed behavior (Haruno et al., 2004; Tricomi, Delgado, & Fiez, 2004; Yin, Ostlund, Knowlton, & Balleine, 2005). The putamen (dorsolateral striatum) plays a central role in the acquisition and expression of habitual actions (Tricomi, Balleine, & O'Doherty, 2009).

p0110   The action—outcome evaluations underlying goal-directed learning additionally integrate information about states and outcomes stemming from a more extensive network of regions, including the prefrontal cortex. The prefrontal cortex exhibits strong connectivity to the striatum and lesions to corticostriatal connections disrupt the acquisition of goal-directed behavior (Hart, Bradfield, & Balleine, 2018). A number of functions critical for goal-directed action are thought to be supported by distinct areas of the prefrontal cortex. The ventromedial prefrontal cortex (vmPFC) is broadly implicated in the representation of value signals that inform motivated behavior (Dayan, Niv, Seymour, & Daw, 2006; Gläscher, Hampton, & O'Doherty, 2009; Hampton, Bossaerts, & O'Doherty, 2006; Tanaka et al., 2006; Valentin, Dickinson, & O'Doherty, 2007). The orbitofrontal cortex (OFC) has been proposed to represent associations between actions and their specific outcomes, forming a cognitive model of the task (McDannald et al., 2012; Schuck, Cai, Wilson, & Niv, 2016; Stalnaker, Cooch, & Schoenbaum, 2015; Valentin et al., 2007; Wilson, Takahashi, Schoenbaum, & Niv, 2014). Recruitment of this cognitive model enables expectations of a specific outcome to guide behavior. The OFC has been found to mediate goal-directed behavior in studies using assays such as contingency degradation and outcome revaluation (Gottfried, O'Doherty, & Dolan, 2003; Izquierdo, Suda, & Murray, 2004; McDannald et al., 2012; Valentin et al., 2007). The dorsolateral prefrontal cortex is also engaged during goal-directed learning (Smittenaar, FitzGerald, Romei, Wright, & Dolan, 2013) and may reflect the contribution of working memory and cognitive control processes, which are critical for the timely retrieval and maintenance of action—outcome associations recruited to obtain a current goal (Barch et al., 1997; Miller & Cohen, 2001). Thus, the prefrontal cortex supports diverse cognitive processes that underpin goal-directed behavior.

p0115   The hippocampus is also widely implicated in goal-directed behavior due to its central role in the learning of relationships between stimuli, events, and contexts (Shohamy & Turk-Browne, 2013; Zeithamova, Schlichting, & Preston, 2012). Learning and memory processes supported by the hippocampus are critical for the construction of the mental models of the environment that underpin goal-directed action (Pennartz, Ito, Verschure, Battaglia, & Robbins, 2011; Pfeiffer & Foster, 2013). Connectivity between the hippocampus, striatum, and vmPFC appears to support the integration of state information with knowledge about potential rewards or goal states during choice (Pennartz et al., 2011; Wimmer & Shohamy, 2012).

p0120   Over the course of development from childhood to adulthood, the neural circuitry underlying goal-directed behavior undergoes striking changes. The volume of cortical

gray matter in the brain increases in early childhood, during which there is an overproduction of synapses (Huttenlocher, 1990). This proliferative period is followed by decreases in the thickness of cortex beginning in middle childhood (approximately 8 years of age) that continue into late adolescence or early adulthood, depending on the cortical region (Huttenlocher & Dabholkar, 1997; Mills et al., 2016). Cortical thinning, thought to reflect synaptic pruning as well as other cellular changes, occurs in a topographically organized and hierarchical manner. Sensory and motor regions undergo thinning first, followed by higher-order association cortices, with anterior and lateral regions of the prefrontal cortex exhibiting continued thinning into young adulthood (Gogtay et al., 2004; Shaw et al., 2008). In contrast to the marked decline in cortical gray matter from childhood to adulthood, the nonlinear volumetric changes in gray matter volume of subcortical structures—including the striatum and the hippocampus—are less pronounced in magnitude, and highly individually variable (Raznahan et al., 2014; Wierenga, Langen, Oranje, & Durston, 2014).

p0125   While subcortical structures exhibit greater volumetric stability, patterns of white matter connectivity between cortical and subcortical structures exhibit dynamic changes over the course of development. White matter volume in the brain increases into young adulthood (Giedd et al., 1999; Mills et al., 2016). These increases are thought to reflect myelination of white matter tracts, which increases their speed of information transmission. Studies examining the functional consequences of these developmental changes suggest that connectivity between the prefrontal cortex and striatum increases from childhood to young adulthood and contributes to age-related improvements in cognitive processes that underpin goal-directed behavior (van den Bos, Cohen, Kahnt, & Crone, 2012; van Duijvenvoorde, Achterberg, Braams, Peters, & Crone, 2016; Somerville & Casey, 2010; Somerville, Hare, & Casey, 2011). Connectivity between the prefrontal cortex and subcortical structures, including the amygdala and the hippocampus, exhibits similar age-related increases (Blankenship, Redcay, Dougherty, & Riggins, 2017; Gabard-Durnam et al., 2014). Network analyses of whole brain dynamics suggest that subcortical connectivity patterns become increasingly differentiated from childhood to adulthood (Gu et al., 2015).

p0130   Developmental changes are also evident in the dopaminergic system (Wahlstrom, White, & Luciana, 2010), which is thought to modulate behavioral flexibility (Grace, Floresco, Goto, & Lodge, 2007). Dopaminergic neurons innervate many regions of the brain that have been implicated in cognitive processes that support goal-directed behavior, including the hippocampus, striatum, and prefrontal cortex (Grace et al., 2007; Roshan Cools, 2008; Shohamy & Adcock, 2010), and this neuromodulatory system undergoes dynamic changes over the course of development (Wahlstrom et al., 2010). As modulation of dopaminergic signaling is implicated in the balance between model-based and model-free choice in adults (Deserno et al., 2015; Wunderlich, Smittenaar, & Dolan, 2012), changes in dopamine neurotransmission may also play a role in differences in behavioral control across developmental time points.

p0135   While studies to date have identified substantial structural and functional develop-ment in the brain circuits implicated in goal-directed learning, few studies have directly related these changes to behavioral indices of goal-directed versus habitual action selec-tion. However, insights into the neurocognitive development of goal-directed decision-making can be gleaned from studies examining age-related changes in the cognitive processes that support goal-directed action.

p0140   In the following sections, we first discuss developmental trajectories in the learning processes that support the construction of cognitive models, including discussion of the neural mechanisms underpinning these changes wherever possible. We then turn our attention to the cognitive processes involved in the *use* of cognitive models, focusing on the central role of proactive cognitive control.

s0030 THE CONSTRUCTION OF COGNITIVE MODELS

p0145 Goal-directed actions are characterized by sensitivity to their contingent outcomes. Thus, integration of a mental representation of the structure of the environment with one's cur-rent goals is critical for the development of goal-directed decision-making. The capacity to infer causal relationships between one's actions and contingent outcomes and use this knowledge to explain or predict such outcomes is evident in children as young as 2 years (Gopnik, 2004). However, the formation of a cognitive model that can support planning in complex environments involves not only the capacity for causal inference but also a diverse array of learning and memory processes that enable the identification and repre-sentation of regularities in the environment. Events that tend to cooccur or follow in sequence can be learned through experience. Such relational associations discovered in distinct learning episodes can be assembled together, continually augmenting one's mental model with newly learned information. Knowledge of the rewarding properties of a given event, stimulus, or state can be used to prioritize that outcome as a goal. This reward information can be integrated with causal knowledge about which actions might lead to this desired outcome or to generalize reward value to states that are similar to one in which reward has been directly experienced. Below, we discuss studies that explore the development of these learning and memory processes, which provide a foundation for goal-directed action.

p0150   Through statistical learning, individuals can recognize events that tend to occur in sequence or covary with high probability. The ability to discover statistical regularities from a continuous stream of sensory experience is evident from infancy (Fiser & Aslin, 2002; Saffran, Aslin, & Newport, 1996) and has been proposed to play a central role in functions such as language acquisition. Learning of these regularities enables prediction of upcoming events, a necessary precursor to goal-directed actions that depend on those anticipated outcomes. While some studies have observed equivalent performance on sta-tistical learning tasks from childhood to adulthood (Amso & Davidow, 2012), in other

studies, performance has been found to improve with age (Potter, Bryce, & Hartley, 2017; Schlichting, Guarino, Schapiro, Turk-Browne, & Preston, 2017), suggesting that learning of more complex sequential structures may improve over development. Consistent with extensive evidence in adults that statistical learning of sequential information depends on the hippocampus and other medial temporal lobe structures (Covington, Brown-Schmidt, & Duff, 2018; Davachi & DuBrow, 2015; Preston, Shrager, Dudukovic, & Gabrieli, 2004; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014), developmental improvements in statistical learning parallel the structural development of the hippocampus (Schlichting et al., 2017).

p0155   Beyond the ability to directly extract regularities from a single learning experience, associations can also be inferred across learning episodes. The ability to associate distinct learning experiences is critical for generalizing knowledge derived from past experience to new situations. Developmental studies using paradigms testing associative inference suggest that the ability to infer relations between overlapping experiences improves with age. In one such study (Schlichting et al., 2017), participants aged 6 to 30 years completed an associative inference task in which several sets of novel object triads (e.g., A, B, and C) were presented in pairs that shared an overlapping element (e.g., pair AB and pair BC). In a subsequent test phase, participants were shown one of the items previously paired with the overlapping element (e.g., item A) and asked to select the object that was indirectly related to it through their common association. The choice set included the target object (i.e., item C), as well as two unassociated objects (e.g., items D and E). Inference improved with age, even when accuracy of recall for the direct item pairings was taken into account. This suggests that the ability to flexibly integrate the shared aspects of distinct learned relationships—a key process underlying abstract knowl-edge representation (Preston, Molitor, Pudhiyidath, & Schlichting, 2017)—improves into adulthood. Moreover, associative inference and statistical learning, assessed in the same cohort, were positively correlated, and both were associated with developmental changes in hippocampal structure (Schlichting et al., 2017). Studies in adults suggest that associative inference involves functional integration of the hippocampus and vmPFC (Zeithamova, Dominick, & Preston, 2012). Thus, developmental changes in the ability to flexibly inte-grate learned associations likely stem in part from refinement of functional connectivity between the hippocampus and the prefrontal cortex, which increases into adulthood (Menon, Boyett-Anderson, & Reiss, 2005; Ofen, Chai, Schuil, Whitfield-Gabrieli, & Gabrieli, 2012).

p0160   Other studies have similarly found that the ability to transitively integrate relational premises (e.g., "A comes before B and B comes before C. Which comes first, A or C?") also improves with age (Halford, 1984) and that the age at which children show competence at such inferences depends on the degree of relational complexity (i.e., the number of premises to be integrated) involved in the judgment (Halford, Andrews, Dalton, Boag, & Zielinski, 2002). Medial temporal lobe regions and the rostrolateral

prefrontal cortex are implicated in the learning and integration of premises during transitive inference in adults (Wendelken & Bunge, 2010). Thus, as with associative inference, developmental changes in prefrontal–hippocampal connectivity might also contribute to age-related improvements in integrating learned relations through transitive inference.

p0165   The recognition that stimuli or contexts are related can promote the generalization of associations or responses learned in one context to another. Studies of acquired equivalence (or "functional equivalence") directly assess such generalization. In these tasks, two stimuli or contexts are associated with the same outcomes or responses. Then, participants learn a novel association for one of the stimuli, and the transfer of this knowledge to the second stimulus is assessed in a subsequent generalization test phase. In one study examining acquired equivalence in 4- and 5-year-olds (Smeets, Barnes, & Roche, 1997), many children exhibited poor learning of the initial associations. However, if these associations were learned, children typically exhibited generalization in the test phase. A more recent study of participants, aged 3 to 52 years, similarly found that while the learning and retrieval of the initial pairings improved into adulthood, generalization for learned pairs was present from 6 years of age and did not differ from that of adults (Braunitzer et al., 2017). Another set of studies demonstrated that 8-month-old infants can transfer a set of learned associations to a novel stimulus based on its associative similarity, providing further evidence for the early emergence of generalization, even for more complex contextually dependent associations (Werchan, Collins, Frank, & Amso, 2015, Werchan, Collins, Frank, & Amso, 2016). Concurrent near-infrared spectroscopy in these infants revealed greater recruitment of the prefrontal cortex in infants who exhibited better learning and transfer, as well as higher eye blink rate during the task, a putative measure of striatal dopamine function (Jongkees & Colzato, 2016; Karson, 1983). This evidence that generalization ability is supported by dopaminergic innervation of frontostriatal circuitry is consistent with the selective impairment in generalization ability in adult patients for whom striatal dopamine is depleted (Myers et al., 2003).

p0170   To inform goal-directed actions, learned sequential regularities or associations must be integrated with information about currently valued outcomes. One way in which this can be accomplished is by recruiting these representations to prospectively envision potential "paths" from a current state to a goal. Studies across species have identified representations of potential future trajectories encoded in hippocampal activity (Johnson & Redish, 2007), which can be recruited to simulate both experienced and novel paths toward goals (Brown et al., 2016; Gupta, van der Meer, Touretzky, & Redish, 2010; Pfeiffer & Foster, 2013). The OFC is also proposed to play a central role in representing the sequential structure or "state space" of a current task, supporting such simulations (Schuck et al., 2016). These simulations are proposed to enable online comparison of the value of alternative courses of action, facilitating goal-directed action selection (Buckner, 2010; Pezzulo, van der Meer, Lansink, & Pennartz, 2014; van der Meer,

Kurth-Nelson, & Redish, 2012). Consistent with this proposed role for prospection in goal-directed evaluation, a neuroimaging study in adults has shown that individuals whose choices reflected a more model-based evaluation process also exhibited neural signatures of mentally invoking future outcomes when making choices (Doll, Duncan, Simon, Shohamy, & Daw, 2015). No study to date, in humans or animal models, has used neural decoding approaches to characterize the capacity to prospectively simulate future states prior to adulthood. However, developmental studies using self-report methods have found that the ability to envision or simulate future events improves with age into adulthood (Atance, 2015; Coughlin, Lyons, & Ghetti, 2014, Coughlin, Robins, & Ghetti, 2017), and the ability to recruit past learning to prospectively select a goal-directed action similarly improves as children age (Redshaw & Suddendorf, 2013; Suddendorf, Nielsen, & von Gehlen, 2011). This work suggests a potential role for age-related changes in prospective simulation ability in the development of goal-directed decision-making.

p0175   Integration of learned relations to support goal pursuit can also occur in a retrospective manner. Studies in animal models suggest that receipt of reward can promote spontaneous retrieval, or "replay," of the events that preceded or are associated with an outcome (Carr, Jadhav, & Frank, 2011). This prioritized retrieval has been proposed as a mechanism for transferring value information backward across states, allowing mentally simulated experience to "retrain" learned values for remotely associated stimuli (Shohamy & Daw, 2015). Importantly, such a mechanism might enable behavioral adaptation to changes in outcome valuation or identification of a newly rewarding action (Sutton, 1991). Empirical support for this proposal comes from neuroimaging studies in adults showing that the degree to which a stimulus, previously associated with a reward predictive cue, is invoked during reward learning predicts the preference for that stimulus in a subsequent choice phase (Kurth-Nelson, Barnes, Sejdinovic, Dolan, & Dayan, 2015; Wimmer & Shohamy, 2012). As with prospective integration, to our knowledge, no studies in children or adolescents have directly examined retrospective simulation using neuroimaging approaches, and prospective and retrospective integration are often not distinguishable in many behavioral paradigms. Thus, future studies are required to better understand whether there are developmental shifts in the mechanisms by which novel value-based associations are integrated with prior learned associations.

p0180   Mounting evidence suggests that sleep plays an important role in developmental changes in memory integration and consolidation (Fischer, Wilhelm, & Born, 2007; Wilhelm, Prehn-Kristensen, & Born, 2012). Reactivation of prioritized memories, and their associated relationships, occurs not only during learning of novel reward associations but also in "offline" periods of sleep or rest following learning (Carr et al., 2011; Kudrimoti, Barnes, & McNaughton, 1999; Wilson & McNaughton, 1994). Such offline memory integration has been proposed to underpin qualitative changes in relational memory following sleep in adults (Ellenbogen, Hu, Payne, Titone, & Walker, 2007;

Stickgold & Walker, 2013). Children not only spend more time in sleeping than adults but also spend a disproportionately larger amount of time in slow-wave sleep (Ohayon, Carskadon, Guilleminault, & Vitiello, 2004), the stage during which these reactivation events occur. While no studies have directly examined sleep-dependent changes in the transfer of reward value in children, sleep-dependent facilitation of memory for emotional, relative to neutral, information is more robust in children than adults (Prehn-Kristensen et al., 2013), suggesting that sleep might have unique effects on the transformation or integration of valenced information in younger individuals. Moreover, as nearly all the studies discussed in earlier sections that have revealed developmental changes in goal-directed behavior were conducted in a single experimental session, it is unknown whether rest- or sleep-dependent memory integration and consolidation processes might facilitate subsequent goal-directed behavior.

p0185   Collectively, the findings reviewed in this section highlight developmental changes in the learning processes that play a central role in the construction of a cognitive model of a task or environment. However, while the construction of such a model is necessary for goal-directed action, it is not sufficient. *Using* this mental model further requires the ability to recognize time points at which prior knowledge could be productively leveraged to purse a current goal, as well as the selective retrieval of goal-relevant information.

p0190   Importantly, several developmental studies have observed dissociations between the ability to acquire the structural knowledge of a task necessary for goal-directed action and actual goal-directed performance on the task itself (Decker et al., 2016; Zelazo, Frye, & Rapus, 1996). Such a dissociation was present in the two-step reinforcement learning task described earlier (Decker et al., 2016). Children's habitual (model-free) choices demonstrated that they were learning stimulus—reward associations in the second stage, but they did not make model-based choices at the first stage of the task. While children did not show evidence of *using* transition structure information to make goal-directed choices, they were able to learn the task transition structure. Children, like adolescents and adults, could explicitly report the transition structure of the task (i.e., "Which planet did this spaceship usually travel to?") and were also slower to respond following rare transitions. Slower second-stage responses following rare transitions may reflect a violation of the expectation that the more frequent transition would occur and thus, reveal knowledge of the task's probabilistic transitions. In adolescents and adults, this reaction time measure of task structure knowledge correlated with the degree to which they exhibited goal-directed choices at the first-stage of each trial. However, this was not the case in children.

p0195   What cognitive factors might account for such a failure to recruit learned action—outcome knowledge to pursue a goal? The gradual development of relational integration discussed above may have hindered children's integration of learned transition structure knowledge with learned reward associations to construct a cognitive model of the task (Potter et al., 2017). However, goal-directed choice also requires the timely retrieval

of this knowledge at the first stage of the task, when it can be leveraged proactively to support goal pursuit. Below, we discuss the developmental changes in the cognitive processes that support the ability to use learned knowledge about task structure to take goal-directed action.

s0035 USING COGNITIVE MODELS TO ENABLE GOAL-DIRECTED ACTION

p0200 To act to bring about a desired outcome, one must monitor the environment to determine when opportunities for goal-relevant actions arise. Such forward-looking behavior relies on cognitive control, or the ability to maintain and flexibly update a mental model of the task at hand, while preventing interference from irrelevant stimuli. Cognitive control is a key component of goal-directed behavior, as it allows for current goals to be retrieved and used to guide one's actions. A large body of work has identified substantial age-related changes in the recruitment of cognitive control (Diamond, 2006; Luna, 2009; Somerville & Casey, 2010), which likely play a critical role in the development of goal-directed decision-making.

p0205   A distinction is commonly made between two forms of cognitive control: proactive and reactive (Braver, 2012). In proactive control, a mental representation of a goal is invoked and sustained during a preparatory period prior to making a goal-directed action. Proactive control is typically engaged in anticipation of a cognitively demanding task, increasing resistance of goal representations to interference from goal-irrelevant stimuli. In contrast, reactive control occurs when a transient representation of a goal is evoked by the presence of a stimulus that either signals conflict with goal pursuit or directly activates a goal representation. For example, individuals might accomplish a goal of buying milk, eggs, and apples through two distinct cognitive control processes. Those who engage proactive control may leave their house with a mental list of groceries to buy, maintaining the goal representation during the drive to the grocery store and even when distracted at the store by goal-irrelevant stimuli such as bananas and bread. On the other hand, those recruiting reactive control may go to the grocery store and remember the items that they intended to buy only after entering the produce or dairy section of the store, which reactivates the goal representation. While both reactive and proactive control can support the pursuit of a goal, only proactive control allows for multistep planning to attain a goal. For example, in the two-step task described previously, for participants to use the information about the task structure from the previous trial to make a goal-directed choice at the first stage, proactive control is necessary to plan which spaceship should be chosen to get to the planet that has the greatest potential for reward. Reactive control does not support planning at the first stage to maximize reward but instead would be recruited at the second stage when it may be too late to select the most rewarding option.

p0210   A common task used to assess reactive versus proactive control is the AX-CPT, an adaptation of the continuous performance test (CPT). In the AX-CPT paradigm,

proactive attention to contextual cues can facilitate quick correct responses when a cue is highly predictive of a target probe. On each trial, a cue ("A" or "B", which may be letters, pictures, or any other type of visual stimulus) appears, followed by a short delay period, and then a probe ("X" or "Y"). While all cue and probe combinations occur, certain cue–probe pairs ("AX") appear more frequently than others ("AY", "BX", or "BY"). A target response should be made when A is followed by X, while a nontarget response should be made for all other combinations. The disproportionate number of AX trials allows individuals to develop both an expectancy for the X probe following an A cue and a prepotent response for X probes. Critically, the pattern of performance on this paradigm can be used to distinguish which cognitive control strategy was employed. Individuals using proactive control will have worse performance on the AY trials, as the target response would have been prepared during the delay period and is more likely to be emitted than for individuals relying on reactive control. In contrast, performance in individuals using reactive control will be worse on the BX trials as the preceding B cue has not been maintained, causing them to make an erroneous target response to the X cue. In contrast, those using proactive control would have already prepared their nontarget response.

p0215   In a child-friendly adaptation of the AX–CPT task, Chatham, Frank, and Munakata (2009) found that 3-year-olds primarily relied on reactive control while 8-year-olds engaged proactive control, suggesting a developmental shift from reactive to proactive control. To investigate whether younger children are capable of invoking proactive control, a cued task-switching paradigm was adapted to create three trial types: "proactive impossible," "proactive possible," and "proactive encouraged" (Chevalier, Martis, Curran, & Munakata, 2015). For "proactive impossible" trials, only reactive control could be employed, as the cue that informed the correct action was presented simultaneously with the target, rendering proactive preparation impossible. For "proactive possible" trials, the cue preceded the probe, enabling proactive control, but remained on the screen following probe onset, facilitating reactive control. For "proactive encouraged" trials, the cue disappeared from the screen when the probe appeared, requiring cue information to be maintained in memory in order to respond correctly. Although reactive or proactive control strategies could be used for both of these trial types, reactive control was disincentivized for the "proactive encouraged" trials by making reactive control more difficult. Using this manipulation, Chevalier et al. (2015) found that whereas 10-year-old children tended to engage proactive control whenever possible, 5-year-old children engaged proactive control only when reactive control was made more challenging.

p0220   Collectively, these studies suggest that younger children tend to engage reactive control, and over the course of development, shift toward preferential engagement of proactive control (Blackwell & Munakata, 2014; Chatham et al., 2009; Chevalier et al., 2015). The age at which this developmental shift occurs depends on the complexity of

the task (Chatham et al., 2009; Chevalier et al., 2015; Church, Bunge, Petersen, & Schlaggar, 2017). However, preferential engagement of reactive control early in development does not reflect an inability to engage proactive control, as younger children are able to engage proactive control if incentivized. While both cognitive control strategies can be observed early in development, the evaluation process that determines which strategy to engage may improve with age (Chevalier et al., 2015). Age-related increases in weighing the costs and benefits of each strategy according to task demands (i.e., computing the expected value of exerting each form of control; Shenhav, Botvinick, & Cohen, 2013) may support the increased reliance on proactive control (Chevalier et al., 2015). Adults may more effectively integrate value computations to inform when cognitive control should be optimally deployed (Insel, Kastman, Glenn, & Somerville, 2017).

p0225   Similar age-related trajectories are evident in working memory gating, a process that limits the information transferred into and out of working memory (Amso, Haas, McShane, & Badre, 2014; Unger, Ackerman, Chatham, Amso, & Badre, 2016). Input gating refers to the selection of task-relevant information passed into working memory, whereas output gating refers to the retrieval of information from working memory to inform action selection. A recent study suggests that children are as effective as adolescents at input gating, although they engage this strategy less often (Unger et al., 2016). Instead, children tend to rely on output gating, despite reduced efficacy in their use of this process when a subset of information must be retrieved from working memory. Selective output gating of specific items from working memory, rather than retrieving all maintained information, may impose greater cognitive demands and undergo more protracted development. Age-related improvements in output gating, along with increased reliance on input gating, may promote the deployment of working memory in a manner that better supports goal-directed behavior. Input and output gating of working memory are proposed to respectively support the engagement of proactive and reactive control. Thus, these findings corroborate the previously discussed evidence of a developmental shift toward increased reliance on proactive control.

p0230   The age-related transition toward preferential reliance on proactive control is a key component of goal-directed decision-making. In the two-step task described earlier (Decker et al., 2016), the age-related dissociation between knowledge and action may reflect developmental changes in cognitive control. Although children demonstrated verbal knowledge of the task transition structure, they were unable to translate this knowledge into goal-directed actions. This dissociation disappeared with age, potentially as individuals shifted from engaging reactive to proactive cognitive control. Development of proactive control in older participants may have supported the timely retrieval of transition structure knowledge at the first stage of the task, when it could be productively used to select the most likely path to a previously received reward. This

proposal is consistent with evidence in adults that a tendency to engage proactive control is predictive of greater model-based choice in the two-step task (Otto, Skatova, Madlon-Kay, & Daw, 2015).

p0235   The development of proactive control may also improve the ability to use explicitly communicated rules to support goal-directed behavior. Rules are consistent stimulus—response mappings that apply under specific contextual conditions (Bunge, 2004; Dayan, 2007). In certain contexts, specific actions may be consistently advantageous (e.g., stopping when a traffic light is red), obviating the need for evaluation of their consequences. In such conditions, action selection can be productively informed by rules. Rules are often learned through explicit communication. For example, a child may learn to be quiet in a classroom through explicit instruction from a teacher, or a sign stating "Use your inside voice" displayed at the entrance. Use of explicitly communicated rules requires translation of a verbally encoded policy for action into a representation of both the preconditions that must be satisfied in order for the rule to apply (i.e., "what is the stimulus or context?"), as well as the action that should be performed under those conditions (i.e., "What is the response?"). Successful deployment of rules often requires verification, prior to action, that the preconditions for its application are satisfied.

p0240   Importantly, successful deployment of a rule to accomplish a goal may engage either a habitual or a goal-directed learning process, with important implications for the flexibility of the rule-guided behavior. While adults are adept at using rules to guide behavior (Cole, Laurent, & Stocco, 2013), the ability to act on the basis of explicitly communicated rules exhibits marked developmental changes. Three-year-olds are able to use explicitly instructed rules to guide behavior (e.g., "sort cards into piles based on their shape"), but they have difficulty adjusting their behavior when the task demands change and conflict with the original rules ("now sort the cards into piles based on their color"). While 3-year-olds exhibit difficulty adhering to this new rule, they show explicit knowledge of the rule, and can repeat the rule when asked, suggesting that the challenge lies in implementing this rule. By 4 years of age, children in this task are able to accomplish this switch (Zelazo et al., 1996). However, studies employing more complex tasks to assess rule-guided behavior have observed that such switch costs decrease gradually with age and persist into young adulthood (Davidson, Amso, Anderson, & Diamond, 2006). While rule use requires translation of instruction into a behavioral procedure, individuals may differ in the extent to which successful deployment of a rule harnesses a habitual learning process. To the extent that performance of a first-learned rule in younger individuals results in the formation of a habitual stimulus—response association (as opposed to an action—outcome representation), this may undermine their capacity to flexibly alter behavior in accordance with novel rules.

p0245   Proactive control and reactive control evoke dissociable patterns of neural and physiological activity. Frontal and parietal regions are broadly implicated in the selection and

maintenance of goal-relevant information during the engagement of cognitive control (MacDonald, 2000; Miller & Cohen, 2001). Supporting the role of prefrontal development in proactive control, a longitudinal study in nonhuman primates found that the age-related enhancement of goal representation in prefrontal neurons, from adolescence to adulthood, was correlated with the ability to plan goal-directed actions (Zhou et al., 2016). Differences in the timing of activation in frontoparietal regions have been found to relate to the type of cognitive control engaged. Younger children exhibit greater activation in frontoparietal regions right before a goal is obtained, as is characteristic of reactive control. Greater recruitment of these regions during the preparatory period parallels age-related increases in proactive control (Andrews-Hanna et al., 2011; Church et al., 2017; Manzi, Nessler, Czernochowski, & Friedman, 2011). Similar temporal differences in pupil dilation, an index of cognitive effort, parallel developmental shifts from reactive to proactive control (Chatham et al., 2009; Chevalier et al., 2015). Younger children, who preferentially relied on reactive control, exhibited greater pupil dilation in response to the probe, whereas greater pupil dilation following the cue, in preparation of the probe, was evident in older children engaging proactive control. Developmental changes in frontoparietal circuitry and the temporal dynamics of pupil dilation reflect the shift from a tendency to rely on reactive control to greater engagement of proactive cognitive control across development (Bunge & Wright, 2007).

p0250   Collectively, this work suggests that developmental increases in proactive control, as well as effective deployment of reactive control, play an important role in the use of learned cognitive models or instructed task rules to inform action selection. The developmental shift from reactive to proactive control affords flexible, adaptive actions in which individuals are not simply reacting to goal-relevant cues but instead maintaining goal-relevant information in anticipation of the goal. Proactive control confers the obvious benefit of advanced preparation prior to pursuit of a goal (Braver, 2012). However, engaging in proactive cognitive control involves a narrowing of the focus of attention to task-relevant information, which may have the side effect of impeding the learning of information that is not currently goal-relevant (Thompson-Schill, Ramscar, & Chrysikou, 2009). Consistent with this notion, whereas adults instructed to make a suboptimal reward-driven choice learn action values that are biased by this inaccurate instruction, children and adolescents' learning appears to more accurately reflect their experienced reward statistics (Decker, Lourenco, Doll, & Hartley, 2015). Similarly, children instructed to attend to a specific attribute of a multidimensional cue are better than adults at learning about its uninstructed attributes (Plebanek & Sloutsky, 2017). Accordingly, reliance on reactive versus proactive cognitive control may confer distinct advantages across development, with increasing engagement of proactive control promoting goal-directed behavior.

s0040   SUMMARY AND OPEN QUESTIONS

p0255   In this chapter, we examined how goal-directed decision-making changes over the course of development, adopting a conceptual framework stemming from animal learning theory, which distinguishes goal-directed from habitual instrumental action. We presented findings in human and animal models suggesting that the choices and actions of younger individuals are less sensitive to changes in outcome value or degradation of action–outcome contingency, two key assays of goal-directedness. These age-related differences in goal-directed choice were corroborated by evidence that the recruitment of model-based computations to evaluate reward-driven actions increases with age, whereas use of model-free computations appears stable from late childhood to adulthood. We discussed how developmental changes in the cognitive processes that support the construction and recruitment of a cognitive model of a task might contribute to such developmental changes in goal-directed decision-making. Our review of this literature suggests that developmental improvements in the capacity to flexibly integrate learned associations might facilitate the construction of mental models that inform goal-directed behavior. Increases in the tendency to engage cognitive control proactively, in anticipation of goal-relevant information or choices, promotes the effective use of these cognitive "maps" to guide action. We proposed that the development of structural connectivity and functional integration between the prefrontal cortex, the striatum, and the hippocampus plays a central underlying role in these cognitive and behavioral changes.

p0260   Many gaps in our understanding were highlighted in the course of this discussion that represent fruitful avenues for future research. Developmental studies directly relating assays of goal-directed behavior to neural structure and function are needed to improve our understanding of the mechanisms underlying the development of goal-directed action. Characterizing age-related changes in functional and structural connectivity between the prefrontal cortex, hippocampus, and striatum, and their functional consequences for the construction and use of cognitive models, will help to relate developmental changes in goal-directed behavior to their neural substrates. The relationship between the dopaminergic system and the control of instrumental action across development also has yet to be well characterized. An improved understanding of how memory integration and consolidation processes change over development, whether reward plays a unique role developmentally in modulating such processes, and the role of sleep in these developmental changes will help to clarify how cognitive model formation contributes to the emergence goal-directed behavior. Investigating the neurocognitive mechanisms through which prospective representations of goals are elicited will help to elucidate how and when proactive cognitive control is recruited to support goal pursuit. Research addressing these questions would greatly augment our understanding of developmental changes in the key cognitive components of goal-directed behavior that were the focus of this chapter.

p0265   While our discussion here focused on neurocognitive developmental changes in the construction and use of cognitive models that may facilitate goal-directed action, additional processes likely influence the development of goal-directed behavior. For example, we have not addressed the question of how the arbitration between goal-directed and habitual action evaluation might shift over development. In adulthood, goal-directed and habitual evaluation processes have been proposed to operate in parallel, effectively competing for control over behavior (Balleine & O'Doherty, 2010). This competitive model is supported by studies in adult animals demonstrating that lesions to the circuitry supporting goal-directed evaluation do not eliminate instrumental behavior but instead render it habitual—insensitive to changes in outcome value or action—outcome contingency (Yin et al., 2005). Conversely, lesions to the circuitry implicated in habit learning restore the sensitivity of previously habitual actions to such changes (Yin, Knowlton, & Balleine, 2004). These findings suggest that while typically only one action evaluation system is revealed in an individual's behavior, both systems might carry out action evaluation using their respective computations and undergo some type of arbitration process to determine whether goal-directed or habitual behavior is expressed.

p0270   Alternative accounts of this arbitration process propose that the action evaluation strategy selected for behavioral expression may depend on the accuracy of its predictions (Daw et al., 2005), the costs and benefits of its computations (Kool, Gershman, & Cushman, 2017), or some combination of these factors (e.g., speed—accuracy trade-offs; Keramati, Dezfouli, & Piray, 2011). While a developmental bias toward habitual action in younger individuals could reflect difficulty in carrying out goal-directed evaluation, it might also reflect an age-related shift in such an arbitration process. For example, known developmental increases in cognitive processing speed (Kail & Salthouse, 1994; Luna, Garver, Urban, Lazar, & Sweeney, 2004) may gradually confer a competitive advantage to model-based evaluation if arbitration between these two strategies is determined via optimization of speed—accuracy trade-offs. This would be consistent with evidence that adults with higher processing speed make more model-based choices (Schad et al., 2014). The computationally intensive calculations involved in model-based evaluation (particularly for complex multistep plans) may simply take too long in children to effectively compete with the more efficient model-free computations of action value. Future studies directly probing such arbitration processes are essential for our understanding of developmental changes in the balance between these two learning systems in the control of behavior.

p0275   The distinction drawn between goal-directed and habitual action implies a dual-system perspective of behavioral control. However, there is clearly greater diversity, beyond these two "systems" in the cognitive processes that inform our motivated behavior. Through Pavlovian learning, stimuli predictive of positive or negative environmental events can acquire the capacity to elicit evolutionarily "prepared" behavioral

responses. These Pavlovian reactions (e.g., rodents freezing in anticipation of a threat) can be thought of as implementing a default behavioral response in motivated contexts. Importantly, Pavlovian behaviors can facilitate or interfere with instrumental action through a process known as Pavlovian—instrumental transfer (Estes, 1943; Lovibond, 1983). For example, instrumental actions that lead to reward can be facilitated by the action invigoration typically elicited by Pavlovian reward anticipation. However, the tendency toward the inhibition of action typically elicited by Pavlovian anticipation of threat might impair performance of the same instrumental behavior. Thus, the functioning of the Pavlovian learning system is proposed to strongly modulate the performance of goal-directed instrumental action (Dayan et al., 2006). Consistent with this proposal, adults who exhibit strong Pavlovian interference with instrumental behavior also show reduced reliance on model-based evaluation (Sebold et al., 2016). Pavlovian learning emerges early in development and exhibits dynamic developmental shifts in its expression (Hartley & Lee, 2015), suggesting that age-related changes in the interaction between Pavlovian and instrumental learning may be critical for understanding the development of goal-directed decision-making.

p0280   Goal-directed and habitual instrumental learning are distinct means of action selection that both integrate over multiple past episodes to derive action values. Recent theoretical proposals and empirical data in adults suggest that value predictions derived from single episodic memories can also drive action selection in a manner that differs from both forms of instrumental evaluation (Bornstein, Khaw, Shohamy, & Daw, 2017; Gershman & Daw, 2017; Lengyel & Dayan, 2008; Ludvig, Madan, & Spetch, 2015). This mechanism for action selection has been proposed to be particularly influential in situations where individuals have little prior experience, which may be true of more choice contexts encountered at earlier developmental stages. However, developmental changes in this behavioral control process remain largely unexplored. The neurocircuitry implicated in both Pavlovian and episodic behavioral control exhibits substantive developmental changes and overlaps with the neural circuits involved in instrumental action evaluation, suggesting the potential for dynamic interactions between these evaluation systems across development. Thus, a mechanistic account of the development of goal-directed behavior must expand beyond the narrow focus presented here to examine the diverse means by which an individual can pursue a goal.

p0285   In recent years, a large body of research has focused on understanding the computational, cognitive, and neural mechanisms underlying goal-directed action. The vast majority of this work has focused on characterizing these processes in adulthood, while the changes in these processes over development have remained largely unexplored. In this chapter, we reviewed the small number of studies that have directly examined developmental changes in goal-directed behavior. Leveraging our knowledge from studies conducted in adult humans and animals, as well as our understanding of neurocognitive

development more broadly, we discussed the mechanisms that might contribute to age-related increases in goal-directed action while highlighting the many gaps in our current understanding. Our hope is that in the coming years, increased efforts to characterize the development of goal-directed decision-making, across multiple levels of investigation, will begin to fill these gaps.

REFERENCES

Amso, D., & Davidow, J. (2012). The development of implicit learning from infancy to adulthood: Item frequencies, relations, and cognitive flexibility. *Developmental Psychobiology, 54*, 664−673.

Amso, D., Haas, S., McShane, L., & Badre, D. (2014). Working memory updating and the development of rule-guided behavior. *Cognition, 133*, 201−210.

Andrews-Hanna, J. R., Mackiewicz Seghete, K. L., Claus, E. D., Burgess, G. C., Ruzic, L., & Banich, M. T. (2011). Cognitive control in adolescence: Neural underpinnings and relation to self-report behaviors. *PLoS One, 6*, e21598.

Atance, C. M. (2015). Young children's thinking about the future. *Child Development Perspectives, 9*, 178−182.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 27*, 8161−8165.

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 35*, 48−69.

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia, 35*, 1373−1380.

Blackwell, K. A., & Munakata, Y. (2014). Costs and benefits linked to developments in cognitive control. *Developmental Science, 17*, 203−211.

Blankenship, S. L., Redcay, E., Dougherty, L. R., & Riggins, T. (2017). Development of hippocampal functional connectivity during childhood. *Human Brain Mapping, 38*, 182−201.

Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications, 8*, 15958.

Braunitzer, G., Öze, A., Eördegh, G., Pihokker, A., Rózsa, P., Kasik, L., … Nagy, A. (2017). The development of acquired equivalence from childhood to adulthood—a cross-sectional study of 265 subjects. *PLoS One, 12*, e0179525.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16*, 106−113.

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., … Wagner, A. D. (2016). Prospective representation of navigational goals in the human hippocampus. *Science, 352*, 1323−1326.

Buckner, R. L. (2010). The role of the hippocampus in prediction and imagination. *Annual Review of Psychology, 61*, 27−48.

Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affective & Behavioral Neuroscience, 4*, 564−579.

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology, 17*, 243−250.

Carr, M. F., Jadhav, S. P., & Frank, L. M. (2011). Hippocampal replay in the awake state: A potential substrate for memory consolidation and retrieval. *Nature Neuroscience, 14*, 147.

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences, 106*, 5529−5533.

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience, 27*, 1125–1136.

Church, J. A., Bunge, S. A., Petersen, S. E., & Schlaggar, B. L. (2017). Preparatory engagement of cognitive control networks increases late in childhood. *Cerebral Cortex, 27*, 2139–2153.

Cole, M. W., Laurent, P., & Stocco, A. (2013). Rapid instructed task learning: A new window into the human brain's unique capacity for flexible cognitive control. *Cognitive, Affective & Behavioral Neuroscience, 13*, 1–22.

Coughlin, C., Lyons, K. E., & Ghetti, S. (2014). Remembering the past to envision the future in middle childhood: Developmental linkages between prospection and episodic memory. *Cognitive Development, 30*, 96–110.

[**AU2**]   Coughlin, C., Robins, R. W., & Ghetti, S. (2017). Development of episodic prospection: factors underlying improvements in middle and late childhood. *Child Development* (epub ahead of print).

Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The necessity of the Hippocampus for statistical learning. *Journal of Cognitive Neuroscience*, 1–20.

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences, 19*, 92–99.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*, 2037–2078.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*, 1204–1215.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*, 1704.

Dayan, P. (2007). Bilinearity, rules, and prefrontal cortex. *Frontiers in Computational Neuroscience, 1*.

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks: The Official Journal of the International Neural Network Society, 19*, 1153–1160.

Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective & Behavioral Neuroscience, 15*, 310–320.

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners tracking the developmental emergence of model-based reinforcement learning. *Psychological Science, 27*, 848–858.

Deserno, L., Huys, Q. J. M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A. A., … Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 1595–1600.

Diamond, A. (2006). The early development of executive functions. In *Lifespan cognition: Mechanisms of change* (pp. 70–95). New York: Oxford University Press.

Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society B, 308*, 67–78.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron, 80*, 312–325.

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience, 18*, 767–772.

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology, 22*, 1075–1081.

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences, 104*, 7723–7728.

Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental Psychology, 32*, 150–155.

Fischer, S., Wilhelm, I., & Born, J. (2007). Developmental differences in sleep's role for implicit off-line learning: Comparing children with adults. *Journal of Cognitive Neuroscience, 19*, 214–227.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 458–467.

Gabard-Durnam, L. J., Flannery, J., Goff, B., Gee, D. G., Humphreys, K. L., Telzer, E., … Tottenham, N. (2014). The development of human amygdala functional connectivity at rest from 4 to 23 years: A cross-sectional study. *NeuroImage, 95*, 193–207.

Gershman, S., & Daw, N. (2017). Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology, 68*, 101–128.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., … Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2*, 861–863.

Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex, 19*, 483–495.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., … Toga, A. W. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences, 101*, 8174–8179.

Gopnik, A. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*, 3–32.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science, 301*, 1104–1107.

Grace, A. A., Floresco, S. B., Goto, Y., & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in Neurosciences, 30*, 220–227.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience, 31*, 359–387.

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., & Redish, A. D. (2010). Hippocampal replay is not a simple function of experience. *Neuron, 65*, 695–705.

Gu, S., Satterthwaite, T. D., Medaglia, J. D., Yang, M., Gur, R. E., Gur, R. C., & Bassett, D. S. (2015). Emergence of system roles in normative neurodevelopment. *Proceedings of the National Academy of Sciences, 112*, 13681–13686.

Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology, 16*, 65–93.

Halford, G. S., Andrews, G., Dalton, C., Boag, C., & Zielinski, T. (2002). Young children's performance on the balance scale: The influence of relational complexity. *Journal of Experimental Child Psychology, 81*, 417–445.

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 26*, 8360–8367.

Hart, G., Bradfield, L. A., & Balleine, B. W. (2018). Prefrontal cortico-striatal disconnection blocks the acquisition of goal-directed action. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 1311–1322.

Hartley, C. A., & Lee, F. S. (2015). Sensitive periods in affective development: Nonlinear maturation of fear learning. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 40*, 50–60.

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., … Kawato, M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: A functional magnetic resonance imaging study of a stochastic decision task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 24*, 1660–1665.

Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia, 28*, 517–527.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*, 167–178.

Insel, C., Kastman, E. K., Glenn, C. R., & Somerville, L. H. (2017). Development of corticostriatal connectivity constrains goal-directed behavior during adolescence. *Nature Communications, 8*, 1605.

Izquierdo, A., Suda, R. K., & Murray, E. A. (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 24*, 7540–7548.

Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 27*, 12176–12189.

Jongkees, B. J., & Colzato, L. S. (2016). Spontaneous eye blink rate as predictor of dopamine-related cognitive function—a review. *Neuroscience and Biobehavioral Reviews, 71*, 58–82.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86*, 199–225.

Karson, C. N. (1983). Spontaneous eye-blink rates and dopaminergic systems. *Brain: A Journal of Neurology, 106*, 643–653.

Kenward, B., Folke, S., Holmberg, J., Johansson, A., & Gredebäck, G. (2009). Goal directedness and decision making in infants. *Developmental Psychology, 45*, 809–819.

Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology, 7*.

Klossek, U. M. H., Russell, J., & Dickinson, A. (2008). The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *Journal of Experimental Psychology: General, 137*, 39–51.

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science, 28*, 1321–1333.

Kudrimoti, H. S., Barnes, C. A., & McNaughton, B. L. (1999). Reactivation of hippocampal cell assemblies: Effects of behavioral state, experience, and EEG dynamics. *The Journal of Neuroscience: The Official Journal of Society for Neuroscience, 19*, 4090–4101.

Kurth-Nelson, Z., Barnes, G., Sejdinovic, D., Dolan, R., & Dayan, P. (2015). Temporal structure in associative retrieval. *eLife, 4*, e04919.

Lengyel, M., & Dayan, P. (2008). Hippocampal contributions to control: The third way. In *Advances in neural information processing systems* (pp. 889–896).

Lovibond, P. F. (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. *Journal of Experimental Psychology. Animal Behavior Processes, 9*, 225.

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2015). Priming memories of past wins induces risk seeking. *Journal of Experimental Psychology: General, 144*, 24–29.

Luna, B. (2009). Developmental changes in cognitive control through adolescence. In P. Bauer (Ed.), *Advances in child development and behavior* (pp. 233–278).

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development, 75*, 1357–1372.

MacDonald, A. W. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*, 1835–1838.

Manzi, A., Nessler, D., Czernochowski, D., & Friedman, D. (2011). The development of anticipatory cognitive control processes in task-switching: An ERP study in children, adolescents and young adults. *Psychophysiology, 48*, 1258–1275.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron, 38*, 339–346.

McDannald, M. A., Takahashi, Y. K., Lopatina, N., Pietras, B. W., Jones, J. L., & Schoenbaum, G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *The European Journal of Neuroscience, 35*, 991–996.

Menon, V., Boyett-Anderson, J. M., & Reiss, A. L. (2005). Maturation of medial temporal lobe response and connectivity during memory encoding. *Cognitive Brain Research, 25*, 379–385.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Mills, K. L., Goddings, A.-L., Herting, M. M., Meuwese, R., Blakemore, S.-J., Crone, E. A., … Sowell, E. R. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage, 141*, 273–281.

Myers, C. E., Shohamy, D., Gluck, M. A., Grossman, S., Kluger, A., Ferris, S., … Schwartz, R. (2003). Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *Journal of Cognitive Neuroscience, 15*, 185–193.

Naneix, F., Marchand, A. R., Scala, G. D., Pape, J.-R., & Coutureau, E. (2012). Parallel maturation of goal-directed behavior and dopaminergic systems during adolescence. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32*, 16223—16232.

Ofen, N., Chai, X. J., Schuil, K. D. I., Whitfield-Gabrieli, S., & Gabrieli, J. D. E. (2012). The development of brain systems associated with successful memory retrieval of scenes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32*, 10012—10020.

Ohayon, M. M., Carskadon, M. A., Guilleminault, C., & Vitiello, M. V. (2004). Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: Developing normative sleep values across the human lifespan. *Sleep, 27*, 1255—1273.

Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science, 24*, 751—761.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences, 110*, 20941—20946.

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive control predicts use of model-based reinforcement learning. *Journal of Cognitive Neuroscience, 27*, 319—333.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38*, 329—337.

Pennartz, C. M. A., Ito, R., Verschure, P. F. M. J., Battaglia, F. P., & Robbins, T. W. (2011). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends in Neurosciences, 34*, 548—559.

Pezzulo, G., van der Meer, M. A. A., Lansink, C. S., & Pennartz, C. M. A. (2014). Internally generated sequences in learning and executing goal-directed behavior. *Trends in Cognitive Sciences, 18*, 647—657.

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place cell sequences depict future paths to remembered goals. *Nature, 497*, 74—79.

Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science, 28*, 723—732.

Potter, T. C. S., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Developmental Cognitive Neuroscience, 25*, 272—280.

Prehn-Kristensen, A., Munz, M., Molzow, I., Wilhelm, I., Wiesner, C. D., & Baving, L. (2013). Sleep promotes consolidation of emotional memory in healthy children but not in children with attention-deficit hyperactivity disorder. *PLoS One, 8*, e65098.

Preston, A. R., Molitor, R. J., Pudhiyidath, A., & Schlichting, M. L. (2017). Schemas. In H.Eichenbaum (Series Ed.) & J. H. Byrne (Vol. Ed.) (2nd ed., *Memory systems: Vol. III. Learning and memory: A comprehensive reference* (pp. 125—132). New York: Elsevier.

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus, 14*, 148—152.

Raznahan, A., Shaw, P. W., Lerch, J. P., Clasen, L. S., Greenstein, D., Berman, R., … Giedd, J. N. (2014). Longitudinal four-dimensional mapping of subcortical anatomy in human development. *Proceedings of the National Academy of Sciences, 111*, 1592—1597.

Redshaw, J., & Suddendorf, T. (2013). Foresight beyond the very next event: Four-year-olds can link past and deferred future episodes. *Frontiers in Psychology, 4*.

Roshan Cools. (2008). Role of dopamine in the motivational and cognitive control of behavior. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry, 14*, 381—395.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926—1928.

Schad, D. J., Jünger, E., Sebold, M., Garbusow, M., Bernhardt, N., Javadi, A.-H., … Rapp, M. A. (2014). Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Frontiers in Psychology, 5*.

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience, 26*, 1736—1747.

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience, 29*, 37—51.

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron, 91*, 1402–1412.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*, 1593–1599.

Sebold, M., Schad, D. J., Nebe, S., Garbusow, M., Jünger, E., Kroemer, N. B., … Rapp, M. A. (2016). Don't think, just feel the music: Individuals with strong Pavlovian-to-instrumental transfer effects rely less on model-based reinforcement learning. *Journal of Cognitive Neuroscience, 28*, 985–995.

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., … Rapoport, J. L. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 28*, 3586–3594.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron, 79*, 217–240.

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences, 14*, 464–472.

Shohamy, D., & Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences, 5*, 85–90.

Shohamy, D., & Turk-Browne, N. B. (2013). Mechanisms for widespread hippocampal involvement in cognition. *Journal of Experimental Psychology: General, 142*, 1159–1170.

Smeets, P. M., Barnes, D., & Roche, B. (1997). Functional equivalence in children: Derived stimulus–response and stimulus–stimulus relations. *Journal of Experimental Child Psychology, 66*, 1–17.

Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., & Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron, 80*, 914–919.

Somerville, L. H., & Casey, B. (2010). Developmental neurobiology of cognitive control and motivational systems. *Current Opinion in Neurobiology, 20*, 236–241.

Somerville, L. H., Hare, T., & Casey, B. (2011). Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. *Journal of Cognitive Neuroscience, 23*, 2123–2134.

Stalnaker, T. A., Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nature Neuroscience, 18*, 620–627.

Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience, 16*, 139.

Suddendorf, T., Nielsen, M., & von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science, 14*, 26–33.

Sutton, R. S. (1991). Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bulletin, 2*, 160–163.

Tanaka, S. C., Samejima, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., & Doya, K. (2006). Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. *Neural Networks: The Official Journal of the International Neural Network Society, 19*, 1233–1241.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science, 18*, 259–263.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189.

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *The European Journal of Neuroscience, 29*, 2225–2232.

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron, 41*, 281–292.

Unger, K., Ackerman, L., Chatham, C. H., Amso, D., & Badre, D. (2016). Working memory gating mechanisms explain developmental change in rule-guided behavior. *Cognition, 155*, 8–22.

Valentin, V. V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 27*, 4019–4026.

van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum–medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral Cortex, 22*, 1247–1255.

van der Meer, M., Kurth-Nelson, Z., & Redish, A. D. (2012). Information processing in decision-making systems. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry, 18*, 342–359.

van Duijvenvoorde, A. C. K., Achterberg, M., Braams, B. R., Peters, S., & Crone, E. A. (2016). Testing a dual-systems model of adolescent brain development using resting-state connectivity analyses. *NeuroImage, 124*, 409–420.

Wahlstrom, D., White, T., & Luciana, M. (2010). Neurobehavioral evidence for changes in dopamine system activity during adolescence. *Neuroscience and Biobehavioral Reviews, 34*, 631–648.

Wendelken, C., & Bunge, S. A. (2010). Transitive inference: Distinct contributions of rostrolateral prefrontal cortex and the hippocampus. *Journal of Cognitive Neuroscience, 22*, 837–847.

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2015). 8-month-old infants spontaneously learn and generalize hierarchical rules. *Psychological Science, 26*, 805–815.

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2016). Role of prefrontal cortex in learning and generalizing hierarchical rules in 8-month-old infants. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 36*, 10314–10322.

Wierenga, L. M., Langen, M., Oranje, B., & Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *NeuroImage, 87*, 120–126.

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation – what can be learnt from children? *Neuroscience and Biobehavioral Reviews, 36*, 1718–1728.

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science, 265*, 676–679.

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron, 81*, 267–279.

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science, 338*, 270–273.

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron, 75*, 418–424.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience, 7*, 464–476.

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *The European Journal of Neuroscience, 19*, 181–189.

Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *The European Journal of Neuroscience, 22*, 513–523.

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron, 75*, 168–179.

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential reasoning: Building memories to navigate future decisions. *Frontiers in Human Neuroscience, 6*.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37–63.

Zhou, X., Zhu, D., King, S. G., Lees, C. J., Bennett, A. J., Salinas, E., … Constantinidis, C. (2016). Behavioral response inhibition and maturation of goal representation in prefrontal cortex after puberty. *Proceedings of the National Academy of Sciences of the United States of America, 113*, 3353–3358.

Non-Print Items

Abstract
Throughout our lives, we face the ongoing challenge of discovering which actions are beneficial and which are not. In order to maximize reward and minimize punishment across diverse environments, individuals must learn to flexibly take actions that are likely to yield a desired outcome. This type of "goal-directed" action selection is distinguished from a "habitual" tendency to simply repeat actions that have been rewarded in the past. In this chapter, we adopt a theoretical framework stemming from animal learning theory that distinguishes goal-directed from habitual instrumental action to discuss the development of goal-directed decision-making and the neurocognitive processes that support its use. We begin by describing experimental assays of goal-directed behavior and studies employing these paradigms to examine developmental changes in goal-directed action. We then review the neural circuitry implicated in goal-directed evaluation and action selection and discuss the changes within this circuitry across development. Finally, we discuss how changes in the ability to construct and use cognitive models of one's environment contribute to developmental improvements in goal-directed decision-making.