Trajectories of Risk Learning and Real-World Risky Behaviors during Adolescence

John Mujia Wang

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Psychology

Pearl H. Chiu, Committee Chair
Brooks King-Casas
Jungmeen Kim-Spoon
Lusha Zhu

August 12, 2020
Roanoke, VA

Keywords: Reinforcement Learning, Adolescence, fMRI, longitudinal, computational modeling

Trajectories of Risk Learning and Real-World Risky Behaviors during Adolescence

John Mujia Wang

Abstract

Adolescence is a transition period during which individuals have increasing autonomy in decision-making for themselves (Casey, Jones, & Hare, 2008), often choosing among options about which they have little knowledge and experience. This process of individuation and independence is reflected as real-world risk taking behaviors (Silveri et al., 2004), including higher motor accidents, unwanted pregnancies, sexually transmitted diseases, drug addictions, and death (Casey et al., 2008). The extent to which adolescents continue to display increased behaviors with negative consequences during this period of life depends critically on their ability to explore and learn potential consequences of actions within novel environments. This learning is not limited to the value of the outcome associated with making choices, but extends to the levels of risk taken in making those choices. While the existing adolescence literature has focused on neural substrates of risk preferences, how adolescents behaviorally and neurally learn about risks remain unknown. Success or failure to learn the potential variability of these consequences, or the risks involved, in ambiguous decisions is hypothesized to be a crucial process to allow the individuals to make decisions based on their risk preferences. The alternative in which adolescents fail to learn about the risks involved in their decisions leaves the adolescent in a state of continued exploration of the ambiguity, reflected as continued risk-taking behavior.

This dissertation comprises 2 papers. The first paper is a perspective paper outlining a paradigm that risk taking behavior observed during adolescents may be a product of each adolescent's abilities to learn about risk. The second paper builds on the hypothesis of the perspective paper by first examining neural correlates of risk learning and quantifying individual risk learning abilities and then examining longitudinal risk learning developmental trajectories in relation to real-world risk-trajectories in adolescent individuals.

Trajectories of Risk Learning and Real-World Risky Behaviors during Adolescence

John Mujia Wang

General Audience Abstract

Adolescence is a transition period during which individuals have increasing autonomy in decision-making for themselves, often choosing among options about which they have little knowledge and experience. This process of individuation and independence begins with the adolescent exploring their world and those options they are ignorant of. This is reflected as real-world risk-taking behaviors, including higher motor accidents, unwanted pregnancies, sexually transmitted diseases, drug addictions, and death. We hypothesized and tested the premise that whether adolescents who succeeded or fail to learn about the negative consequences of their actions while exploring will continue to partake in behaviors with negative consequences. This learning is not limited to the value of the outcome associated with making choices, but extends to the range of possible outcomes of the choices or the risks involved. Indeed, the failure to learn the risks involved in decisions with no known information show continued and greater risk-taking behavior, perhaps remaining in a state of continued exploration of the unknown.

Acknowledgements

I am grateful to my advisors, Pearl and Brooks, for their support, encouragement, patience, understanding, and scientific insights throughout this long process. They were essential in my development as a scientist and researcher. I am also thankful for Lusha for teaching the analysis, techniques, and skills that is the foundation of this dissertation. To Jungmeen for her encouragement and discussions critical to this dissertation.

To Kevin and Dongil who are the older brothers I never had. To Vanessa, Katie, Zhuoya, Mark, and the rest of the CKC lab and bullpen where we been through it all together. And lastly to Meg who was there through it all.

Table of Contents

Introduction

Adolescence is a time period where adolescents start to partake in more real-world risky behaviors (Burnett, Bault, Coricelli, & Blakemore, 2010; Lejuez, Aklin, Zvolensky, & Pedulla, 2003; Silveri et al., 2004; Steinberg, 2008). These behaviors are known objectively to have a probability of detrimental consequences for the adolescent. Examples of such behaviors and their negative consequences include drug use leading to long term dependence, sexual behaviors leading to sexually transmitted diseases or unwanted pregnancies, the combination of excessive alcohol consumption and other activities like driving leading to injury or death of themselves or even others, and antisocial behavior leading to the harm of others (Casey et al., 2008). These are examples of some immediate consequences of risk-taking behavior, but subsequent vulnerable positions the adolescent may find themselves in could set off a chain of cascading or snowballing events that can have broader unforeseen negative consequences for the individual over the course of their development (Masten & Cicchetti, 2010). With the costly consequences from increased real-world risk-taking behavior for adolescents, it is imperative to understand the reasons behind these behaviors. Such understanding may help to inform overarching educational policies and individual strategies of intervention and prevention to curtail these real-world risk-taking behaviors.

Existing research on adolescents have found several factors that are related to their real-world risk-taking behaviors. These include, but are not limited to, emotional, inhibition, and impulsivity measurements (Blakemore & Robbins, 2012; Pharo, Sim, Graham, Gross, & Hayne, 2011; Robbins & Bryan, 2004; Whelan et al., 2014); adolescents' religiosity (Sinha, Cnaan, & Gelles, 2007); adolescents' environments (Gardner & Steinberg, 2005; Jaccard, Blanton, & Dodge, 2005; Kendler et al., 2012; Telzer, Ichien, & Qu, 2015); and adolescents' genetic and biological markers (Diamond, 2011). While these lines of research provide excellent relational analyses of constructs that are involved in adolescent real-world risk-taking behavior, we can only theorize the underlying motivations and mechanisms that drive such behaviors. Additionally, adolescence is a transitional period where individuals undergo many biological changes, including neurological changes that may directly involve those mechanisms or even explain changes in real-world risk-taking behavior.

To address questions surrounding adolescent real-world risk-taking behaviors, this dissertation will aim to i) distinguish relevant mechanisms underlying why adolescents exhibit real-world risk-taking behaviors, ii) quantify measurements of the relevant mechanisms, iii) examine longitudinal trajectories of those mechanisms, and iv) identify how longitudinal trajectories of those mechanisms relate to changes to real-world risk-taking behavior in adolescents.

Developmental theories of adolescence

During the development of a human, adolescence is described by various literature to take place from 10 to 25 years of age. Biologically, the onset of adolescence is marked by the onset of puberty and the dramatic changes in hormones (Crone & Dahl, 2012; Peper & Dahl, 2013). Subsequently, adolescents also undergo new drives and motivations as part of their developing affective processes for rewards and threats along with a later developing cognitive control process (Crone & Dahl, 2012). This combination of neural development is theorized in the dual systems model of reward and control (Casey et al., 2008; Shulman et al., 2016; Steinberg, 2008). Specifically, the affective process are localized in the limbic reward system and the ventral striatal dopamine tract while the cognitive control processes are localized to prefrontal cortex (Steinberg, 2008).

Developing earlier, the ventral striatal dopamine tract in the limbic system follows a quadratic trajectory of development peaking during adolescence (Crone & Dahl, 2012; Luciana, Wahlstrom, Porter, & Collins, 2012; Wahlstrom, White, & Luciana, 2011). Existing studies have found elevated ventral striatum activity during adolescence (Burnett et al., 2010). Elevated ventral striatum activity during this developmental timeframe have been linked to increased reward sensitivity (Galván et al., 2006; Galván & McGlennen, 2013), impulsivity (van den Bos, Rodriguez, Schweitzer, & McClure, 2015), and emotion (Hare et al., 2008). The functionality of the ventral striatal dopamine tract is involved in motivation, drives (Robinson & Berridge, 1993), rewards, and prediction error (Montague, Hyman, & Cohen, 2004; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Schultz, Dayan, & Montague, 1997). Taken together, examining the ventral striatal dopamine pathway with real-world risk-taking behavior is a prudent next step. Indeed, adolescents exhibiting problematic behavior based on a self-report survey were found to have higher striatal signals in anticipation of rewards (Bjork, Smith, Chen, & Hommer, 2011)

supporting the possibility that hyperactive ventral striatal limbic systems has an important role in real-world risk-taking behavior (Casey, 2015; Steinberg, 2008). However, other studies found no such relationship between ventral striatal activity and real-world risk-taking behavior (Braams, van Duijvenvoorde, Peper, & Crone, 2015). Additional research is needed to distinguish real-world risk-taking behavior with striatal reward activity.

Concurrently, cognitive control processes of the prefrontal cortex are hypothesized to develop at a linear trajectory and peak later than the reward limbic system (Diamond, 2002). Cognitive control processes include emotional regulation and working memory, both of which have been found to correlate positively with adolescent individuals' prefrontal cortex activity (Crone, 2009). Additionally, cognitive control has been found to regulate striatal activity and to a stronger degree as adolescents age (van den Bos, Cohen, Kahnt, & Crone, 2012; van den Bos et al., 2015). The anatomical connectivity between the prefrontal cortex and striatum has also been found to increase with age during adolescence (van den Bos et al., 2012, 2015). These findings suggest improved executive functions may inhibit striatal reward and motivation signal, subsequently leading to decreased real-world risk-taking behavior. Consistent with this suggestion, higher impulse control and executive functioning scores from neuropsychological tests were found to predict lower real-world risky-behavior among adolescents (Pharo et al., 2011).

Existing literature has found evidence that the dual theory components of ventral striatal reward activity and cognitive control processes of the prefrontal cortex are important networks related to adolescents exhibiting real-world risk-taking behavior. However, these results do not address the underlying mechanisms involved in decision-making, and the question of how these components lead adolescents to partake in real-world risk-taking behaviors remains unclear. With the combination of computational modeling and functional magnetic resonance imaging, it may be possible to identify these underlying mechanisms.

Computational approach to learning

The power of computational modeling is the ability to identify and quantify cognitive mechanisms that may not be measurable based on theory alone. Specifically, it may allow us to hypothesize the processes adolescents may take that may lead them to partake in real-world risky behaviors.

In the real-world, when an individual arrives at a point where they must make a decision, they often have a number of options with a range of possible outcomes to choose from. In the lab environment, a 'multi-armed bandit' task is used to represent this situation, in which the individual is shown multiple slot machines they can choose from (Sutton & Barto, 1998). In the initial stage, the value associated with each possible choice is unknown. That is, the possible outcomes and probabilities of the outcomes associated with each of the choices are unknown; this lack of information makes each of the choices ambiguous (Preuschoff & Bossaerts, 2007). Each individual has to explore the outcome of the choices by making them. When they encounter the outcome from their choice, they may learn from that outcome, which will subsequently update their expected outcome (i.e., expected value) for that choice and be applied to subsequent choices. From there, the individual may have an expected value for some of the choices while others remain unknown. The individual would then have to decide between exploring the remaining unknown option(s) or continuing to choose the option (i.e., exploiting) for which they have an expected value (Sutton & Barto, 1998).

A key component of basic decision-making processes is how the individual updates their expected values from the received outcome, in other words, how they learn. The Temporal Difference Model of Reinforcement Learning (TDRL; Montague, Dayan, & Sejnowski, 1996; Montague et al., 2004) is a classic approach in the existing literature used to quantify the learning process. The main focus of TDRL is that the expected value of a specific choice is updated by the prediction error from selecting that choice each time it has been made. A prediction error is the difference between the realized outcome for that choice and what the expected value is for that choice. Additionally, the magnitude of how much the prediction error will update subsequent expected values of the associated choice is determined by an estimated parameter defined as a learning rate. The learning rate is a quantifiable construct indicating that an individual is learning and updating their valuations. Individuals can then use the expected values, once they have learned, to maximize their outcomes for future choices. Using fMRI, a BOLD signal correlated with prediction errors while the individual views outcomes has been robustly identified in the striatum of adult individuals (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Pessiglione et al., 2006; Rolls, McCabe, & Redoute, 2008; Schultz et al., 1997). Adolescents' prediction error signals correlated negatively with the quadratic function of age (Cohen et al., 2010), and the pattern matches the limbic developmental trajectory in adolescents (Crone &

Dahl, 2012), possibly suggesting faster learning as a function of their striatal developmental trajectory. However as mentioned above, inconsistent striatal signals in reward learning tasks (van den Bos et al., 2012) require additional research (Gee et al., 2018).

As reviewed above, when a person has to make a decision with multiple unknown choices, the expected values associated with each choice are unknown. In the previous reward valuation learning model, the subsequent expected values were learned based on the prediction errors from experience outcomes. Another way to look at the initial decision is that each of the possible choices are inherently ambiguous, as the probabilities of the outcomes are unknown (Epstein, 1999; Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005; Huettel, Stowe, Gordon, Warner, & Platt, 2006). All decisions made by the individual at this point are ambiguous and for the purpose of exploring. For an adolescent, all choices are ambiguous from their perspective while there may be a probability of an objectively negative consequences. Such probabilities of negative outcomes remain unknown to the adolescent unless they experience the outcome and learn from it. So, the decision to explore ambiguous options to the adolescent is necessary and may appear as real-world risk-taking behavior to outside observers. Exploration of these ambiguous options seen as real-world risk-taking behavior can be seen as not maladaptive (Romer, Reyna, & Satterthwaite, 2017). Furthermore, adolescent individuals show elevated novelty-seeking behavior (Burnett et al., 2010) aiding their exploration of the world. Using the Iowa Gambling Task where individuals made decisions regarding several decks of cards with ambiguous outcomes (Bechara, Damasio, & Damasio, 2000), adolescents showed increased exploratory behavior than adults (Christakou et al., 2013).

As the individual realizes outcomes over time, they may update their valuations – learn – in the form of a distribution of possible outcomes rather than a specific expected value. A known expected distribution is referred to as 'Risk' in the behavioral economics literature (D'Acremont, Lu, Li, Van der Linden, & Bechara, 2009). This allows for informed decisions based on the risk associated with those possible choices. In reward learning, a reward prediction error is utilized to update the expected value of the associated choice and has been neurally associated with striatal signals. Similarly, others have hypothesized that there may be a risk prediction error that helps in the updating of expected risk (Preuschoff & Bossaerts, 2007; Preuschoff, Quartz, & Bossaerts, 2008). As with reward prediction error, the risk prediction error is the difference between the current trial's updated variance and the expected variance associated with the choice. Also

similar to reward learning, the rate an individual can update his or her expected variance with their risk prediction error will more quickly decrease the ambiguity of their choices. Subsequently, that individual would have an estimate of risk associated with the choice and decide based on those risks.

Alternatively, if adolescents fail to learn about the risks, it can lead them to continued risk-taking behavior. Specifically, adolescents who do not learn the risk associated with their choices remain in a constant state of exploration as the probability of outcomes associated with the choices remain ambiguous, even for the choices with negative consequences. So continued exploration means the adolescent would display continuous real-world risk-taking behavior. Inability to learn the risks would change an adaptive behavior of exploration into a maladaptive behavior of continued risk-taking behavior.

When attempting to isolate a risk prediction error in adults using a gambling task by keeping track of changes in probabilities, the anterior insula signal was found to reflect the risk prediction error (Preuschoff et al., 2008). Additional research using the Iowa Gambling Task (Bechara et al., 2000) found risk prediction error to correlate positively with activity in the insula and inferior frontal gyrus, replicating previous results. Activity in the inferior frontal gyrus was also positively correlated with risk averse participants in a second level analysis (D'Acremont et al., 2009). Furthermore, the dorsal ACC and anterior insula were found to correlate with the "volatility" of a choice (Behrens, Woolrich, Walton, & Rushworth, 2007), an estimated parameter representing the reliability of previous information and analogous to risk prediction error. However, research into risk prediction errors using model-based analyses that track trial-by-trial variance has been sparse, including only theory papers (Preuschoff & Bossaerts, 2007) and limited neuroimaging. The extant research suggests that the anterior insula and dorsal medial prefrontal cortex play an important role in risk learning in the form of risk prediction errors (D'Acremont et al., 2009; Preuschoff et al., 2008). There are currently no direct risk-learning studies of adolescents, even though adolescence is described as a period of increased risk-taking behavior. Taken together, a hypothesis of the underlying mechanism involved in adolescent real-world behavior can be formed, that is the slower an adolescent learns about the risks associated with their choices, the more they need to continue exploring ambiguous choices and take additional real-world risk-taking behavior in order learn.

Overview

The present work intends to investigate how computational models of reinforcement learning can distinguish the underlying mechanisms in adolescents that lead to increased real-world risk-taking behavior. Paper 1 expands these ideas suggesting that risk-taking behavior observed during adolescents may be a product of each adolescent's abilities to learn about risk. The paper 2 builds on the hypothesis by first examining neural correlates of risk learning and quantifying individual risk learning abilities and then examining longitudinal risk learning developmental trajectories in relation to real-world risk-trajectories in adolescent individuals. Specifically, the study uses a computational model to quantify a parameter that captures the mechanism of risk learning in adolescents. Secondly, the study will examine the developmental trajectories of risk learning during adolescents. Thirdly, the study will compare developmental trajectories of risk learning during adolescents with their corresponding trajectories of real-world risk-taking behavior.

Paper 1

Title: Importance of Risk Learning in Minimizing Recurrent Real-World Risky Behavior

Abstract:

Adolescence is a transition period during which individuals have greater autonomy in decision-making for themselves (Casey, Jones, & Hare, 2008), often choosing among options about which they have little knowledge and experience. This process of individuation and independence is reflected as real-world risk taking behaviors (Silveri et al., 2004), including higher motor accidents, unwanted pregnancies, sexually transmitted diseases, drug addictions, and death (Casey et al., 2008). The extent to which adolescents continue to display increased behaviors with negative consequences during this period of life depends critically on their ability to explore and learn potential consequences of actions within novel environments. This learning is not limited to the value of the outcome associated with choices per se, but extends to learning the risks in those choices. Risk, here, is equivalent to the mathematical variance of possible outcomes associated with a choice. Exploring, likewise, is making a decision in the state of ambiguity, where decisions are made without knowing the value or risk of the outcome. While the existing adolescence literature has focused on neural substrates of risk or ambiguity preferences, how adolescents behaviorally and neurally learn about those risks remain unknown. Success or failure to learn the value and risk from an ambiguous environment is a crucial process that individuals must before they can make decision based on their preferences. The alternative in which adolescents fail to learn about risks involved in their decisions leaves adolescents in a state of continued exploration of options which remain ambiguous, reflected as continued risk-taking behavior. This review will focus on the neural and behavioral substrates of risk learning, which may provide a key in understanding escalated risk-taking behaviors shown by adolescents.

Introduction

Adolescence is a period in development during which individuals display increased risk-taking behaviors (Silveri et al., 2004). Concurrently, adolescents experience increased rates of negative consequences during this period of life, including higher motor accidents, unwanted pregnancies, sexually transmitted diseases, drug addiction, and death (Casey et al., 2008). Increased rates of risk-taking behavior exhibited by adolescents have been linked to numerous factors including, but not limited to emotional, inhibition, and impulsivity measurements (Blakemore & Robbins, 2012; Pharo, Sim, Graham, Gross, & Hayne, 2011; Robbins & Bryan, 2004; Whelan et al., 2014); adolescents' religiosity (Sinha, Cnaan, & Gelles, 2007); adolescents' environments (Gardner & Steinberg, 2005; Jaccard, Blanton, & Dodge, 2005; Kendler et al., 2012; Telzer, Ichien, & Qu, 2015); and adolescents' genetic and biological markers (Diamond, 2011). The general direction of these previous studies has been to identify the factors that lead adolescents to make highly risky choices, which can have dire consequences in their lives. The hope of these research topics is to possibly inform policy, therapy, or interventions that can decrease adolescents' risky behaviors. While the current research into adolescents' risk-taking does an excellent job describing the circumstances in which adolescents are more or less likely to partake in risky behavior in the real world, further research is needed to identify the underlying motivations and mechanisms that leads adolescents to actively choose risky behaviors. Some of these choices may not necessarily be maladaptive, while some may continuously lead adolescent to risky behaviors.

The aim of the current paper is to delve into possible motivations and mechanisms underlying why adolescents might continuously partake in risky behaviors. During the course of this paper, the following topics will be discussed in the context of adolescent development: i) The benefits of exploration during adolescence and the necessary negative consequences that accompanies exploration; ii) learning as a crucial complement to exploration and existing literature into adolescents' learning processes; iii) linking risk and ambiguity preference to risk learning; and iv) relating dual systems model of reward and executive functioning during development to findings in research about learning.

Exploration and its necessary negative consequences

Exploration is a process by which an individual gathers information by choosing the unknown and experiences new outcomes, which may help the individual gain knowledge that guides future choices. Without exploration, individuals may never comprehend the value of the choices available to them and never discover choices that may be better as they would fail to collect any new information. During the time of development including adolescence, exploration is more imperative as individuals are forming their world views and the basis of their decision making (Casey et al., 2008) while possessing minimal knowledge of the world. In fact adolescents will utilize experience gained through exploration over explicit given instructions showing the priority they place in personal experience and exploration (Decker, Lourenco, Doll, & Hartley, 2015).

Exploratory behavior during adolescence has been found to cultivate greater benefits later in life such that experimenting with drugs, but not continued use, during adolescence leads to more positive social relationships than those who completely abstained (Oliva, Keyes, Iacono, & McGue, 2012). This idea is further supported by educational policy and childhood development literature showing that exploration through unstructured play in the form of school recess is beneficial for children socially, emotionally, and cognitively (Ramstetter, Murray, & Garner, 2010). Furthermore in the social domain, willingness to explore the unknown, in the form of ambiguity tolerance, has been found to predict prosocial behaviors in that individuals more likely to cooperate with and trust (Vives & FeldmanHall, 2018), which may lead to social reciprocity and social fulfillment.

With countless number of choices with varying outcomes a person will face during their life time, exploration becomes a necessity for collecting information about the unknown. Unfortunately, exploring the outcomes of unknown choices have the inherent potential of turning out for the worse, and the realization of the negative outcomes only takes place once the option is explored. So with exploration, negative outcomes become necessary experiences. From a third person objective perspective, it might appear that an individual exhibits negative risky behavior. However to the contrary, these real-world risky taking behavior from exploration motivated by sensation seeking is not maladaptive and emphasized as part of experience accumulation needed to assume adult roles and behaviors including making wise decisions (Romer, Reyna, & Satterthwaite, 2017).

Altogether, exploration is a necessary process for adolescents to gain information about their environments that allows them to better survive for the future. However, exploration may exhibit as real-world risk-taking behavior from an objective standpoint because of the negative possible outcomes that will inherently occur. These real-world risky behaviors are necessary and not maladaptive.

Learning as a crucial complement to exploration and current research examining learning during adolescence

Thus far this paper has focused on the importance and necessity of exploration for adolescence to gain information. Learning is the process of integrating the information into knowledge that the individual can use in future decisions (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Dayan & Balleine, 2002; Sutton & Barto, 1998). From a reinforcement learning framework, learning is the process of mapping the outcomes or the resulting state of the world to the actions or decisions made by an individual. If the individual fails to learn, then no new knowledge is gained and the individual would be in the same spot when faced with the same choices. This could result in a continued state of making the same decisions, and a continued state of exploration in the case of no knowledge of the choices. Negative outcomes from repeated objectively real-world risky behaviors from a state of continuous state of exploration is now maladaptive. So, the learning ability of an individual provides a line to differentiate maladaptive and not maladaptive risky behaviors.

A classic approach to study learning is in the form of the Temporal Difference Model of Reinforcement Learning (TDRL; Montague, Dayan, & Sejnowski, 1996; Montague, Hyman, & Cohen, 2004). TDRL is a computation that posits individuals track expected value of the choices available to them and update those valuations by the prediction errors from their choice. Prediction errors are the difference between the realized outcome and the original expected value for that choice. Functional neuroimaging studies have isolated the striatum to track prediction errors in adult individuals (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Rolls, McCabe, & Redoute, 2008; Schultz, Dayan, & Montague, 1997). Learning is quantified as the estimate parameter of learning rate to which individuals update their valuations with prediction errors. Quick learners have higher learning rates, because they use their prediction error signals to a greater extent in updating their future

expected values. Those expected values are found to localize to the Ventral Medial Prefrontal Cortex in adults (VMPFC; Rolls et al., 2008), an area shown to encode valuations across modalities (Chib, Rangel, Shimojo, & O'Doherty, 2009; D. J. Levy & Glimcher, 2011; Lin, Adolphs, & Rangel, 2012; Smith et al., 2010).

Previous studies aiming to differentiate adolescent and adults' neural learning signals failed to provide a clear answer. Several studies reported that they did not find striatal prediction error differences between adults and adolescents (Davidow, Foerde, Galván, & Shohamy, 2016; Hauser, Iannaccone, Walitza, Brandeis, & Brem, 2015; A. H. Javadi, Schmidt, & Smolka, 2014; A. Javadi, Schmidt, & Smolka, 2014; van den Bos, Cohen, Kahnt, & Crone, 2012). However, it has been posited that the learning paradigms in these studies may be too simple to elicit learning differences (Davidow, Insel, & Somerville, 2018). Despite the lack of striatal prediction error signal differences between groups, neural prediction error signals differences in other brain regions (e.g., adolescents with greater Hippocampal prediction error, Davidow et al., 2018; adolescents showing greater Anterior Insula signal to negative prediction errors, Hauser et al., 2015) and lower medial prefrontal cortex connectivity with the striatum in adolescents (van den Bos, Cohen, et al., 2012) were found. Meanwhile, estimated learning rate comparisons between adults and adolescents provided mixed results as well, with adolescents showing lower learning rate estimates (Davidow et al., 2016), higher learning rate for worse than expected outcomes (Hauser et al., 2015), and lower learning rates for better than expected outcomes (van den Bos, Cohen, et al., 2012) than adults. In another study, higher learning rates to worse than expected outcomes showed decreased striatal and medial prefrontal cortex signal for adolescents, but not adults (Christakou et al., 2013). Altogether, the current research does not provide a clear answer as to adolescents' learning abilities and, crucially, has not linked value-based reinforcement learning to real-world risk-taking behavior. A possible explanation may be that the current approach to study learning in adolescents have narrowly focused on value-based learning and not about learning the possible range of outcomes. What makes real-world risk-taking behavior risky is not the average anticipated outcome but outcomes at the end of the spectrum of the possible outcome range that are negative and detrimental to the individual.

Another way to approach initial decisions during exploration is that each of the possible choices have an inherently high uncertainty, as the possibilities of the outcomes are unknown – referred to as ambiguous choices in the economics literature (Epstein, 1999; Hsu, Bhatt,

Adolphs, Tranel, & Camerer, 2005; Huettel, Stowe, Gordon, Warner, & Platt, 2006). Exploration in this framework is choosing the options that are high in ambiguity. Continuing through the explorative process, the adolescent will observe outcomes that can inform not only the expected value of the choice, but also the expected distribution of possible outcomes, which the behavioral economic literature refers to as risk (D'Acremont, Lu, Li, Van der Linden, & Bechara, 2009). If the distribution associated with the choice is large, then the risk associated with the choice is high.

The process of moving a choice from a state of ambiguity where the distribution of outcomes is unknown to a state of risk where the distribution of outcomes is known is learning, specifically risk learning. Merely learning the expected value associated with choices without learning the risk associated would still be uninformative if the outcomes vary greatly. Here we differentiate value learning, where a prediction error is utilized to update the expected value of the associated choice and is robustly associated with neural striatal signals. Whereas risk learning is hypothesized to utilize a risk prediction error in the process of learning to update the expected risks (Preuschoff & Bossaerts, 2007; Preuschoff, Quartz, & Bossaerts, 2008). Existing neuroimaging research examining typical adults found neural substrates of risk prediction error correlations with anterior insula signals (Preuschoff et al., 2008; Wang et al., 2017). Additional research using the Iowa Gambling Task (Bechara, Damasio, & Damasio, 2000) found risk prediction error to correlate positively with activity in the insula and inferior frontal gyrus. The dorsal ACC and anterior insula were found to correlate with the "volatility" of a choice (Behrens, Woolrich, Walton, & Rushworth, 2007), an estimated parameter representing the reliability of previous information and analogous to risk prediction error. The posterior middle prefrontal cortex has also been found to correlate to another risk prediction error proxy in a dynamic learning rate, which is based on the change in magnitude of prediction errors (Krugel, Biele, Mohr, Li, & Heekeren, 2009).

While the existing literature has not examined risk learning across adolescent development, Wang et al (in preparation) isolated an estimated parameter conceptualizing a measure of risk learning ability showing a developmental trajectory where adolescents were better at tracking past risks as they aged. In addition, neuroimaging analyses found the anterior insula positively correlated with calculated risk prediction errors. Adolescents who tracked past risks better showed increased anterior insula risk prediction error signal as well. However, the

anterior insula risk prediction error signal developmental trajectory was flat, meaning additional mechanisms are needed to explain the increase in adolescents' risk learning ability over time. One possibility may be found in construct of cognitively demanding thinking, which is hypothesized to increase with age through adolescence (Davidow et al., 2018). Specifically, adolescents were found to employ simpler computational strategies than adults, who integrated more complex models accounting for counterfactual learning and learning across rewards and punishments (Palminteri, Kilford, Coricelli, & Blakemore, 2016). Adolescents also begin to utilize a mixture of complicated (model-based) and simple (model-free) learning as they aged, reflecting both capability and reliance on more cognitively demanding thinking (Decker, Otto, Daw, & Hartley, 2016; Potter, Bryce, & Hartley, 2017). Thus, we hypothesize that an increase in learning ability during adolescence relies on a combination of risk prediction error signaling localized in the anterior insula and prefrontal cortex increased ability for cognitively demanding thinking.

Wang et al (in preparation) also found individual risk learning trajectories were negatively correlated with individual real-world risky behavior trajectories in adolescents meaning adolescents who exhibited diminished improvements in their risk learning abilities partook in greater real-world risky behaviors. These findings support the notion that risk learning is crucial for transitioning adolescents away from exploratory choices made under ambiguity to informed explorative choices made with expected values and risks. Adolescents who are able to learn the risks associated with their decisions will decrease the occurrences of inherent real-world risk-taking behavior because of less exploration.

In summary, risk learning is a crucial to complement the necessary exploratory behaviors exhibited by adolescents in efforts to gain information about their environment. Without the ability for adolescents to learning the risks associated with their choices, they will remain in a constant state of exploration while exhibiting continued real-world risk-taking behavior inherent to exploration.

Risk learning facilitates the transition from decisions under ambiguity to decisions based on risk preference

Existing studies examining adolescent real-world risk-taking behavior have attributed exploration related to real-world risk-taking behavior as sensation seeking and not maladaptive

(Romer et al., 2017). As discussed in the previous section, adolescents initially make decisions under ambiguity, or a lack of information about the risks associated with the choices available to them. More specifically, choices under ambiguity involves the situation where the possibilities of the outcomes or the range of possible outcomes are unknown (Epstein, 1999; Hsu et al., 2005; Huettel et al., 2006). The preference or utility an individual has for ambiguous choices is described as ambiguity preference (Hsu et al., 2005; Huettel et al., 2006). Typical adults find ambiguous monetary lotteries undesirable (Tymula et al., 2012), meanwhile research examining children and adolescents have not found such aversion to ambiguity. In contrast to adults, adolescents were willing to accept ambiguous conditions to a greater degree (Tymula et al., 2012; van den Bos & Hertwig, 2017) and show a linear increase in ambiguity aversion with age (Blankenstein, Crone, van den Bos, & van Duijvenvoorde, 2016). Diminished ambiguity aversion found in adolescents is consistent with their tendency for exploration and sensation seeking as adolescents are more willing to choose the options with ambiguity in their outcomes or unknown outcomes. Furthermore, adolescents' tolerance of the unknown was found to associate with motivation factors and not cognitive variables (van den Bos & Hertwig, 2017). While adolescents' relative preference for ambiguous choices may be reflected as higher levels of real-world risk-taking behavior (Tymula et al., 2012), it is in line with the notion that exploration is necessary and that real-world risky behaviors as part of exploration are not maladaptive (Romer et al., 2017).

Alternatively, Romer et al. (2017) differentiate maladaptive real-world risk-taking behavior due to impulsivity and the lack of executive control over limbic motivation. This latter classification of real-world risk-taking behavior is based on the choice of an individual to choose the risky option according to their own preference. However, individuals have different levels of risk preference and the maladaptive classification is specific to those individuals who have higher risk preferences. That is, while individuals are making informed decision with known probabilities or range of outcomes, referred to by economists as the risk associated with the decision, each individual will attribute a utility based on the level of risk to their choices. The amount that risk related utility affects their choices is quantified as an individual's risk preference. For example, an individual with a risk averse preference will value a choice with high risk less by discounting the risk related utility. Risk preference is different than ambiguity preference, where the ambiguity utility is in the form of how much the distribution or range of

possible outcomes is unknown. Current research into individuals' preferences for risk and ambiguity have found weak correlations at best between them (I. Levy, Snell, Nelson, Rustichini, & Glimcher, 2010; Tymula, Rosenberg Belmaker, Ruderman, Glimcher, & Levy, 2013), making the clear distinction of separate constructs. In contrast and opposite to findings relating to ambiguity preference, adolescents are more risk averse than adults (Tymula et al., 2012). Developmentally, adolescents show linear ambiguity-aversion increase with age while showing no change in risk aversion (Blankenstein et al., 2016). Furthermore, individual differences in task-related risk-taking behavior were localized in different regions for risk (positively associated with ventral striatum activation) and for ambiguity (negatively associated with insula and dorsomedial prefrontal cortex; Blankenstein, Schreuders, Peper, Crone, & Duijvenvoorde, 2018). So typical adolescents prefer options with a smaller range of possible outcomes or choices with smaller probability of negative outcomes. This means that adolescents making maladaptive choices due to high risk preference are in the minority, and whereas most adolescents are not activity choosing real-world risk-taking behaviors knowing of their risk but choosing such behaviors because they don't know what the outcomes are.

The distinction to relating real-world risk-taking behavior to ambiguity and risk preferences would hypothesize that typical adolescents making choices based on their ambiguity preference show high, necessary, and non-maladaptive real-world risk-taking behavior while typical adolescents making choices based on their risk preferences show low maladaptive real-world risk-taking behaviors. A key point is that an individual cannot make decisions based on their risk preference when there is complete ambiguity. For an individual to be able to make decisions based on their risk preferences, they must decrease the ambiguity involved in their choices and acquire a better understanding for the probabilities. This is the reason that real-world risk-taking behavior cannot be described solely in the context of risk and ambiguity preference. That vital process needed to decrease ambiguity and gain knowledge of the risks involved in their choices is risk learning, as described earlier.

The ability for an individual to learn the risks may determine how quickly they are shifting from making decisions based on ambiguity preference to risk preference. Failure to do so, exaggerated and continued real-world risky-behaviors exhibited in adolescents' choices may be driven by adolescents diminished ambiguity aversion as part of endless exploration and sensation seeking. In this case without risk learning, even exploration and sensation seeking

becomes maladaptive as the purpose of exploration is unfulfilled in its role to provide experience, and nothing is gained by the adolescent. However, for the typical adolescent, successful risk learning means a decrease in exhibited real-world risk-taking behavior since adolescents are typically risk averse, even more so than typical adults. So, successful risk learning allows the adolescent to shift for a high preference for ambiguity to a low preference for risk reflected by corresponding high to low real-world risk-taking behavior. Developmental research showing that ambiguity preference, not risk preference, correlating more with reckless behavior (Blankenstein et al., 2016) further supports the notion that quicker transition to risk preference with better risk learning would decrease reckless behavior. The important adaptive behavior of risk learning needs to be studied in conjunction with risk and ambiguity preferences to better understand their effect on real-world risk-taking behavior through adolescence.

Relating dual systems and learning models of development

Adolescent development has been prominently described by the dual systems model of reward control (Casey et al., 2008; Shulman et al., 2016; Steinberg, 2008). The model comprises a limbic reward development trajectory offset by a reflective executive function development trajectory (Bechara & Damasio, 2005). The limbic component has been localized in the ventral striatal dopamine tract and follows a quadratic trajectory and peaks during adolescence (Crone & Dahl, 2012; Luciana, Wahlstrom, Porter, & Collins, 2012; Wahlstrom, White, & Luciana, 2011). This exaggerated ventral striatal dopamine pathway activity during adolescence has been associated with increased reward sensitivity (Galván et al., 2006; Galván & McGlennen, 2013), impulsivity (van den Bos, Rodriguez, Schweitzer, & McClure, 2015), and emotion (Hare et al., 2008). Adolescents with behavioral issues were found to have increased striatal signals (Bjork, Smith, Chen, & Hommer, 2011), while indirectly related to striatal tract, emotion and impulsivity were found to predict risk-taking behavior (Blakemore & Robbins, 2012; Robbins & Bryan, 2004).

In broader neuroimaging research, ventral striatal dopamine pathway found to be involved in motivation, drives (Robinson & Berridge, 1993), rewards, and prediction error (Montague et al., 2004; Pessiglione et al., 2006; Schultz et al., 1997). The ventral striatal dopamine tract may serve as a critical underlying neural pathway responsible for the important exploratory processes and sensation seeking in adolescents. While the hyperactive ventral striatal

tract may push adolescents toward choices with inherent real-world risks, these are adaptive behaviors necessary for adolescents to gain experience in the real world (Romer et al., 2017). Furthermore, subjective value under ambiguity, which was linked to exploration above, is also associated with the ventral striatal signals (I. Levy et al., 2010). In all, the quadratically developing limbic reward system in the dual system model of adolescent development may be the underlying driving factor for adolescents to explore their environment. As discussed in the earlier sections, such exploration may be necessary and not maladaptive.

The reflective component, second component of the dual systems model, is based around the constructs of cognitive control and executive functioning (Casey et al., 2008; van Duijvenvoorde, Achterberg, Braams, Peters, & Crone, 2016). Traditional executive functioning studies have focused on the prefrontal cortex, specifically the dorsolateral prefrontal cortex, as signals from adolescent individuals' prefrontal cortex correlate with higher executive functioning during emotional regulation and working memory tasks (Crone, 2009), addressing updating and inhibition executive functions specifically (Miyake & Friedman, 2012). The developmental trajectory through adolescence of the prefrontal cortex was hypothesized as a linear trajectory (Diamond, 2002) and it subsequently regulates striatal activity with connectivity that increases with age (van den Bos, Cohen, et al., 2012; van den Bos et al., 2015). The connection between the prefrontal cortex and striatum was also found to increase with age (van den Bos, Crone, & Güroğlu, 2012; van den Bos et al., 2015). Neuropsychological tests of impulse control and executive functioning found to predict lower real-world risky-behavior during adolescence (Pharo et al., 2011) emphasize the importance of cognitive control.

While the literature has connected executive function and cognitive control networks as important aspects of real-world risk-taking behavior, the mechanisms linking the underlying construct and behavior is still unclear. The approach found in the previous research is to relate executive function and cognitive control to neural underpinnings related to risk preference. Risk preference neural imaging studies have indeed found related signals in the inferior frontal gyrus, the lateral OFC, striatum, insula, and ACC (Paulus, Rogalsky, Simmons, Feinstein, & Stein, 2003; Tobler, O'Doherty, Dolan, & Schultz, 2007). However as discussed earlier, risk preferences may not map onto real-world risk-taking behavior since adolescents are found to be more risk averse than adults (Tymula et al., 2012) and their preferences are not related to age (Blankenstein et al., 2016). Examining executive function in relationship to risk preference may

not provide the entire story to explain real-world risk-taking behavior, rather the incorporation of risk learning is needed as well. Indeed, several regions of the executive function and control networks that have been implicated in risk learning, such as the insula, ventral striatum, and prefrontal cortex. This outlines the importance of future studies to examine both risk preference and learning to distinguish how executive functions, cognitive control, and their development influence real-world risk-taking behavior exhibited by adolescents.

Summary

Adolescence is a crucial time period during which individuals explore and learn from the outcomes of choices as they prepare for independent survival. In the process of exploration, adolescents will inevitably make choices with the possibility of serious negative outcomes such as addiction from drug use, sexually transmitted diseases and unwanted pregnancies from sexual behaviors, or even death from certain dangerous actions. However, initial real-world risk-taking behaviors in exploring ambiguous options are not necessary maladaptive and are indeed necessary as the most prevalent way adolescents can learn expected values and risks associated with the possible choices available to them (Decker et al., 2015). Furthermore, novelty-seeking behavior serves as a process for adolescents to learn to detect cues for safety and danger (Casey et al., 2008). Crucially, adolescents must be able to learn the risks from the outcomes they receive or will continue in an endless process of exploration and making choices under ambiguity. This latter scenario would be reflected in continued real-world risk-taking behavior and maladaptive for the adolescent. Adolescents who learn the risks involved in their choices would no longer be making choices under ambiguity, but rather with known risks based on their own preferences for risks. Typically, that would reflect in lower real-world risk-taking behavior as typical adolescents are risk averse compared to adults (Tymula et al., 2012). By understanding the underlying mechanisms that drive adolescents to continue partaking in real-world risky behaviors, we can find means to teach and put adolescents in positions in which they can take risks and learn, but not fall into a spiral of risky behavior.

References

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336–372. https://doi.org/10.1016/j.geb.2004.06.010

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307. https://doi.org/10.1093/cercor/10.3.295

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Bjork, J. M., Smith, A. R., Chen, G., & Hommer, D. W. (2011). Psychosocial problems and recruitment of incentive neurocircuitry: Exploring individual differences in healthy adolescents. *Developmental Cognitive Neuroscience*, *1*(4), 570–577. https://doi.org/10.1016/j.dcn.2011.07.005.Psychosocial

Blakemore, S., & Robbins, T. W. (2012). Decision-making in the adolescent brain. *Nature Neuroscience*, *15*(9), 1184–1191. https://doi.org/10.1038/nn.3177

Blankenstein, N. E., Crone, E. A., van den Bos, W., & van Duijvenvoorde, A. C. K. K. (2016). Dealing With Uncertainty: Testing Risk- and Ambiguity-Attitude Across Adolescence. *Developmental Neuropsychology*, *41*(1–2), 77–92. https://doi.org/10.1080/87565641.2016.1158265

Blankenstein, N. E., Schreuders, E., Peper, J. S., Crone, E. A., & Duijvenvoorde, A. C. K. Van. (2018). Individual differences in risk-taking tendencies modulate the neural processing of risky and ambiguous decision-making in adolescence. *NeuroImage*, *172*(January), 663–673. https://doi.org/10.1016/j.neuroimage.2018.01.085

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*, 111–126. https://doi.org/10.1196/annals.1440.010

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *Journal of Neuroscience*, *29*(39), 12315–12320. https://doi.org/10.1523/JNEUROSCI.2575-09.2009

Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and Psychological Maturation of Decision-making in Adolescence and Young Adulthood. *Journal of Cognitive Neuroscience*, *25*(11), 1807–1823. https://doi.org/10.1162/jocn_a_00447

Crone, E. A. (2009). Executive functions in adolescence: inferences from brain and behavior. *Developmental Science*, *12*(6), 825–830. https://doi.org/10.1111/j.1467-7687.2009.00918.x

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313

D'Acremont, M., Lu, Z. L., Li, X., Van der Linden, M., & Bechara, A. (2009). Neural correlates of risk prediction error during reinforcement learning in humans. *NeuroImage*, *47*(4), 1929–1939. https://doi.org/10.1016/j.neuroimage.2009.04.096

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264–1267. https://doi.org/10.1126/science.1150605

Davidow, J. Y., Foerde, K., Galván, A., & Shohamy, D. (2016). An Upside to Reward

Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence. *Neuron*, *92*(1), 93–99. https://doi.org/10.1016/j.neuron.2016.08.031
Davidow, J. Y., Insel, C., & Somerville, L. H. (2018). Adolescent Development of Value-Guided Goal Pursuit, *xx*, 1–12. https://doi.org/10.1016/j.tics.2018.05.003
Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879. https://doi.org/10.1038/nature04766
Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*(2), 285–298. https://doi.org/10.1016/S0896-6273(02)00963-7
Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective and Behavioral Neuroscience*, *15*(2), 310–320. https://doi.org/10.3758/s13415-014-0332-5
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*(6), 848–858. https://doi.org/10.1177/0956797616639301
Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In *Principles of Frontal Lobe Function* (pp. 466–503). New York, NY: Oxford University Press.
Diamond, A. (2011). Biological and social influences on cognitive control processes dependent on prefrontal cortex. *Progress in Brain Research*, *189*(1), 319–339. https://doi.org/10.1016/B978-0-444-53884-0.00032-4
Epstein, L. G. (1999). Definition of Uncertainty Aversion. *The Review of Economic Studies*, *66*(3), 579–608.
Galván, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *The Journal of Neuroscience*, *26*(25), 6885–6892. https://doi.org/10.1523/JNEUROSCI.1062-06.2006
Galván, A., & McGlennen, K. M. (2013). Enhanced striatal sensitivity to aversive reinforcement in adolescents versus adults. *Journal of Cognitive Neuroscience*, *25*(2), 284–296. https://doi.org/10.1162/jocn_a_00326
Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental Psychology*, *41*(4), 625–635. https://doi.org/10.1037/0012-1649.41.4.625
Hare, T., Tottenham, N., Galván, A., Voss, H. U., Glover, G. H., & Casey, B. J. (2008). Biological Substrates of Emotional Reactivity and Regulation in Adolescence During an Emotional Go-Nogo Task. *Biological Psychiatry*, *63*(10), 927–934. https://doi.org/10.1016/j.biopsych.2008.03.015
Hauser, T. U., Iannaccone, R., Walitza, S., Brandeis, D., & Brem, S. (2015). Cognitive flexibility in adolescence: Neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *NeuroImage*, *104*, 347–354. https://doi.org/10.1016/j.neuroimage.2014.09.018
Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making. *Science*, *310*(5754), 1680–1683. https://doi.org/10.1126/science.1115327
Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural

Signatures of Economic Preferences for Risk and Ambiguity. *Neuron*, *49*(5), 765–775. https://doi.org/10.1016/j.neuron.2006.01.024
Jaccard, J., Blanton, H., & Dodge, T. (2005). Peer influences on risk behavior: an analysis of the effects of a close friend. *Developmental Psychology*, *41*(1), 135–147. https://doi.org/10.1037/0012-1649.41.1.135
Javadi, A. H., Schmidt, D. H. K., & Smolka, M. N. (2014). Differential representation of feedback and decision in adolescents and adults. *Neuropsychologia*, *56*(1), 280–288. https://doi.org/10.1016/j.neuropsychologia.2014.01.021
Javadi, A., Schmidt, D., & Smolka, M. (2014). Adolescents Adapt More Slowly than Adults to Varying Reward Contingencies. *Journal of Cognitive Neuroscience*, 2670–2681. https://doi.org/10.1162/jocn
Kendler, K. S., Sundquist, K., Ohlsson, H., Palmér, K., Maes, H., Winkleby, M. A., & Sundquist, J. (2012). Genetic and Familial Environmental Influences on the Risk for Drug Abuse: A National Swedish Adoption Study. *Archives of General Psychiatry*, *69*(7), 690–697. https://doi.org/10.1001/archgenpsychiatry.2011.2112
Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(42), 17951–17956. https://doi.org/10.1073/pnas.0905191106
Levy, D. J., & Glimcher, P. W. (2011). Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. *Journal of Neuroscience*, *31*(41), 14693–14707. https://doi.org/10.1523/JNEUROSCI.2218-11.2011
Levy, I., Snell, J., Nelson, A. J., Rustichini, A., & Glimcher, P. W. (2010). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, *103*(2), 1036–1047. https://doi.org/10.1152/jn.00853.2009
Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, *7*(3), 274–281. https://doi.org/10.1093/scan/nsr006
Luciana, M., Wahlstrom, D., Porter, J. N., & Collins, P. F. (2012). Dopaminergic modulation of incentive motivation in adolescence: Age-related changes in signaling, individual differences, and implications for the development of self-regulation. *Developmental Psychology*, *48*(3), 844–861. https://doi.org/10.1037/a0027432
Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458
Montague, P. R., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936–1947. Retrieved from http://www.jneurosci.org/content/16/5/1936.short
Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational Roles for Dopamine in Behavioural Control. *Nature*, *431*(7010), 760–767. https://doi.org/10.1038/nature03015
Oliva, E. M., Keyes, M., Iacono, W. G., & McGue, M. (2012). Adolescent Substance Use Groups: Antecedent and Concurrent Personality Differences in a Longitudinal Study. *Journal of Personality*, *80*(3), 769–793. https://doi.org/10.1111/j.1467-6494.2011.00746.x
Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S.-J. (2016). The computational development of reinforcement learning during adolescence. *PLoS Computational Biology*, 1–25. https://doi.org/10.1371/journal.pcbi.1004953

Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S., & Stein, M. B. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage*, *19*(4), 1439–1448. https://doi.org/10.1016/S1053-8119(03)00251-9

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-Dependent Prediction Errors Underpin Reward-Seeking Behaviour in Humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Pharo, H., Sim, C., Graham, M., Gross, J., & Hayne, H. (2011). Risky business: Executive function, personality, and reckless behavior during adolescence and emerging adulthood. *Behavioral Neuroscience*, *125*(6), 970–978. https://doi.org/10.1037/a0025768

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. https://doi.org/10.1016/j.dcn.2016.10.005

Preuschoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. In *Annals of the New York Academy of Sciences* (Vol. 1104, pp. 135–146). https://doi.org/10.1196/annals.1390.005

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *The Journal of Neuroscience*, *28*(11), 2745–2752. https://doi.org/10.1523/JNEUROSCI.4286-07.2008

Ramstetter, C. L., Murray, R., & Garner, A. S. (2010). The crucial role of recess in schools. *Journal of School Health*, *80*(11), 517–526. https://doi.org/10.1111/j.1746-1561.2010.00537.x

Robbins, R. N., & Bryan, A. (2004). Relationships Between Future Orientation, Impulsive Sensation Seeking, and Risk Behavior among Adjudicated Adolescents. *Journal of Adolescent Research*, *19*(4), 428–445. https://doi.org/10.1177/0743558403258860

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev*, *18*(3), 247–291. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8401595

Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex*, *18*(3), 652–663. https://doi.org/bhm097 [pii] 10.1093/cercor/bhm097

Romer, D., Reyna, V. F., & Satterthwaite, T. D. (2017). Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Developmental Cognitive Neuroscience*, *27*(July), 19–34. https://doi.org/10.1016/j.dcn.2017.07.007

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, *17*, 103–117. https://doi.org/10.1016/j.dcn.2015.12.010

Silveri, M. M., Tzilos, G. K., Pimentel, P. J., Yurgelun-Todd, D. A., Pimental, P. J., & Yurgelun-Todd, D. A. (2004). Trajectories of Adolescent Emotional and Cognitive Development: Effects of Sex and Risk for Drug Use. *Annals of the New York Academy of Sciences*, *1021*(Adolescent Brain Development: Vulnerabilities and Opportunities), 363–370. https://doi.org/10.1196/annals.1308.046

Sinha, J. W., Cnaan, R. a, & Gelles, R. W. (2007). Adolescent risk behaviors and religion : Findings from a national study. *Journal of Adolescence*, *30*(2), 231–249.

https://doi.org/10.1016/j.adolescence.2006.02.005
Smith, D. V, Hayden, B. Y., Truong, T.-K., Song, A. W., Platt, M. L., & Huettel, S. a. (2010). Distinct Value Signals in Anterior and Posterior Ventromedial Prefrontal Cortex. *Journal of Neuroscience*, *30*(7), 2490–2495. https://doi.org/10.1523/JNEUROSCI.3319-09.2010
Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*(1), 78–106. https://doi.org/10.1016/j.dr.2007.08.002
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge Univ Press.
Telzer, E. H., Ichien, N. T., & Qu, Y. (2015). Mothers know best: redirecting adolescent reward sensitivity toward safe behavior during risk taking. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsv026
Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, *97*, 1621–1632. https://doi.org/10.1152/jn.00745.2006
Tymula, A., Rosenberg Belmaker, L. A., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences*, *109*(42), 17135–17140. https://doi.org/10.1073/pnas.1207144109
Tymula, A., Rosenberg Belmaker, L. A., Ruderman, L., Glimcher, P. W., & Levy, I. (2013). Like cognitive function, decision making across the life span shows profound age-related changes. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 17143–17148. https://doi.org/10.1073/pnas.1309909110
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*(6), 1247–1255. https://doi.org/10.1093/cercor/bhr198
van den Bos, W., Crone, E. A., & Güroğlu, B. (2012). Brain function during probabilistic learning in relation to IQ and level of education. *Developmental Cognitive Neuroscience*, *2*, S78–S89. https://doi.org/10.1016/j.dcn.2011.09.007
van den Bos, W., & Hertwig, R. (2017). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Scientific Reports*, *7*(January), 40962. https://doi.org/10.1038/srep40962
van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *Proceedings of the National Academy of Sciences*, *112*(29), E3765–E3774. https://doi.org/10.1073/pnas.1423095112
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. https://doi.org/10.1016/j.neuroimage.2015.04.069
Vives, M.-L., & FeldmanHall, O. (2018). Tolerance to ambiguous uncertainty predicts prosocial behavior. *Nature Communications*, *9*(1), 2156. https://doi.org/10.1038/s41467-018-04631-9
Wahlstrom, D., White, T., & Luciana, M. (2011). Neurobehavioral evidence for changes in dopamine system activity during adolescence. *Neuroscience & Biobehavioral Reviews*, *34*(5), 631–648. https://doi.org/10.1016/j.neubiorev.2009.12.007.Neurobehavioral
Wang, Y., Ma, N., He, X., Li, N., Wei, Z., Yang, L., … Zhang, X. (2017). Neural substrates of updating the prediction through prediction error during decision making. *NeuroImage*, *157*(May), 1–12. https://doi.org/10.1016/j.neuroimage.2017.05.041

Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., … Ziesch., V. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature*, *512*(7513), 185–189. https://doi.org/10.1038/nature13402

Paper 2

Title: Diminished Development of Risk Learning in Adolescence is Associated with Increased Real-World Risk-Taking Behavior

John M. Wang[1, 2], Zhuoya Cui[1], Alec Solway[1*], Vanessa Brown[1, 2*] Lusha Zhu[1*], Jungmeen Kim-Spoon[1], Pearl Chiu[1,2], & Brooks King-Casas[1,2,3]

[1]Fralin Biomedical Research Institute at Virginia Tech Carilion, Roanoke, VA
[2]Department of Psychology, Virginia Tech, Blacksburg, VA
[3]Virginia Tech-Wake Forest University School of Biomedical Engineering and Science
* These authors contributed equally.

Abstract

Adolescence is a transition period during which individuals have increasing autonomy in decision-making for themselves, often choosing among options about which they have little knowledge and experience. This process of individuation and independence depends critically on the ability to explore and learn potential consequences of actions within novel environments. Here we suggest that success or failure with learning these consequences contributes to individual differences in real-world risk-taking behaviors observed in adolescence. Specifically, failure to learn means an adolescent will continue to make decisions about options for which outcomes are unknown. To test this possibility, we examined components of learning about mean value and risk (mathematical variance) longitudinally over the course of four years using neural responses and behavioral choices as N = 144 adolescents made decisions about ambiguous options in a reinforcement-learning task. Longitudinally, computational modeling showed adolescents improving their abilities to track past risks as they aged, as quantified by an individually estimated risk learning rate parameter. Neurally, better tracking of past risks was associated with higher anterior insula risk prediction error signal across adolescents at all timepoints. No such relationship was found between ventral striatum value prediction error signal and risk learning rate. The developmental trajectory of risk learning was also negatively correlated with real-world risk-taking trajectory, such that adolescents who improved in tracking of past risks showed a decrease in real-world risky behavior over time. These data indicate that neural mechanisms of risk learning, rather than mean value learning, are related to maladaptive behaviors among adolescents. More specifically, risk learning capacity, typically gained through exploration of novel decision-environments, is suggested as a critical determinant of real-world risk-taking behavior during adolescence.

Introduction

Adolescence is a stage of development during which individuals learn about the environment to prepare for independent survival (Casey, Jones, & Hare, 2008). Learning from the outcomes of choices during this time is crucial to optimize the results of future decisions. However, adolescence is also known as a period during which individuals display increased risk-taking behaviors (Silveri et al., 2004) and subsequently experience increased rates of negative consequences, including motor accidents, unwanted pregnancies, sexually transmitted diseases, drug addictions, and death (Casey et al., 2008). Research suggests that risk-taking behavior and risky decisions shown in adolescents may be the result of exploring the unknown as adolescents were shown to increase exploratory behavior and shift between choices regardless of outcome quality in on a variant of the Iowa Gambling Task (Christakou et al., 2013). Here we posit that the success with which adolescents learn from exploring the unknown may in turn determine whether individuals continue to partake in risk-taking behaviors. Simply put, if an adolescent does not learn from the outcomes associated with risky choices, they may continue to engage in these risk-taking behaviors.

Laboratory-based reinforcement learning paradigms commonly focus on the value associated with possible decisions; here we emphasize the importance of also considering the risk associated with a choice. Specifically, risk learning allows for the utilization of information from the possibility of catastrophic outcomes, such as real-world risk-taking behavior. Such information regarding risks can critically change an individual's decisions in the form of risk preference (for review, see Weber & Johnson, 2009). Risk learning has received limited investigation, possibly because the computational psychiatric approach is relatively in its inaugural stages of its potential uses. In a posited risk learning tasks, 'risk' is equivalent to the mathematical variance of possible outcomes associated with a choice and an update rule in conjunction with a risk prediction error, the difference between the actual variance of the outcome and the expected variance associated with that choice (Preuschoff & Bossaerts, 2007), can be used to track risks associated with a choice.

Such a risk learning model draw upon traditional value based reinforcement learning models, which uses update rules to track the expected values associated with specific options (TDRL; Montague, Dayan, & Sejnowski, 1996; Montague, Hyman, & Cohen, 2004). Signals in

the ventral striatal dopamine system encode value prediction error signals representing the difference of received outcome and expected outcome. In typical individuals, these signals are subsequently used to neurally update future expectations (Jocham, Klein, & Ullsperger, 2014; Montague, King-Casas, & Cohen, 2006; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). Adolescents show ventral striatal prediction error signals, similar to adults (Davidow, Foerde, Galván, & Shohamy, 2016; Hauser, Iannaccone, Walitza, Brandeis, & Brem, 2015; A. H. Javadi, Schmidt, & Smolka, 2014; A. Javadi, Schmidt, & Smolka, 2014; van den Bos, Cohen, Kahnt, & Crone, 2012), and these signals in adolescents have been shown to develop quadratically as individuals aged (Cohen et al., 2010). Developmental theories suggest exaggerated striatal activity to rewards play a role for adolescents at-risk for risky, reward-seeking behavior that can lead to accidental death and the onset of drug addiction (Cohen et al., 2010; Galván et al., 2006; Galván & McGlennen, 2013) with increasing striatal signals to rewards through development (Bjork, Smith, Chen, & Hommer, 2011).

Studies examining the neural substrates of risk prediction error in typical individuals have found correlations with insula activity (Preuschoff, Quartz, & Bossaerts, 2008; Wang et al., 2017). Similarly, insula activity was also found to associate with risk preference in typical individuals (Huettel, Stowe, Gordon, Warner, & Platt, 2006). This suggests risk-taking behaviors may be related to variation in the abilities of individuals to learn about the risks and the neural update of the risks, in the form of a risk prediction error learning signal, associated with choices. Additionally, these studies point to the insula an important region of interest for risk prediction error, possibly dissociable from the ventral striatum in value prediction errors.

Lastly, existing within adolescence literature have found adolescents with a propensity toward ambiguous choices – where the variance of the choice is unknown (Tymula et al., 2012) – and exploration of unknown outcomes (Christakou et al., 2013) which manifest as risk-taking behavior (Tymula et al., 2012). These same adolescents were found to show risk-averse behavior (Tymula et al., 2012), in that they tend to value choices less when the variance of a choice is high. Such results would suggest adolescents would prefer not to partake in behavior with high real-world risk-taking behavior. However, no existing studies have examined the relationship between risk learning and real-world risk-taking behavior. Such examination into risk learning may link the differences in ambiguity and risk preferences exhibited by adolescents. Specifically, we suggest that adolescents' ability to learn risks during the initial process of exploration, as an

exhibition of their ambiguity seeking preference, may be crucial for the adolescent to make informed decisions going forward base on their risk aversive preference. Conversely, adolescents who are unable to learn the risks involved in their options may continue to explore, which will be exhibited as continued risk-taking behavior. We hypothesized that adolescent's individual abilities to learn and corresponding neural learning signal specifically to values, risks, or both will directly affect whether adolescent individuals' real world risky-behaviors.

Results

Participants & task

To test whether the developmental trajectories of neural and behavioral components of value and risk learning are related to real-world risk-taking behaviors during adolescence, adolescent participants were recruited to complete a reinforcement learning task during fMRI scanning (Fig 1a) once per year over the course of four years. 144 individuals (71 females; Time 1 mean age: 14.09 (s.d. 0.55), Time 2: 15.10 (0.56), Time 3: 16.11 (0.56), Time 4: 17.04 (0.56)) completed at least two successful fMRI scan sessions of the task as part of a larger study. The task entailed learning from repeated choices in a 3-arm bandit, depicted as 3 slot machines, over 2 runs with 100 trials in each run. The trials comprised independent and randomly varying outcomes between 0 and 100. The outcome drifted with a random walk to introduce a component of risk to each option. Each trial was self-paced and participants observed the outcome of their chosen option before moving on to the next trial. In order for the adolescent to perform well, they needed to learn the value and risk associated with each slot machine. Adolescents performed well above chance, showing they were learning as the outcomes shifted (Fig. 1b). Adolescents were compensated dependent on the outcomes of 5 trials drawn randomly from each adolescent's actual choices.

Computational modeling

To characterize each individuals' risk learning aptitudes, participants' behavioral choices for each time point were fit to a Kalman filter reinforcement learning model (Daw, O'Doherty,

Dayan, Seymour, & Dolan, 2006) with the inclusion of a risk learning rate parameter (ρ; Fig. 1c). This model provides a trial by trial valuation of the available choices based on a value prediction error, risk prediction error, and estimated parameters including the risk learning rate. Value prediction error is the difference between the outcome and the expected value, while risk prediction error is the difference between the magnitude of the value prediction error and the past expected risk (mathematical variance). The risk learning rate in turn balances the tracking of the past expected risk and new risk information in the form of risk prediction errors for the calculation of the updated expected risk. Subsequently, the expected risk modulates the trial-by-trial value learning rate, which updates the expected value. Individual parameters were estimated using a Bayesian Hierarchical method (Daw, 2011) utilizing Hamiltonian Markov Chain Monte Carlo as implemented in Stan (Carpenter et al., 2017) and fit better than the standard RL model (See Model Fitting and Selection).

Developmental trajectory of risk learning

To evaluate changes in adolescents' risk learning over the four years, individual risk learning rates were regressed onto the age of the individual using a Hierarchical Linear Model (HLM) with random effects of age and subject. There was an significant effect of age on risk learning rate ($\beta = 0.177$, $t(479) = 4.30$, $p < 0.001$; Figure 2), meaning adolescents track past risk information at a greater rate as they age.

Neural correlates of risk and value prediction error

To differentiate between neural correlates of risk and value learning, trial-by-trial value prediction errors and risk prediction errors were calculated using each adolescent's parameters for each time point. To examine risk prediction error neural correlates, risk prediction error was modeled as the second regression while value prediction error was the first regressor. In the whole brain analysis, the risk prediction error signal was found in the bilateral insula and bilateral striatum (Fig. 3a; Table 2) after cluster level FWE correction for multiple comparisons. Similarly for neural correlates of value prediction error, trial by trial value prediction errors were regressed onto the outcome event of each individual's BOLD response as the second regressor

while calculated risk prediction error was the first regressor, accounting for effects of risk prediction error. Meanwhile, the value prediction error signal was found in the right ventral striatum (Fig. 3b; Table 3), again after correcting for multiple comparisons.

Risk learning rate effect on neural correlate of risk and value prediction error

To validate the cognitive construct of risk learning rate parameter, we examined the relationship between each individuals' risk learning rate and their neural value and risk learning signals, neural risk and value prediction error signals (Anatomical regions of interest (ROIs) of the bilateral anterior insula (aal library) and the bilateral ventral striatum (Garrison, Erdeniz, & Done, 2013) were used to extract individual beta values for risk prediction error and value prediction error) was regressed on individual estimates of risk learning rate using HLM regression with random effects for risk learning rate and subject. There was an significant effect of risk learning rate on anterior insula risk prediction error signal ($\beta = 0.142$, $t(470) = 3.20$, $p = 0.001$; Figure 3c) but not for ventral striatum value prediction error signal ($\beta = 0.001$, $t(159) = 0.012$, $p = 0.990$; Figure 3c). These results meant better tracking of past expected risk was related to greater anterior insula risk prediction error signal. Furthermore, the effect was specific to anterior insula risk prediction error signal and not to ventral striatum value prediction error signal.

To verify the validity of risk learning rate and neural risk prediction error signal, we correlated these measures with task performance in the form of average score per trial. Task performance was positively correlated with both risk learning rate (Figure S2; $r = 0.150$, $p < 0.001$) and anterior insula risk prediction error signal (Figure S2; $r = 0.241$, $p < 0.001$). However, ventral striatum value prediction error signal did not correlate with performance (Figure S2; $r = -0.025$, $p = 0.581$). As expected, providing convergent evidence that risk learning, and not value learning, is the vital learning process needed for long term success.

Longitudinal developmental trajectories related by learning

To examine the longitudinal trajectories of neural learning signals, anterior insula risk prediction error signal and ventral striatum value prediction error signal were separately

regressed onto individual's age using HLM regression with random effects for age and subject. From the HLM regression, the group longitudinal trajectory is the estimated group slope coefficients when regressed onto the age variable. There was no effect of age on Anterior Insula risk prediction error signal ($\beta = 0.005$, $t(115) = 0.099$, $p = 0.921$; Figure S3). However as expected from previous literature (van den Bos et al., 2012; van den Bos, Rodriguez, Schweitzer, & McClure, 2015), there was a significant effect of age on Left Ventral Striatum value prediction error signal ($\beta = 0.110$, $t(137) = 2.26$, $p = 0.025$).

To examine real-world risk-taking behaviors in adolescents, the Things I do (Conger & Elder, 1995) self-report survey was collected at each of the four lab visits. The survey asks adolescents whether they have engaged in a list of real-world risk-taking behaviors (see Supplemental for complete list), divided into subscales of Major and Minor Risks. When Major Risks subscale was regressed onto age using HLM regression with random effects for age and subject, there was a significant effect of age ($\beta = 0.319$, $t(137) = 8.24$, $p < 0.001$; Figure 4). There was no effect of age on Minor Risks ($\beta = 0.067$, $t(124) = 1.64$, $p = 0.103$).

Lastly to assess how adolescents' developmental trajectories of risk learning with real world taking behavior, we first had to extract the individual slopes of the separate dependent variable (risk learning rate, anterior insula risk prediction error signal, value prediction error signal, and Things I do subscales) regressed onto age in the aforementioned HLM regression, which is possible with the subject random effect in the regression. Next to we correlated individuals' trajectories of risk learning related to their neural correlates of risk and value learning. Risk learning trajectory was found to correlate negatively with Anterior Insula risk prediction error trajectory (Figure 5a; $r = -0.236$, $p = 0.005$), meaning adolescents who improved at tracking past risks showed decreases in anterior insula risk prediction error signal. A possible explanation for this result is that those who showed lack of improvement in their risk learning rate would exhibit greater and greater anterior insula risk prediction error signal to compensate. There was no relationship between the trajectories of risk learning rate and ventral striatum value prediction error signals however (Figure 5b; $r = 0.100$, $p = 0.234$).

Finally comparing the trajectory of the risk learning rate with changes in the Major Risk subscale of the Things I Do self-report survey (Conger & Elder, 1995), there was a significant negative correlation (Figure 5c, $r = -0.220$, $p = 0.005$). Adolescents who improved in tracking past risks exhibited less real-world risk-taking behavior. This relationship supports the

hypothesis that adolescents who are not encoding risk learning signals in anterior insula are more likely to show externalizing behaviors.

Discussion

Risk learning quantification

Here in this longitudinal study examining risk learning and real-world risk-taking behavior in adolescence, a computational modeling approach captured the progression of how adolescents learned in changing environments over the course of four test sessions across four years. Individual estimates of a risk learning rate (1-ρ) found a longitudinal pattern where over time, adolescents were better at keeping track of past risks during the task (Figure 2). Better tracking of past risks, as quantified by higher risk learning rate, was related to better performance in the reinforcement learning task (Figure S2a) providing validation that risk learning rate is a useful measure of adolescents' learning abilities.

Additionally, separate neural signals encoding both risk learning (risk prediction errors) and value learning (value prediction errors) were identified in adolescents to assess whether learning components are associated with the risk learning rate. First, as expected, the anterior insula and ventral striatum were found to track risk prediction errors, consistent with previous findings of risk prediction error in the anterior insula in a typical population (D'Acremont, Lu, Li, Van der Linden, & Bechara, 2009; Preuschoff et al., 2008; Wang et al., 2017). Meanwhile, the ventral striatum tracked value prediction errors in adolescents consistent with previous studies of both adolescent individuals and typical adults (Cohen et al., 2010; Hauser et al., 2015; Jocham et al., 2014; Montague et al., 2006; Pessiglione et al., 2006). Both risk and value learning signals found in our adolescent sample matched that of our adult control sample (see Supplement).

Adolescents' individually estimated risk learning rate was found to have a significant positive effect on their anterior insula risk prediction error signal, but not with striatal value prediction error. Adolescents engaged their anterior insula when they kept better track of previous risk information related to their choices (Figure 3c, see Supplement). The functionality of the insula is associated with interoception, or the awareness of one's internal states, with the

role of maintaining homeostasis that extends to subjective feelings (Wood & Bechara, 2014). Anatomically, the anterior insula has reciprocal connections (Craig, 2002) with the ventral striatal dopamine system that encodes value prediction error signals. As part of the subjective interoceptive feelings, anterior insula is hypothesized to contribute to motivating factors such as craving or inhibition that may regulate the ventral striatum (Wood & Bechara, 2014). In the present data, the anterior insula signal encodes risk prediction error, a measure updating the stability of the chosen outcome and may serve as an internal measure of homeostatic environment. The relationship between risk learning rate and anterior insula signal may represent the individual possessing a more active response to changes in the homeostatic environment, which subsequently leads to better learning of valuations through downstream modulations on their ventral striatum. This relationship further validates risk learning rate as a measure quantifying individuals' ability to learn about risks.

Developmental trajectories risk learning and neural development

The traditional theory of adolescent neural development is the dual process model (Casey et al., 2008). The model is comprised of an impulsive component offset by a reflective component (Bechara & Damasio, 2005). The impulsive component has been correlates in the ventral striatal dopamine system and shows quadratic development through adolescence (Crone & Dahl, 2012; Luciana, Wahlstrom, Porter, & Collins, 2012; Wahlstrom, White, & Luciana, 2011) and adolescents with behavioral issues have been found to have increased striatal signals (Bjork et al., 2011). Longitudinal analyses of ventral striatal value prediction error signal in this current study supports these findings as adolescents display increasing ventral striatal value prediction error signal over time (see Supplement). In contrast, the reflective component is based around the constructs of cognitive control and executive functioning (Casey et al., 2008; van Duijvenvoorde, Achterberg, Braams, Peters, & Crone, 2016). Traditional executive functioning studies have focused on the prefrontal cortex, specifically the dorsolateral prefrontal cortex, as signals from adolescent individuals' prefrontal cortex correlate with higher executive functioning during emotional regulation and working memory tasks (Crone, 2009), addressing updating and inhibition executive functions specifically (Miyake & Friedman, 2012). The development trajectory through adolescence of the prefrontal cortex was hypothesized as a linear trajectory

(Diamond, 2002) and it subsequently regulates striatal activity with connectivity that increases with age (van den Bos et al., 2012, 2015).

Although anterior insula has not been central to dual process theories of development, data indicate the anterior insula plays a prominent role in human decision-making and the behaviors observed during adolescence (Wood & Bechara, 2014). Previous research has shown the insula directly connects to the dual process system both anatomical and functional connectivity with the ventral striatum (Craig, 2002) and the prefrontal cortex (Cai, Ryali, Chen, Li, & Menon, 2014; Craig, 2002; van Duijvenvoorde et al., 2016). It is also suggested the anterior insula detecting salient events such as errors (Cai et al., 2014) and translating homeostatic states into conscious feelings (Wood & Bechara, 2014), whereas the dual process system of executive function implements the control on impulsive actions.

Current findings in this study shows higher anterior insula risk prediction error signals is associated with better retention of past expected risks. This supports interpretations of the anterior insula risk prediction error acting as a monitoring signal that subsequently can engage individual's executive function systems as improved executive function lowers real-world risk-taking behavior. Longitudinally, adolescents' anterior insula risk prediction error signal did not change over time. This is consistent with the dual theory predictions of longitudinal changes in the striatum (Crone & Dahl, 2012; Luciana et al., 2012; Wahlstrom et al., 2011) and prefrontal cortex (Diamond, 2002). However, there was a negative correlation between adolescents' developmental trajectories for risk learning rate and their anterior insula risk prediction error signal (Figure 5a). This meant adolescents who showed improvement in their abilities to track past risk information exhibited a decrease in their anterior insula risk prediction error signal. One possible explanation is that adolescents who show lack of improvement in their risk learning rate would exhibit greater and greater anterior insula risk prediction error signal. Such changes in anterior insula risk prediction error signal might be an adaptative process signaling the need for the individual's executive functioning systems to attend to such relevant homeostatic information.

For experience-naïve individuals, exploration is needed and novelty-seeking is indeed characteristic of adolescence (Casey et al., 2008). During such exploration, however, real-world risk-taking behaviors may not be negative insofar as exploration is a necessary process whereby adolescents can observe the possible variability of outcomes, or the risks, associated with the

choices available to them. Such behaviors may be facilitated by the development of hyperactive ventral striatum reward systems. From there, adolescents may learn the risks and valuations of their choices, which informs them during future decision. However, adolescents may still also continue to explore among seemingly ambiguous options. Here, we suggest that such continuous exploration manifests as real-world risky behaviors, leading to adolescents at-risk for the negative consequences of these behaviors. Based on our findings, we hypothesized that adolescents who display lower abilities to learn about risks are predisposed to make additional risks reflected by increased real-world risk-taking behavior.

Developmental trajectories comparison of risk learning and real-world risk-taking behavior

To assess real-world risk-taking behavior changes during adolescences in relation to individual's abilities to learn about risks, individuals completed the Things I Do (Conger & Elder, 1995) self-report survey every year during the four years of the study. As expected, there was a significant longitudinal effect of age on the Major Risk subscale showing that adolescents displayed more real-world risk-taking behavior over time (Figure 4). Examples of items in the Major Risk subscale items included whether the individual drank alcohol, whether the individual used tobacco products, or whether the individual threatened the use of violence to achieve their goals. No such longitudinal effects were found in the Minor Risk subscale.

Longitudinal trajectory of adolescents' risk learning rate was negatively correlated with the trajectories of the Major Risk subscale. That is individuals who did not improve at tracking risks showed more real-world risk-taking behaviors over time (Figure 5c). This is consistent with the possibility that adolescents who lack the ability to learn about the risks involved in their choices and show diminished neural risk updating signal, may also not able to overcome their hyperactive impulsivity. They will continue to explore options they have encountered before exhibited as continued risk-taking behavior, perhaps never comprehending the behavior as risky.

Conclusion

In summary, adolescent individuals showed improved abilities to track past risks, quantified by an individually estimated risk learning rate, during a reinforcement-learning task as they aged. During the task, the anterior insula signal encoded risk prediction error while the ventral striatum encoded value prediction error during a reinforcement-learning task. The individually estimated risk learning rate had a significant positive effect on individual's anterior insula risk prediction error signal, not on their ventral striatum value prediction error signal. Furthermore, adolescents' learning rate developmental trajectories was negatively correlated with their real-world risk-taking behavior trajectory meaning individuals who did not improve their risk learning abilities exhibited increased real-world risk-taking behavior. These findings support the hypothesis that risk learning plays a crucial role in adolescents' ability to avoid the continued maladaptive risk-taking behavior that comes with the necessity of exploration. Adolescents with lower anterior insula risk learning signals and lower risk learning rate are ignorant to the risks associated with their choices and leading to continued exploration that is realized as continued risk-taking behavior.

Materials and methods

Participants

144 adolescent individuals ages 13 and 18 were included in the current study over the course of 4 years as part of a larger adolescence longitudinal developmental study. Overall, the study recruited a total of 171 adolescents with 144 completing returning to complete at least two sessions of the continuous learning task. All participants provided informed consent and all procedures were approved by the Institutional Review Board of Virginia Tech. Exclusion criteria consisted of: claustrophobia, history of head injury resulting in loss of consciousness for more than 10 minutes, orthodontia impairing image acquisition, severe psychopathology (e.g., psychosis), and other contraindications to MRI (e.g. pacemaker, aneurysm clips, neurostimulators, cochlear implants, metal in eyes, steel worker, or other implants). All exclusion criteria were assessed through self-report.

67 adult individuals (58 females; mean age 43.40, s.d. 6.71) were recruited from the parents of the adolescents as an adult comparison sample.

Continuous learning task design

Participants performed a three-arm bandit reinforcement-learning task during a yearly session over the course of the four years. As illustrated in **Fig. 1a**, on each trial subjects chose between three slot machine stimuli and subsequently observed the outcome. Trials were presented repeatedly for 100 trials per block for 2 blocks, one block with positive outcomes between 0 and 100 and the other with negative outcomes between -100 and 0. Each trial consisted of 3 slot machines, which required participants to learn the contingencies between stimuli and outcomes within each block. During each trial, the participant was first shown 3 slot machines. Then the participant chose one of the slot machines at their own pace followed by a jittered inter trial interval. Lastly, the outcome of their choice was presented and remained for a jittered interval. The jittered timing was based on a gamma distribution function distribution with minimum of 2 seconds and maximum of 6.5 seconds.

The outcome sequences of each stimulus were from one of 12 independently generated sequences. The outcome for the trial (t) of a sequence was drawn from a Gaussian distribution (standard deviation of 4) around a mean (μ_t) and rounded to the nearest integer. μ_t for each trial diffused in a decaying Gaussian random walk, with $\mu_{t+1} = \lambda\mu_t + (1 - \lambda)\theta + v$ where the decay parameter (λ) was 0.925, the decay center (θ) was 50, and the diffusion noise v was a zero-mean Gaussian (standard deviation = 5).

As implemented, this task took approximately 20 minutes per block, and each participant played two blocks to ensure task comprehension prior to entering the scanner. Participants were informed that their performance in the game determined their payment. Participants were endowed with $10 and the outcome of 5 random trials per block were converted into monetary payout and added to their initial endowment.

Model Fitting and Selection

The hypothesized learning model was first fit to the observed behavioral data with the best fitting model used for subsequent fMRI analysis. The model was a Kalman filter model (Figure 1C; Daw et al., 2006), which implements the Bayesian mean-tracking rule for the outcome sequences, modified to include a trial-by-trial update of risk, or risk prediction error, for the purpose of neuroimaging analysis.

Modified Kalman filter model. The modified Kalman filter assumes the subject is learning the valuations based on the parameters ρ, $\hat{\sigma_d}$, $\hat{\lambda}$, and $\hat{\theta}$. For trial number t, a prior distribution over the true mean payoffs $\mu_{i,t}$ of the choice i are independent Gaussians, $N(\hat{\mu_{i,t}}^{pre}, \hat{\sigma^2}_{i,t}{}^{pre})$. The initial distribution of each choice was set to values used to generate each sequence, $N(50, 5)$. The posterior mean for the option i after the outcome r_t is:

$$\hat{\mu}_{i,t}^{post} = \hat{\mu}_{i,t}^{pre} + \kappa_t \delta_t$$

The prediction error δ_t is:

$$\delta_t = r_{i,t} - \hat{\mu}_i^{pre}$$

The trial-by-trial learning rate κ_t is calculated as follows:

$$\kappa_t = \frac{\hat{\sigma}_{i,t}^{2\,pre}}{\hat{\sigma}_{i,t}^{2\,pre} + \hat{\sigma}_{o,i,t}^2}$$

With the variance of the outcome $\hat{\sigma}^2_{o,i,t}$ being:

$$\hat{\sigma}_{o,i,t}^2 = \hat{\sigma}_{o,i,t-1}^2 + \rho(\delta_t^2 - \hat{\sigma}_{o,i,t-1}^2)$$

The parameter ρ is the updating parameter for the variance (risk) that utilized the risk prediction error δ^2_t - $\hat{\sigma}^2_{o,i,t-1}$.

The posterior variance (risk) for the chosen option is:

$$\hat{\sigma}_{i,t}^{2\,post} = (1 - \kappa_t)\hat{\sigma}_{i,t}^{2\,pre}$$

The posterior mean and variance (risk) for the unchosen options are unchanged. To account for the diffusion of the outcomes, the distributions $N(\hat{\mu}_{i,t+1}^{pre}, \hat{\sigma}^2_{i,t+1}{}^{pre})$ for trial $t+1$ for all choices are given by:

$$\hat{\mu}_{i,t+1}^{pre} = \hat{\lambda}\hat{\mu}_{i,t}^{post} + (1 - \hat{\lambda})\hat{\theta}$$
$$\hat{\sigma}_{i,t+1}^{2\,pre} = \hat{\lambda}^2\hat{\sigma}_{i,t}^{2\,post} + \hat{\sigma}_d^2$$

As noted in Daw et al. (2006), error-driving learning rule in the Modified Kalman Filter model is the same as standard Q-learning models. The main difference is the updating of trial-by-trial

expected variance (risk) $\hat{\sigma^2}_{i,t}$ in relationship the updated variance of the outcome $\hat{\sigma^2}_{o,i,t}$ to determine trial-by-trial learning rates κ_t.

Lastly, the softmax action selection function was used for the choice rule. The probability of selecting choice *i* at time *t* was estimated as follows:

$$P_{i,t} = \frac{\exp(\tau\, \hat{\mu}_{i,t}^{pre})}{\sum_j \exp(\tau \hat{\mu}_{j,t}^{pre})}$$

Inverse temperature τ is an estimated exploration parameter that quantifies the balance between the exploitation of the higher valued option and exploration of the other option for information.

Model Fitting. Reinforcement learning models and parameters were fit using hierarchical Bayesian estimation (Daw, 2011) utilizing Hamiltonian Markov Chain Monte Carlo implemented in Stan (Carpenter et al., 2017) and R.

Group level parameters were specified as normally distributed. ρ, $\hat{\lambda}$, *and* α of their respective models were inverse logit transformed to be between 0 and 1. β and $\hat{\sigma_d}$ were bound with the minimum of 0 to be positive. ρ was given a non-centered parameterization to aid in estimation by specifying mean, scale, and error distributions for each parameter (Betancourt & Girolami, 2015). Estimation for the group level was based on each timepoint.

Prior mean and variance distributions for all parameters are listed in Table S1a. Error distributions, which were estimated for each subject are also listed in Table S1a. Each subject's parameter consisted of a group estimated mean value plus the combined value of the group estimated scale value multiplied by the individually estimated error value. Group level estimates are listed in Table S1b.

Eight MCMC chains were run for each model, with 8000 samples per chain (6000 after discarding warm up samples). Chains from the RL models used in analyses were inspected for convergence

and showed good mixing, with all values of the potential scale reduction factor (Gelman & Rubin, 1992) less than 1.1.

Model and parameter recovery was performed for validation according to Palminteri et al. (Palminteri, Wyart, & Koechlin, 2017). Data were simulated based on the estimated distributions of each parameters from the participants' data. These simulated data were then fit to the model and the resulting recovered parameters were compared to the simulated parameters. All parameters were recoverable, with simulated differences in parameter values similar to recovered differences see Table S1c.

Model Comparison. In addition to the modified Kalman filter model detailed above, we fitted three other models for comparison. The other models tested were the original Kalman model (Daw et al., 2006), the standard Q-learning model (Sutton & Barto, 1998), and the volatile Kalman filter (Piray & Daw, 2019). The original Kalman model and the standard Q-learning model were only fit as a benchmark for our modified Kalman model, but they do not include a trial by trial risk prediction error calculation. The Volatile Kalman filter model does include a trial by trial prediction error calculation, so it was serves at the main model of comparison for our current model.

The models were compared using cumulative integrated Bayesian Information Criterion (iBIC; Huys et al., 2012):

$$\text{iBIC} = -2\sum_{i} log\frac{1}{K}\sum_{K=1}^{K} p(A_i|h^K) + |M|log\ (|A|)$$

where $|M|$ is the number of hyper parameters fitted (mean and variance of group distribution for each parameter), and $|A|$ is the total number of observed choices made by all participants, and K is the total number of samplers, and i is the participant index. This approach enables penalizing models with greater number of parameter and provides information about how well the model fits given the posterior distributions of each parameter. Model with the lowest iBIC is preferred.

Each model was fit across all subjects' behavioral data (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Schonberg, Daw, Joel, & O'Doherty, 2007; Schonberg et al., 2010) and the subsequent iBIC fits for the models are as follows: modified Kalman: 43267; volatile Kalman: 49602; Kalman: 43026; Q-learning: 49577.

The modified Kalman had the best fit among models including a risk prediction error, so it was chosen. The goodness of fit of the modified Kalman model is reflected on Figure 1d, which compares the calculated probabilities against percentage of time the choice was selected. While the original Kalman model had a the best fit overall, it did not include a quantification of risk prediction error, which was needed in order to examine the neural correlates of risk learning.

Imaging analysis

MRI collection. Participants were scanned on a 3T Siemens Tim Trio MR scanner. Echoplanar images were collected in 34 4-mm slices at a 30° hyperangulation from the anterior-posterior commisure (AC-PC) line (TR = 2000 ms, TE = 30 ms, flip angle = 90°, matrix = 64 x 64, voxel size = 3.4 x 3.4 x 4.0 mm3). A high resolution (1 mm3) anatomical Magnetization Prepared Rapid Gradient Echo (MPRAGE) T1 image (TR = 1200 ms, TE = 2.66 ms, flip angle = 12°) was collected to aid in registration.

Pre-processing. Pre-processing of the imaging data was completed using statistical parametric mapping software (SPM8; Wellcome Department of Imaging Neuroscience, University College London, UK). Images were first corrected temporally for slice timing, and then for movement using least squares minimization without higher-order corrections for spin history and normalized to stereotaxic MNI (Montreal Neurological Institute) space by calculating a multiplication matrix for segmented grey and white matter and CSF separately. Images were then resampled every 3.4 mm using 4th Degree B-spline interpolation and smoothed with a 6 mm Gaussian kernel.

First-level fMRI processing. The general linear model (GLM) and the theory of Gaussian random fields implemented in SPM8 (Friston et al., 1994) were used to perform statistical analysis on the

individual and group level. For the first level individual analysis, onset times for stimuli and outcomes of each trial were modeled as separate punctate events.

Trial-by-trial expected values were modeled as parametric regressors onto the response events. For outcome events, calculated risk prediction error and value prediction error based on individuals' estimated parameter were modeled as parametric regressors. The effects of the risk prediction error and value prediction error, respectively, were examined after accounting for effect of the other regressor. Effects due to run number, time in scanner, and head movement parameters were modeled out as nuisance covariates for each time point.

Within Group Whole Brain Analysis. To examine the effects of risk prediction error and value prediction error, within group second level contrast in SPM8 was used. Results were thresholded with a voxel level uncorrected $p < 0.001$, unless noted, and significant clusters were defined using a family-wise-error correction.

Correlation Analysis of Neural Risk Prediction Error and Individual Measures. To directly relate neural prediction error signal and risk prediction error signal to individual measures, individual beta values for risk prediction error and prediction error were first extracted from anatomical regions of interest (ROIs) including bilateral anterior insula (AAL library) and bilateral ventral striatum (Garrison et al., 2013). Anatomical masks of the ventral striatum were constructed using WFU-pickatlas (Maldjian, Laurienti, Kraft, & Burdette, 2003) including the structures of the caudate, putamen, and globus pallidus. Also included in the mask was the nucleus accumbens, which was defined in the meta-analysis of prediction errors by Garrison et al. (2013) as part of the striatum. Extracted individuals' ROI signals were then correlated with individual risk learning rate estimates ($1-\rho$).

Developmental trajectory analysis with hierarchical linear modeling

To determine the adolescents risk learning trajectory across time, individual risk learning rates were regressed onto the age of the individual using a Hierarchical Linear Model (HLM) with random effects of age and subject. To examine relationships among real-world risk taking behavior

and learning, Things I do (Conger & Elder, 1995) self-reported survey was collected for each time point. The survey asked adolescents whether they may have partaken in a list of established real-world risk-taking behaviors (see Supplement for complete list), which can be divided into subscales of Major and Minor Risks. Similarly, Major and Minor Risk trajectories were determined by regressing each respective subscale onto the age of the individual using a HLM with random effects of age and subject. Lastly, individual slopes for risk learning rate on age was correlated with the respective Major and Minor subscale slopes on age.

References

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336–372. https://doi.org/10.1016/j.geb.2004.06.010

Betancourt, M., & Girolami, M. (2015). Hamiltonian Monte Carlo for Hierarchical Models. In *Current Trends in Bayesian Methodology with Applications* (pp. 79–101). Chapman and Hall/CRC. https://doi.org/10.1201/b18502-5

Bjork, J. M., Smith, A. R., Chen, G., & Hommer, D. W. (2011). Psychosocial problems and recruitment of incentive neurocircuitry: Exploring individual differences in healthy adolescents. *Developmental Cognitive Neuroscience*, *1*(4), 570–577. https://doi.org/10.1016/j.dcn.2011.07.005.Psychosocial

Cai, W., Ryali, S., Chen, T., Li, C.-S. R., & Menon, V. (2014). Dissociable Roles of Right Inferior Frontal Cortex and Anterior Insula in Inhibitory Control: Evidence from Intrinsic and Task-Related Functional Parcellation, Connectivity, and Response Profile Analyses across Multiple Datasets. *Journal of Neuroscience*, *34*(44), 14652–14667. https://doi.org/10.1523/JNEUROSCI.3048-14.2014

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*, 111–126. https://doi.org/10.1196/annals.1440.010

Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and Psychological Maturation of Decision-making in Adolescence and Young Adulthood. *Journal of Cognitive Neuroscience*, *25*(11), 1807–1823. https://doi.org/10.1162/jocn_a_00447

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., & Poldrack, R. A. (2010). A unique adolescent response to reward prediction errors. *Nature Neuroscience*, *13*(6), 669–671. https://doi.org/10.1038/nn.2558

Conger, R. D., & Elder, G. H. (1995). Familes in Troubled Times: Adapting to Change in Rural America. HeinOnline.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*, *3*(8), 655–666. https://doi.org/10.1038/nrn894

Crone, E. A. (2009). Executive functions in adolescence: inferences from brain and behavior. *Developmental Science*, *12*(6), 825–830. https://doi.org/10.1111/j.1467-7687.2009.00918.x

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313

D'Acremont, M., Lu, Z. L., Li, X., Van der Linden, M., & Bechara, A. (2009). Neural correlates of risk prediction error during reinforcement learning in humans. *NeuroImage*, *47*(4), 1929–1939. https://doi.org/10.1016/j.neuroimage.2009.04.096

Davidow, J. Y., Foerde, K., Galván, A., & Shohamy, D. (2016). An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence. *Neuron*, *92*(1), 93–99. https://doi.org/10.1016/j.neuron.2016.08.031

Daw, N. D. (2011). Trial-by-trial Data Analysis Using Computational Models. In *Decision Making, Affect, and Learning* (pp. 3–38). Oxford, England: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199600434.003.0001

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879. https://doi.org/10.1038/nature04766

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In *Principles of Frontal Lobe Function* (pp. 466–503). New York, NY: Oxford University Press.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping*, *2*(4), 189–210. https://doi.org/10.1002/hbm.460020402

Galván, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *The Journal of Neuroscience*, *26*(25), 6885–6892. https://doi.org/10.1523/JNEUROSCI.1062-06.2006

Galván, A., & McGlennen, K. M. (2013). Enhanced striatal sensitivity to aversive reinforcement in adolescents versus adults. *Journal of Cognitive Neuroscience*, *25*(2), 284–296. https://doi.org/10.1162/jocn_a_00326

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction Error in Reinforcement Learning: A Meta-Analysis of Neuroimaging Studies. *Neuroscience & Biobehavioral Reviews*, *37*(7), 1297–1310. https://doi.org/10.1016/j.neubiorev.2013.03.023

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Hauser, T. U., Iannaccone, R., Walitza, S., Brandeis, D., & Brem, S. (2015). Cognitive flexibility in adolescence: Neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *NeuroImage*, *104*, 347–354. https://doi.org/10.1016/j.neuroimage.2014.09.018

Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural Signatures of Economic Preferences for Risk and Ambiguity. *Neuron*, *49*(5), 765–775. https://doi.org/10.1016/j.neuron.2006.01.024

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3), e1002410. https://doi.org/10.1371/journal.pcbi.1002410

Javadi, A. H., Schmidt, D. H. K., & Smolka, M. N. (2014). Differential representation of feedback and decision in adolescents and adults. *Neuropsychologia*, *56*(1), 280–288. https://doi.org/10.1016/j.neuropsychologia.2014.01.021

Javadi, A., Schmidt, D., & Smolka, M. (2014). Adolescents Adapt More Slowly than Adults to Varying Reward Contingencies. *Journal of Cognitive Neuroscience*, 2670–2681. https://doi.org/10.1162/jocn

Jocham, G., Klein, T. A., & Ullsperger, M. (2014). Differential Modulation of Reinforcement Learning by D2 Dopamine and NMDA Glutamate Receptor Antagonism. *Journal of Neuroscience*, *34*(39), 13151–13162. https://doi.org/10.1523/JNEUROSCI.0757-14.2014

Luciana, M., Wahlstrom, D., Porter, J. N., & Collins, P. F. (2012). Dopaminergic modulation of incentive motivation in adolescence: Age-related changes in signaling, individual differences, and implications for the development of self-regulation. *Developmental Psychology*, *48*(3), 844–861. https://doi.org/10.1037/a0027432

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An Automated Method

for Neuroanatomic and Cytoarchitectonic Atlas-Based Interrogation of fMRI Data Sets. *NeuroImage*, *19*(3), 1233–1239. https://doi.org/10.1016/S1053-8119(03)00169-1
Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458
Montague, P. R., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936–1947. Retrieved from http://www.jneurosci.org/content/16/5/1936.short
Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational Roles for Dopamine in Behavioural Control. *Nature*, *431*(7010), 760–767. https://doi.org/10.1038/nature03015
Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging Valuation Models in Human Choice. *Annual Review of Neuroscience*, *29*(1), 417–448. https://doi.org/10.1146/annurev.neuro.29.051605.112903
Palminteri, S., Wyart, V., & Koechlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. *Trends in Cognitive Sciences*, *xx*(6), 1–10. https://doi.org/10.1016/j.tics.2017.03.011
Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-Dependent Prediction Errors Underpin Reward-Seeking Behaviour in Humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051
Piray, P., & Daw, N. D. (2019). A transparent model for learning in volatile environments. *BioRxiv*, 701466. https://doi.org/10.1101/701466
Preuschoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. In *Annals of the New York Academy of Sciences* (Vol. 1104, pp. 135–146). https://doi.org/10.1196/annals.1390.005
Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *The Journal of Neuroscience*, *28*(11), 2745–2752. https://doi.org/10.1523/JNEUROSCI.4286-07.2008
Schonberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement Learning Signals in the Human Striatum Distinguish Learners from Nonlearners during Reward-Based Decision Making. *Journal of Neuroscience*, *27*(47), 12860–12867. https://doi.org/10.1523/JNEUROSCI.2496-07.2007
Schonberg, Tom, O'Doherty, J. P., Joel, D., Inzelberg, R., Segev, Y., & Daw, N. D. (2010). Selective Impairment of Prediction Error Signaling in Human Dorsolateral but not Ventral Striatum in Parkinson's Disease Patients: Evidence from a Model-Based fMRI Study. *NeuroImage*, *49*(1), 772–781. https://doi.org/10.1016/j.neuroimage.2009.08.011
Silveri, M. M., Tzilos, G. K., Pimentel, P. J., Yurgelun-Todd, D. A., Pimental, P. J., & Yurgelun-Todd, D. A. (2004). Trajectories of Adolescent Emotional and Cognitive Development: Effects of Sex and Risk for Drug Use. *Annals of the New York Academy of Sciences*, *1021*(Adolescent Brain Development: Vulnerabilities and Opportunities), 363–370. https://doi.org/10.1196/annals.1308.046
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge Univ Press.
Tymula, A., Rosenberg Belmaker, L. A., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences*, *109*(42), 17135–17140. https://doi.org/10.1073/pnas.1207144109

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*(6), 1247–1255. https://doi.org/10.1093/cercor/bhr198

van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *Proceedings of the National Academy of Sciences*, *112*(29), E3765–E3774. https://doi.org/10.1073/pnas.1423095112

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. https://doi.org/10.1016/j.neuroimage.2015.04.069

Wahlstrom, D., White, T., & Luciana, M. (2011). Neurobehavioral evidence for changes in dopamine system activity during adolescence. *Neuroscience & Biobehavioral Reviews*, *34*(5), 631–648. https://doi.org/10.1016/j.neubiorev.2009.12.007.Neurobehavioral

Wang, Y., Ma, N., He, X., Li, N., Wei, Z., Yang, L., … Zhang, X. (2017). Neural substrates of updating the prediction through prediction error during decision making. *NeuroImage*, *157*(May), 1–12. https://doi.org/10.1016/j.neuroimage.2017.05.041

Weber, E. U., & Johnson, E. J. (2009). Decisions Under Uncertainty: Psychological, Economic, and Neuroeconomic Explanations of Risk Preference. In *Neuroeconomics* (pp. 127–144). Elsevier. https://doi.org/10.1016/B978-0-12-374176-9.00010-5

Wood, S. M. W., & Bechara, A. (2014). The neuroscience of dual (and triple) systems in decision making. In V. F. Reyna & V. Zayas (Eds.), *The Neuroscience of Risky Decision Making* (pp. 177–202). https://doi.org/http://dx.doi.org/10.1037/14322-008

Figures

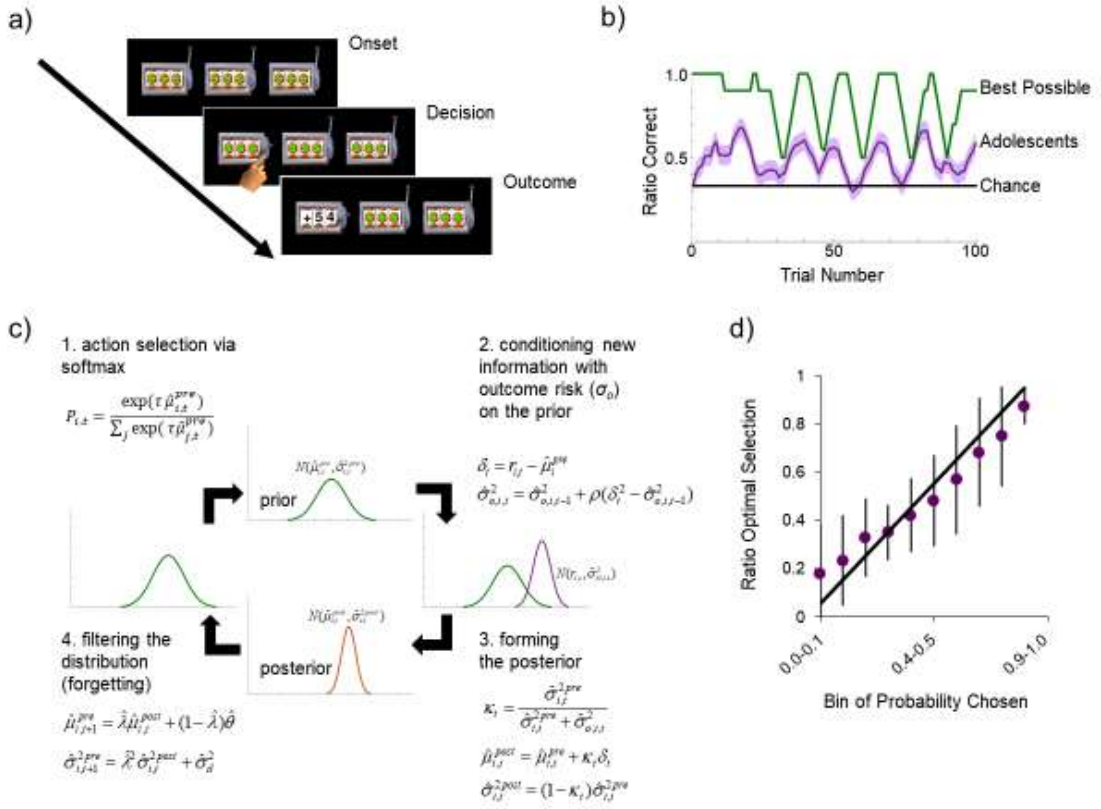

Figure 1. **Experimental design and behavioral model**. a) Participants reinforcement learning task with independent and randomly varying outcomes that drifted with a random walk while fMRI scanning collected their neural data. The task entailed learning from repeated choices in a 3-arm bandit over 2 runs with 100 trials in each run. Each trial was self-paced and participants observed the outcome of their chosen option before moving on to the next trial. Adolescents were compensated dependent on their performance on 5 random trials. b) Adolescents actual behavior from one specific pseudo-random block compared against chance and a moving window showing the trend a best possible choice in a moving window of 10 trials. Adolescents performed well above chance, showing they were learning as the outcomes shifted. In order for the adolescent to perform well, they needed to learn the variance associated with each slot. c) Participants' behavioral choices were fit to a Kalman filter reinforcement learning model (Daw et al., 2006) which included a risk update parameter (1-ρ) that represented the amount of past risk associated with each choice retained. d) The model predicted probability of selecting the better option was a good fit with adolescents' actual behavior (% better option selected across subjects).

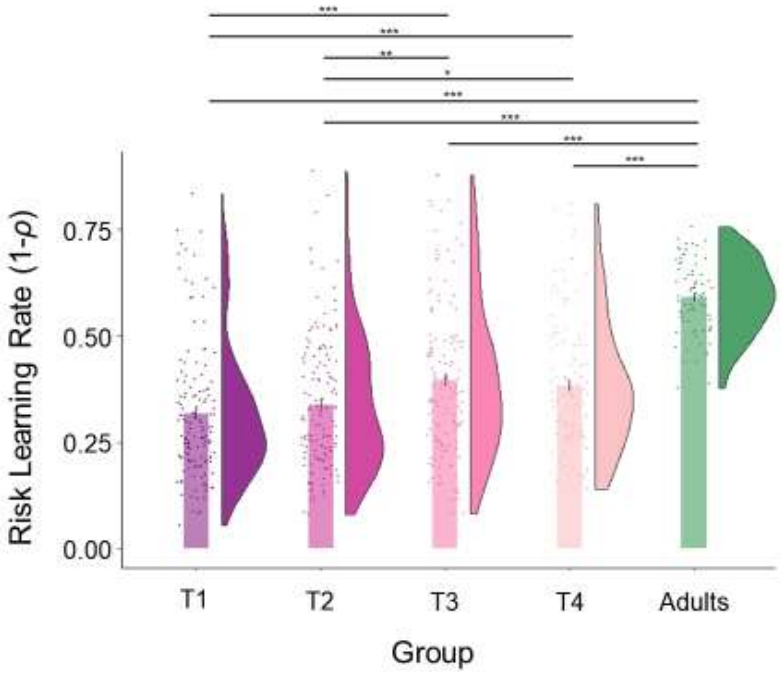

Figure 2. **Risk Learning Rate Increases with Time During Adolescence**. Adolescents' individually estimated risk learning rate (1-ρ) regressed onto age using a Hierarchical Linear Model with random effects of age and subject showed a significant group effect with higher risk learning rate over time ($\beta = 0.177$, $t(479) = 4.30$ $p < 0.001$). This meant adolescents improved tracking past risks as they developed. Estimated risk learning rate in adults were closer to estimates of adolescents when they were older.

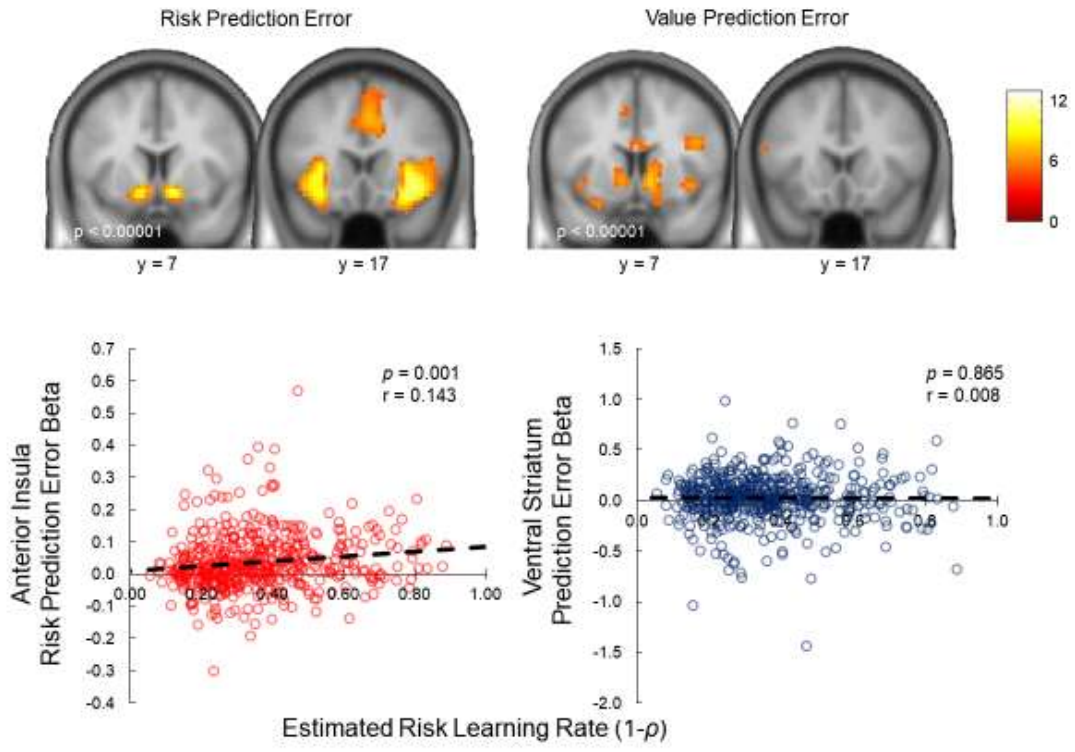

Figure 3. **Risk prediction error localized in the Anterior Insula and Ventral Striatum**. a) The anterior insula and ventral striatum were found to track risk prediction errors. b) The ventral striatum also tracked the value prediction errors. c) There was a significant positive effect of adolescents' individually estimated risk learning rate (1-ρ) on their anterior insula risk prediction error signal ($\beta = 0.142$, $t(470) = 3.20$, $p = 0.001$) using HLM regression with random effects of age and subject, but not the ventral striatum value prediction error signal ($\beta = 0.001$, $t(159) = 0.01$, $p = 0.990$).

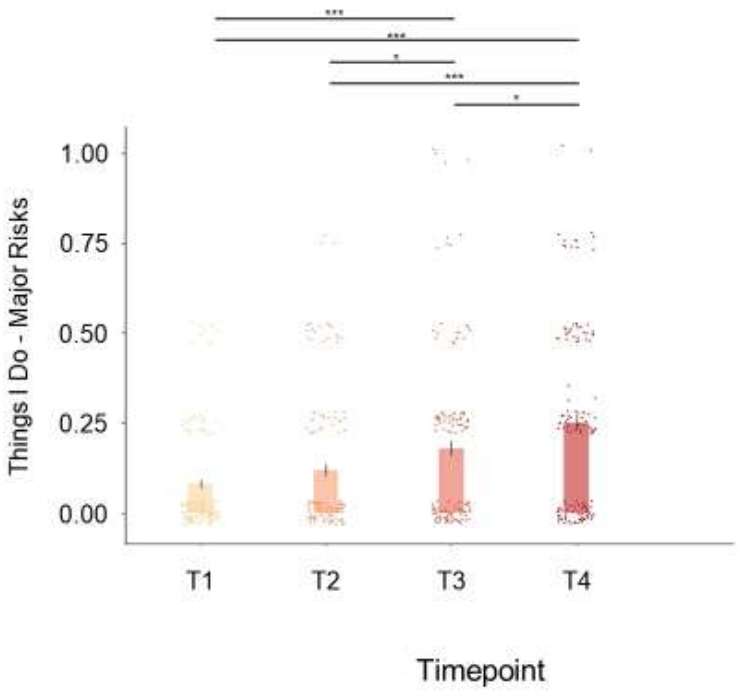

Figure 4. **Adolescents Displayed Increased Real-World Risk-Taking Behavior Over Time.** Major Risk subscale score of the self-report survey Things I Do (Conger & Elder, 1995) was regressed onto age using a Hierarchical Linear Model with random effects of age and subject showing a significant group effect ($\beta = 0.319$, $t(137) = 8.24$ $p < 0.001$). Adolescents displayed more real-world risk-taking behavior over time.

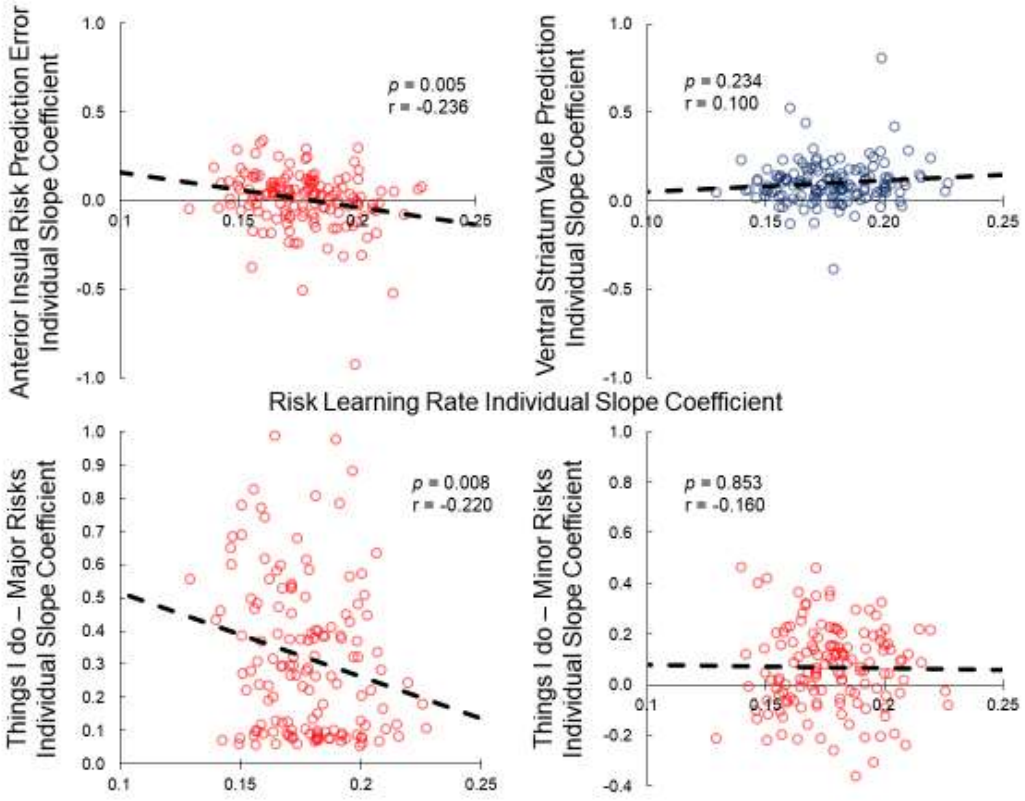

Figure 5. **Individual Risk Learning Developmental Trajectories Negative Correlated with Anterior Insula Risk Prediction Error Signal Trajectory and Major Risk Subscale Trajectory.** a) Risk Learning trajectory was found to correlate negatively with Anterior Insula risk prediction error trajectory (r = -0.236, p = 0.005), meaning adolescents who improved at tracking past risks showed an actual decrease in Anterior Insula risk prediction error signal. b) There was no relationship between the trajectories of risk learning rate and Ventral Striatum value prediction error signals however (r = 0.100, p = 0.234). c) Risk learning rate trajectory was negatively correlated (r = -0.220, p = 0.005) with the developmental trajectory of the Major Risk subscale of Things I Do self-report survey (Conger & Elder, 1995). Adolescents who improved in tracking past risks exhibited less real-world risk-taking behavior.

Tables

a)

Parameter	Prior Mean Distribution	a	b	Subject Level Random Effects	Prior Standard Deviation Distribution	a	b	Parameter Uncentered	Parameter Transformed
ρ	normal	0	1	yes	cauchy	0	1	yes	yes
τ	normal	5	10	yes	cauchy	0	5	no	no
λ	normal	0.5	0.3	yes	cauchy	0	0.3	no	no
θ	normal	5	3	yes	cauchy	0	3	no	no
σ_θ	normal	7.69	1	no	NA	NA	NA	NA	NA

b)

Parameter	T1	T2	T3	T4
ρ	0.684 (0.158)	0.663 (0.159)	0.604 (0.182)	0.618 (0.154)
τ	3.601 (1.568)	4.238 (1.626)	4.692 (2.022)	4.525 (2.120)
λ	0.462 (0.129)	0.454 (0.115)	0.499 (0.140)	0.521 (0.138)
θ	5.910 (2.309)	5.515 (1.679)	5.442 (1.935)	5.355 (1.590)
σ_θ	10.705	10.534	8.267	9.024

c)

Parameter	r	p
ρ	0.560	<.001
τ	0.793	<.001
λ	0.728	<.001
θ	0.935	<.001

Table 1. Model Parameters. a) Prior distribution for estimations. b) Estimations for each year of data collection. c) Parameter recovery correlations.

Risk Prediction Error	Cluster Size	Cluster *p* FWE	Peak T	Voxel p FDR	x	y	z	Region
	868	0	11.27	0	34	21	-2	R Anterior Insula
			10.72	0	31	17	-13	R Anterior Insula
			7.64	0	10	7	-2	R Ventral Striatum
	335	0	11.14	0	-3	-30	25	Posterior Cingulate
			5.81	0.014	4	4	25	Middle Cingulate
	209	0	10.07	0	-37	14	-9	L Anterior Insula
			9.71	0	-34	17	-2	L Anterior Insula
			9.35	0	-30	17	-13	L Anterior Insula
	771	0	9.9	0	0	38	15	Dorsal Anterior Cingulate
			8.22	0	4	24	35	Dorsal Anterior Cingulate
			6.33	0.002	7	14	59	R Supplemental Motor Area
	449	0	9.8	0	34	-85	-6	R Inferior Occipital Lobe
			7.29	0	44	-58	-16	R Inferior Temporal Lobe
			6.73	0	34	-51	-23	R Fusiform Gyrus
	529	0	9.38	0	14	-64	35	R Precuneus
			8	0	34	-61	38	R Angular Gyrus
			7.86	0	-7	-68	35	L Precuneus
	378	0	8.11	0	-24	-95	-6	L Inferior Occipital Lobe
			7.61	0	-34	-85	-9	L Inferior Occipital Lobe
			6.58	0.001	-37	-64	-13	L Fusiform Gyrus
	181	0	7.64	0	-30	-64	38	L Middle Occipital Lobe
			6.42	0.001	-41	-54	45	L Inferior Parietal Lobe
			5.63	0.047	-47	-41	45	L Inferior Parietal Lobe
	37	0	6.65	0.001	44	7	25	R Frontal Inferior Operculum
	24	0	6.49	0.001	-20	-37	-2	L Hippocampus
	21	0	5.29	0.223	38	34	15	R Dorsolateral Prefrontal Cortex
			5.25	0.223	44	27	21	R Dorsolateral Prefrontal Cortex

Table 2. Risk prediction error whole brain activation table.

Value Prediction Error	Cluster Size	Cluster p FWE	Peak T	Voxel p FDR	x	y	z	Region
	41	0	10.14	0	10	7	-9	R Ventral Striatum
	366	0	9.02	0	-44	-64	-2	L Middle Temporal Lobe
			8.04	0	-30	-41	-13	L Fusiform Gyrus
			6.54	0.001	-47	-47	-13	L Inferior Temporal Lobe
	44	0	9.01	0	-10	7	-9	L Ventral Striatum
	130	0	7.49	0	-7	-41	38	Middle Cingulum Gyrus
			6.83	0	-7	-54	15	L Precuneus Gyrus
	157	0	7.2	0	-7	41	-13	Ventromedial Prefrontal Cortex
			6.01	0.014	-3	31	1	Ventral Anterior Cingulate
			5.4	0.119	7	38	-13	Ventromedial Prefrontal Cortex
	122	0	7.14	0	31	-37	-16	R Fusiform Gyrus
			6.92	0	48	-61	-9	R Inferior Temporal Gyrus
			5.23	0.223	44	-47	-13	R Inferior Temporal Gyrus
	57	0	6.66	0	68	-20	4	R Superior Temporal Gyrus
			5.8	0.014	68	-10	-2	R Superior Temporal Gyrus
	118	0	6.5	0.001	-47	34	-2	L Inferior Prefrontal Cortex
			6.28	0.002	-37	38	-13	L Inferior Prefrontal Cortex
			5.91	0.014	-24	31	-19	L Inferior Prefrontal Cortex
	125	0	6.41	0.001	-47	-61	21	L Middle Temporal Lobe
			6.35	0.002	-51	-68	25	L Angular Gyrus
			5.48	0.119	-41	-71	32	Middle Cingulum Gyrus
	30	0	5.93	0.014	-17	31	45	L Middle Occipital Gyrus
	22	0	5.22	0.223	51	-7	15	R Precentral Gyrus
			5.2	0.223	55	-7	25	R Precentral Gyrus
			5.09	0.599	65	-3	25	R Postcentral Gyrus

Table 3. Value prediction error whole brain activation chart.

Supplemental Materials

Standard Q-learning model. In the standard model, the initial expected values $Q_{i,0}$ for the possible choices i were set to 50. For trial number t, the outcome was represented by r_t with the expected value represented by $Q_{i,t}$. The prediction error δ_t, which measures the difference in outcome r_t and expectation $Q_{i,t}$, for a trial was defined as the following:

$$\delta_t = r_t - Q_{i,t}$$

For the standard Q-learning model algorithm, the model-based parameter estimated was learning rate α, which quantifies how much weight the prediction error δ_t from current trials is given in updating the following trials' expected value $Q_{a,t+1}$. The standard Q-learning model was the null hypothesis model. Each trial by trial expected value Q for option i was calculated as follows:

$$Q_{i,t+1} = Q_{i,t} + \alpha * \delta_t$$

A softmax action selection function was used for the choice rule. The probability of selecting choice a at time t was estimated as follows:

$$P_{i,t} = \frac{\exp(\beta Q_{i,t})}{\sum_j \exp(\beta Q_{j,t})}$$

Inverse temperature β is an exploration parameter that quantifies the balance between the exploitation of the higher valued option and exploration of the other option for information. Lower inverse temperatures represent greater exploration, and is also consistent with increased random choice behavior.

Risk Learning Rate correlated with Neural Risk Prediction Error Signal. To address the question of whether neural correlated of value and risk learning signal are related to the risk learning rate, a second level analysis was performed. Anatomical regions of interest (ROIs) of the bilateral anterior insula (aal library) and the bilateral ventral striatum (Garrison et al., 2013) were used to

extract individual beta values for risk prediction error and value prediction error. Individuals' ROI signals were then correlated with individual risk update parameter estimates. Bilateral anterior insula risk prediction error signal was significantly correlated with the risk update parameter (Fig. 2c; $r = 0.208$, $p = 0.012$; Right Anterior Insula: $r = 0.228$, $p = 0.006$; Left Anterior Insula: $r = 0.179$, $p = 0.030$) such that adolescents with greater anterior insula signal kept track of past risk information to a greater extent. Ventral striatum signal for risk prediction error also correlated with risk learning rate ($r = 0.163$, $p = 0.049$), but ventral striatum signal for value prediction error did not (Fig 2c; $r = 0.032$, $p = 0.720$).

Model Agnostic Analysis. Adolescents selected the most optimal choice on 51.8% of the trials (s.e. 1.2%) on the task. This was well above the chance performance of the 33.3% (see Fig 1b for sample subject performance; t_{145}=15.55, p < 2.2*10^-16).

Adult Comparison. 67 adult individuals (58 females; mean age 43.40, s.d. 6.71) were recruited from the parents of the adolescents as an adult comparison sample. Similar to the adolescents, a whole brain analysis showed anterior insular, ventral striatum, and dorsal medial prefrontal cortex BOLD signal to be correlated with calculated risk prediction error (Figure S5a; Table S1b). Meanwhile, the ventral striatal BOLD signal correlated with calculated value prediction error signal (Figure S5b, Table S1c). In a between group analysis, no differences were observed between adolescents and adult neural encoding or risk or value prediction error signals. (Figure S5c).

Appendix A: List of Items for Things I Do self-report Risk Taking survey

1. Ridden in a car without a seatbelt *
2. Ridden on a bike without a helmet *
3. Done something dangerous on a dare *
4. Carried a weapon somewhere
5. Threatened to beat up someone to make them do something +
6. Taken part in a gang fight
7. Skipped school without permission +

8. Had a fist fight with another person *

9. Purposely set a fire in a building or in any other place

10. Hurt an animal on purpose

11. Smoked a cigarette or used tobacco +

12. Drunk a bottle or glass of beer or other alcohol (not counting having a drink with the family on holidays or religious occasions) +

13. Used or smoked marijuana, grass, pot, weed

14. Taken or stolen something not yours worth a lot, like a video game

15. Taken or stolen something not yours worth little, like candy

16. Gotten into someplace like a movie or game without paying *

17. Run away from home *

18. Broken into a building to take or steal something

+ denotes major risk subscale item

* denotes minor risk subscale item

Supplemental Figures

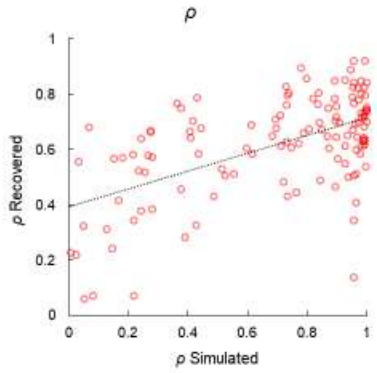

Figure S1. Model recovery for risk learning rate (ρ). The recovered parameter was significant correlated data simulated from a set of real subject estimations with noise.

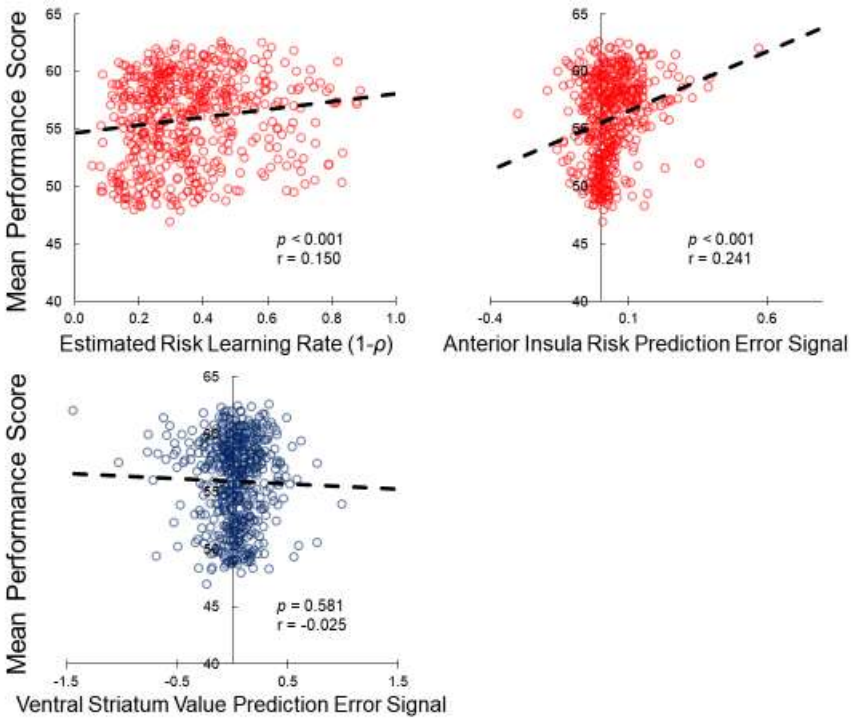

Figure S2. Correlation with performance during the continuous learning task. a) Estimated risk learning rate was significantly positively correlated with mean performance score in the task. b) Anterior insula risk prediction error signal was significantly positively correlated with mean performance score. c) Ventral striatum value prediction error signal was not correlated with mean performance score.

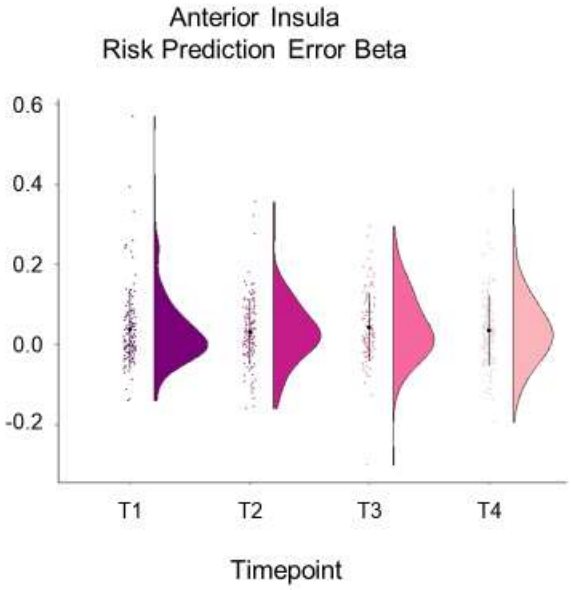

Figure S3. Anterior insula risk prediction error signal does not change over time.

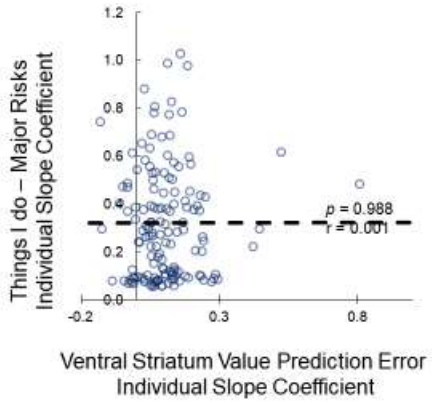

Figure S4. Ventral striatum value prediction error trajectory does not correlate with Major Risk trajectory.

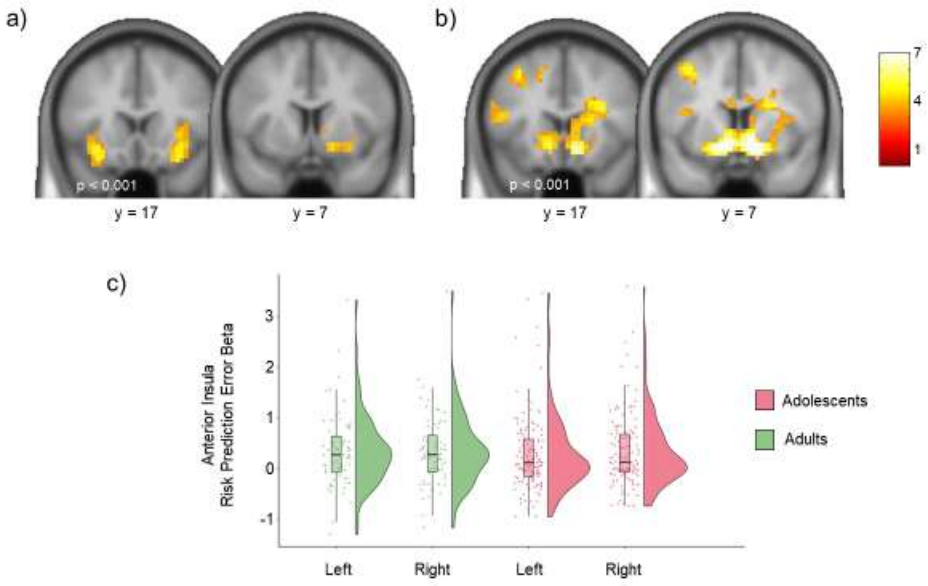

Figure S5. Adult risk and value prediction error signals. a) Risk prediction error signal found in the anterior insula, ventral striatum, and dorsal ACC in adults, similar to adolescents. b) Value prediction error signal found in the ventral striatum and dorsal ACC in adults. c) There is no different between adults and adolescents from their first time point in their anterior insula risk prediction error signal.

Supplemental Table

Parameter	Mean	Standard Error
ρ	0.579	0.009
β	4.108	0.155
λ	0.555	0.012
θ	4.891	0.102
σ_a	10.465	NA

Risk Prediction Error	Cluster Size	Cluster p FWE	Z-score	Voxel p FDR	x	y	z	region
	159	0	5	0.1	31	21	-16	R Anterior Insula
			4.11	0.365	17	7	-9	R Striatum
			3.93	0.446	31	21	-2	R Anterior Insula
	76	0.002	4.89	0.1	-30	17	-13	L Anterior Insula
			4.19	0.365	-37	10	-8	L Anterior Insula
			3.78	0.498	-30	27	-2	L Anterior Insula
	159	0	4.45	0.36	0	44	35	Dorsal Medial Prefrontal Crotex
			4.36	0.36	7	61	28	Dorsal Medial Prefrontal Crotex
			4.3	0.365	-7	48	45	Dorsal Medial Prefrontal Crotex
	49	0.017	4.35	0.36	-37	-51	-19	L Fusiform Gyrus
			3.67	0.491	-34	-61	-13	L Fusiform Gyrus
			3.65	0.497	-30	-54	-16	L Fusiform Gyrus
	45	0.025	4.08	0.365	4	-30	-8	Brainstem
			3.69	0.491	-3	-30	4	Brainstem
	76	0.002	4.04	0.365	-3	-34	25	Posterior Cingulate
			3.69	0.491	0	-24	25	Middle Cingulate
			3.51	0.598	0	-24	38	Middle Cingulate
	59	0.007	3.97	0.434	31	-68	-9	R Fusiform Gyrus
			3.78	0.498	48	-61	-16	R Inferior Temporal Lobe
			3.69	0.491	31	-51	-26	R Cerebellum

Value Prediction Error	Cluster Size	Cluster p FWE	Z-score	Voxel p FDR	x	y	z	region
	6962	0	Inf	0	14	7	-13	R Striatum
			7.3	0	-10	7	-9	L Striatum
			6.32	0	-20	38	42	Dorsal Medial Prefrontal Crotex
	154	0	5.99	0	38	-68	-40	R Cerebelum
			3.75	0.223	17	-81	-40	R Cerebelum
	69	0.004	4.28	0.047	51	-64	18	R Middle Temporal Lobe
			3.95	0.223	48	-58	25	R Superior Temporal Lobe
			3.48	0.599	38	-54	25	R Angular Gyrus

Table S1. Adult model estimation and whole brain activation charts. a) Estimated parameters for adults. b) Risk prediction whole brain error activation table for adults. c) Value prediction whole brain error activation table for adults.

General Discussion

Summary of studies

The goal of these papers was to examine underlying mechanisms that contribute to adolescents exhibiting real-world risk-taking behavior. The ability to distinguish such behaviors may inform educational policies on a broader scale and individual strategies of intervention and prevention.

In Paper 1, we reviewed work relating real-world risk-taking behavior with existing literature in adolescent development and risky decision making. In doing so, we offered an alternative paradigm for understanding the mechanism of real-world risk-taking behavior. Specifically, we first establish the necessity, preference (Christakou et al., 2013), and benefit of exploration for adolescents to learn about their environment (Decker, Lourenco, Doll, & Hartley, 2015; Oliva, Keyes, Iacono, & McGue, 2012; Ramstetter, Murray, & Garner, 2010; Vives & FeldmanHall, 2018). At the same time, such necessary exploration will inherently lead to real-world risk-taking behavior. These risk-taking behaviors are a form of sensation seeking and adaptive (Romer et al., 2017), however. Lastly, we emphasize the importance of risk learning to shift adolescents from decisions based on ambiguity to decisions based on expected risk. The inability to learn about risks would leave an adolescent in a continuous state of exploration and, subsequently, a continued state of taking real-world risky behaviors.

In Study 2, we used computational modeling to quantify adolescents' ability to track past risks in the form of a risk learning rate parameter. Such parameter was directly related to task performance, as in better tracking of past risk leading to better performance. Examining the fMRI signals, there was a significant effect of risk learning rate on anterior insula risk prediction error signal across all subjects and time points, meaning adolescents with better tracking of past risks showed greater anterior insula signal to risk prediction error; of note, there was no effect on ventral striatum value prediction error signal. Such results may indicate individuals possessing a more active interoceptive response would exhibit better learning of risks and subsequently better learning of values. There was a significant linear effect of age on the risk learning rate parameter, meaning adolescents improved their risk tracking as they aged. Individual trajectories of risk learning rate were negatively correlated with real-world risk-taking behavior trajectory, that is

adolescents who showed diminished improvements to their risk learning abilities ended up partaking in more real-world risk-taking behavior.

Overall, our current findings are consistent with the dual process theory of adolescent development (Casey et al., 2008; Shulman et al., 2016; Steinberg, 2008) comprising the early developing reward limbic system and the lagging executive functioning and control prefrontal system (Bechara & Damasio, 2005). Specifically, we found an increase in ventral striatal value prediction error signal over time. Increase in ventral striatal dopamine tract activity during adolescence has been associated with increased reward sensitivity (Galván et al., 2006; Galván & McGlennen, 2013), impulsivity (van den Bos, Rodriguez, Schweitzer, & McClure, 2015), and emotion (Hare et al., 2008). We did not find a developmental change in the prefrontal cortex. However, the dorsal anterior cingulate and right dorsolateral prefrontal cortex signal positively correlated with risk prediction error signal. Additionally, the anterior insula was also positively correlated with risk prediction error signal. Higher anterior insular risk prediction error signal was associated with better retention of past risk in the learning process. The anterior insula, here, may act as a monitoring signal, detecting salient events such as errors (Cai et al., 2014) and translating homeostatic states into conscious feelings (Wood & Bechara, 2014), that can engage individuals' executive function systems. This is supported by previous research showing the insula anatomically and functionally connects with the prefrontal cortex (Cai, Ryali, Chen, Li, & Menon, 2014; Craig, 2002; van Duijvenvoorde et al., 2016). While the anterior insula is not specified as a central developmental component in the dual process theory, the anterior insula may in fact play a prominent role in human decision-making and the risk-taking behaviors observed during adolescence (Wood & Bechara, 2014).

Taken together, these findings indicate that risk learning is crucial for adolescents in preventing continued real-world risk-taking behavior. If adolescents fail to learn the risks associated with their choices, they may have difficulties overcoming their propensity to sensation seek (Burnett et al., 2010; Steinberg, 2008) and exhibit continued exploration of choices that remain ambiguous, leading to continued real-world risk-taking behavior.

Limitations and future directions

While the current findings provide for greater understanding of the mechanisms of real-world risk-taking behavior in adolescents, future work can further delineate the effects of risk

learning on real-world risk-taking behaviors. First, the current sample was collected from a typical developing population of adolescents. As such, our data may not be generalizable to high-risk individuals with high levels of real-world risk-taking behavior. Future studies should examine high-risk adolescents to identify whether the same mechanisms are observed. Such work would provide insight for intervention and treatment. For example, incorporating components of cognitive behavioral therapy by encouraging the acknowledgement of one's own mental states and learning for them.

Second, the present study assesses adolescents using a model-free reinforcement learning paradigm. While we were successful in identifying and quantifying learning mechanisms that play a vital role in adolescents' real-world risk-taking behaviors, the current task and model does not incorporate and test all components of learning. Our current task does not include other circumstances of learning that an adolescent may encounter in the real world. For instance, adolescents do not learn in a vacuum and can learn socially from peers or family members. Previous research in real-world risk-taking behaviors have found peers to have a significant influence on adolescents' behaviors (Gardner & Steinberg, 2005; Jaccard et al., 2005). Similarly parents present at the time of a decision may inhibit risk-taking behavior in adolescence (Telzer et al., 2015). Additionally, adolescents are instructed with many rules and axioms regarding which behaviors are beneficial to them and which are detrimental. How adolescents learn under these circumstances may provide additional insight into the mechanisms within adolescents during a decision. Existing studies have found adolescents and children will utilize experience gained through exploration over explicit instructions to a higher degree in adults (Decker et al., 2015); the mechanisms involved and how they relate to real-world risk-taking behavior remains unknown. Furthermore model-based learning paradigms examining the balance of prior expectations (Huys et al., 2015), two-step learning (Gläscher, Daw, Dayan, & O'Doherty, 2010; Kool, Cushman, & Gershman, 2016), and decision tree pruning (Huys et al., 2012) can provide insight into how executive function and control influence adolescents' real-world risk-taking behaviors.

Lastly, the current study theorizes deficits in risk learning leading to downstream ambiguous choices and continued exploration of those choices (Christakou et al., 2013) exhibited as real-world risk-taking behavior. Such a hypothesis cannot be tested with the current paradigm.

Future studies combining ambiguous preference and risk learning components are needed to address this hypothesis.

Conclusion

The current research utilized a reinforcement learning paradigm to examine the underlying mechanisms involved in real-world risk-taking behavior exhibited by adolescents. Specifically, adolescents improved their abilities to track past risks associated with choices as they aged, quantified by an individually estimated risk learning rate. Better tracking of past risks had a positive effect on both task performance and anterior insula risk prediction error signal. Lastly, adolescents who showed diminished improvements in tracking past risks over time exhibited an increase in real-world risk-taking behavior. This result supports the hypothesis that individuals who fail to learn the risks associated with their choices will continue to make decision on ambiguous choices with the only recourse of continued exploration. Such continued exploration only leads to unknowingly choosing the risky behavior and is maladaptive.

References

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307. https://doi.org/10.1093/cercor/10.3.295

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Bjork, J. M., Smith, A. R., Chen, G., & Hommer, D. W. (2011). Psychosocial problems and recruitment of incentive neurocircuitry: Exploring individual differences in healthy adolescents. *Developmental Cognitive Neuroscience*, *1*(4), 570–577. https://doi.org/10.1016/j.dcn.2011.07.005.Psychosocial

Blakemore, S., & Robbins, T. W. (2012). Decision-making in the adolescent brain. *Nature Neuroscience*, *15*(9), 1184–1191. https://doi.org/10.1038/nn.3177

Braams, B. R., van Duijvenvoorde, a. C. K., Peper, J. S., & Crone, E. a. (2015). Longitudinal Changes in Adolescent Risk-Taking: A Comprehensive Study of Neural Responses to Rewards, Pubertal Development, and Risk-Taking Behavior. *Journal of Neuroscience*, *35*(18), 7226–7238. https://doi.org/10.1523/JNEUROSCI.4764-14.2015

Burnett, S., Bault, N., Coricelli, G., & Blakemore, S. J. (2010). Adolescents' heightened risk-seeking in a probabilistic gambling task. *Cognitive Development*, *25*(2), 183–196. https://doi.org/10.1016/j.cogdev.2009.11.003

Casey, B. J. (2015). Beyond Simple Models of Self-Control to Circuit-Based Accounts of Adolescent Behavior. *Annual Review of Psychology*, *66*(1), 295–319. https://doi.org/10.1146/annurev-psych-010814-015156

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*, 111–126. https://doi.org/10.1196/annals.1440.010

Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and Psychological Maturation of Decision-making in Adolescence and Young Adulthood. *Journal of Cognitive Neuroscience*, *25*(11), 1807–1823. https://doi.org/10.1162/jocn_a_00447

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., & Poldrack, R. A. (2010). A unique adolescent response to reward prediction errors. *Nature Neuroscience*, *13*(6), 669–671. https://doi.org/10.1038/nn.2558

Crone, E. A. (2009). Executive functions in adolescence: inferences from brain and behavior. *Developmental Science*, *12*(6), 825–830. https://doi.org/10.1111/j.1467-7687.2009.00918.x

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313

D'Acremont, M., Lu, Z. L., Li, X., Van der Linden, M., & Bechara, A. (2009). Neural correlates of risk prediction error during reinforcement learning in humans. *NeuroImage*, *47*(4), 1929–1939. https://doi.org/10.1016/j.neuroimage.2009.04.096

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264–1267. https://doi.org/10.1126/science.1150605

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

15(2), 310–320. https://doi.org/10.3758/s13415-014-0332-5
Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In *Principles of Frontal Lobe Function* (pp. 466–503). New York, NY: Oxford University Press.
Diamond, A. (2011). Biological and social influences on cognitive control processes dependent on prefrontal cortex. *Progress in Brain Research*, *189*(1), 319–339. https://doi.org/10.1016/B978-0-444-53884-0.00032-4
Epstein, L. G. (1999). Definition of Uncertainty Aversion. *The Review of Economic Studies*, *66*(3), 579–608.
Galván, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *The Journal of Neuroscience*, *26*(25), 6885–6892. https://doi.org/10.1523/JNEUROSCI.1062-06.2006
Galván, A., & McGlennen, K. M. (2013). Enhanced striatal sensitivity to aversive reinforcement in adolescents versus adults. *Journal of Cognitive Neuroscience*, *25*(2), 284–296. https://doi.org/10.1162/jocn_a_00326
Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental Psychology*, *41*(4), 625–635. https://doi.org/10.1037/0012-1649.41.4.625
Gee, D. G., Bath, K. G., Johnson, C. M., Meyer, H. C., Murty, V. P., van den Bos, W., & Hartley, C. A. (2018). Neurocognitive Development of Motivated Behavior: Dynamic Changes across Childhood and Adolescence. *The Journal of Neuroscience*, *38*(44), 9433–9445. https://doi.org/10.1523/JNEUROSCI.1674-18.2018
Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016
Hare, T., Tottenham, N., Galván, A., Voss, H. U., Glover, G. H., & Casey, B. J. (2008). Biological Substrates of Emotional Reactivity and Regulation in Adolescence During an Emotional Go-Nogo Task. *Biological Psychiatry*, *63*(10), 927–934. https://doi.org/10.1016/j.biopsych.2008.03.015
Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making. *Science*, *310*(5754), 1680–1683. https://doi.org/10.1126/science.1115327
Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural Signatures of Economic Preferences for Risk and Ambiguity. *Neuron*, *49*(5), 765–775. https://doi.org/10.1016/j.neuron.2006.01.024
Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3), e1002410. https://doi.org/10.1371/journal.pcbi.1002410
Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., … Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, 201414219. https://doi.org/10.1073/pnas.1414219112
Jaccard, J., Blanton, H., & Dodge, T. (2005). Peer influences on risk behavior: an analysis of the effects of a close friend. *Developmental Psychology*, *41*(1), 135–147. https://doi.org/10.1037/0012-1649.41.1.135

Kendler, K. S., Sundquist, K., Ohlsson, H., Palmér, K., Maes, H., Winkleby, M. A., & Sundquist, J. (2012). Genetic and Familial Environmental Influences on the Risk for Drug Abuse: A National Swedish Adoption Study. *Archives of General Psychiatry*, *69*(7), 690–697. https://doi.org/10.1001/archgenpsychiatry.2011.2112

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When Does Model-Based Control Pay Off? *PLOS Computational Biology*, *12*(8), e1005090. https://doi.org/10.1371/journal.pcbi.1005090

Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(42), 17951–17956. https://doi.org/10.1073/pnas.0905191106

Lejuez, C. W., Aklin, W. M., Zvolensky, M. J., & Pedulla, C. M. (2003). Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *Journal of Adolescence*, *26*(4), 475–479. https://doi.org/10.1016/S0140-1971(03)00036-8

Luciana, M., Wahlstrom, D., Porter, J. N., & Collins, P. F. (2012). Dopaminergic modulation of incentive motivation in adolescence: Age-related changes in signaling, individual differences, and implications for the development of self-regulation. *Developmental Psychology*, *48*(3), 844–861. https://doi.org/10.1037/a0027432

Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, *22*(3), 491–495. https://doi.org/10.1017/S0954579410000222

Montague, P. R., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936–1947. Retrieved from http://www.jneurosci.org/content/16/5/1936.short

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational Roles for Dopamine in Behavioural Control. *Nature*, *431*(7010), 760–767. https://doi.org/10.1038/nature03015

Oliva, E. M., Keyes, M., Iacono, W. G., & McGue, M. (2012). Adolescent Substance Use Groups: Antecedent and Concurrent Personality Differences in a Longitudinal Study. *Journal of Personality*, *80*(3), 769–793. https://doi.org/10.1111/j.1467-6494.2011.00746.x

Peper, J. S., & Dahl, R. E. (2013). The Teenage Brain. *Current Directions in Psychological Science*, *22*(2), 134–139. https://doi.org/10.1177/0963721412473755

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-Dependent Prediction Errors Underpin Reward-Seeking Behaviour in Humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Pharo, H., Sim, C., Graham, M., Gross, J., & Hayne, H. (2011). Risky business: Executive function, personality, and reckless behavior during adolescence and emerging adulthood. *Behavioral Neuroscience*, *125*(6), 970–978. https://doi.org/10.1037/a0025768

Preuschoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. In *Annals of the New York Academy of Sciences* (Vol. 1104, pp. 135–146). https://doi.org/10.1196/annals.1390.005

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *The Journal of Neuroscience*, *28*(11), 2745–2752. https://doi.org/10.1523/JNEUROSCI.4286-07.2008

Ramstetter, C. L., Murray, R., & Garner, A. S. (2010). The crucial role of recess in schools. *Journal of School Health*, *80*(11), 517–526. https://doi.org/10.1111/j.1746-1561.2010.00537.x

Robbins, R. N., & Bryan, A. (2004). Relationships Between Future Orientation, Impulsive Sensation Seeking, and Risk Behavior among Adjudicated Adolescents. *Journal of Adolescent Research*, *19*(4), 428–445. https://doi.org/10.1177/0743558403258860

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev*, *18*(3), 247–291. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8401595

Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex*, *18*(3), 652–663. https://doi.org/bhm097 [pii] 10.1093/cercor/bhm097

Romer, D., Reyna, V. F., & Satterthwaite, T. D. (2017). Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Developmental Cognitive Neuroscience*, *27*(July), 19–34. https://doi.org/10.1016/j.dcn.2017.07.007

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, *17*, 103–117. https://doi.org/10.1016/j.dcn.2015.12.010

Silveri, M. M., Tzilos, G. K., Pimentel, P. J., Yurgelun-Todd, D. A., Pimental, P. J., & Yurgelun-Todd, D. A. (2004). Trajectories of Adolescent Emotional and Cognitive Development: Effects of Sex and Risk for Drug Use. *Annals of the New York Academy of Sciences*, *1021*(Adolescent Brain Development: Vulnerabilities and Opportunities), 363–370. https://doi.org/10.1196/annals.1308.046

Sinha, J. W., Cnaan, R. a, & Gelles, R. W. (2007). Adolescent risk behaviors and religion : Findings from a national study. *Journal of Adolescence*, *30*(2), 231–249. https://doi.org/10.1016/j.adolescence.2006.02.005

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*(1), 78–106. https://doi.org/10.1016/j.dr.2007.08.002

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge Univ Press.

Telzer, E. H., Ichien, N. T., & Qu, Y. (2015). Mothers know best: redirecting adolescent reward sensitivity toward safe behavior during risk taking. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsv026

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*(6), 1247–1255. https://doi.org/10.1093/cercor/bhr198

van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *Proceedings of the National Academy of Sciences*, *112*(29), E3765–E3774. https://doi.org/10.1073/pnas.1423095112

Vives, M.-L., & FeldmanHall, O. (2018). Tolerance to ambiguous uncertainty predicts prosocial behavior. *Nature Communications*, *9*(1), 2156. https://doi.org/10.1038/s41467-018-04631-9

Wahlstrom, D., White, T., & Luciana, M. (2011). Neurobehavioral evidence for changes in dopamine system activity during adolescence. *Neuroscience & Biobehavioral Reviews*, *34*(5), 631–648. https://doi.org/10.1016/j.neubiorev.2009.12.007.Neurobehavioral

Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., … Ziesch., V. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature*, *512*(7513), 185–189. https://doi.org/10.1038/nature13402