HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2019 November 01.

Published in final edited form as:

Neuroimage. 2018 November 01; 181: 568–581. doi:10.1016/j.neuroimage.2018.06.058.

Novel insights from the Yellow Light Game: Safe and risky decisions differentially impact adolescent outcome-related brain function

Zdeňa A. Op de Macks[a,1], **Jessica E. Flannery**[a,b,1], **Shannon J. Peake**[a], **John C. Flournoy**[a,b], **Arian Mobasser**[a,b], **Sarah L. Alberti**[c], **Philip A. Fisher**[a,b], **Jennifer H. Pfeifer**[a,b,*]

[a]Prevention Science Institute, University of Oregon, Eugene, OR, USA

[b]Department of Psychology, University of Oregon, Eugene, OR, USA

[c]Oregon Health & Science University, Portland, OR, USA

Abstract

Changes across the span of adolescence in the adolescent reward system are thought to increase the tendency to take risks. While developmental differences in decision and outcome-related reward processes have been studied extensively, existing paradigms have largely neglected to measure how different types of decisions modulate reward-related outcome processes. We modified an existing decision-making paradigm (the Stoplight Task; Chein et al., 2011) to create a flexible laboratory measure of decision-making and outcome processing, including the ability to assess modulatory effects of safe versus risky decisions on reward-related outcome processes: the Yellow Light Game (YLG). We administered the YLG in the MRI scanner to 81 adolescents, ages 11–17 years, recruited from the community. Results showed that nucleus accumbens activation was enhanced for (1) risky > safe decisions, (2) positive > negative outcomes, and (3) outcomes following safe decisions compared to outcomes following risky decisions, regardless of whether these outcomes were positive or negative. Outcomes following risky decisions (compared to outcomes following safe decisions) were associated with enhanced activity in cortical midline structures. Furthermore, while there were no developmental differences in risk-taking behavior, more pubertally mature adolescents showed enhanced nucleus accumbens activation during positive > negative outcomes. These findings suggest that outcome processing is modulated by the types of decisions made by adolescents and highlight the importance of investigating processes involved in safe as well as risky decisions to better understand the adolescent tendency to take risks.

Keywords

fMRI; Risk taking; Adolescence; Reward; Ventral striatum; Driving game

[*]Corresponding author. Department of Psychology, 1227 University of Oregon, Eugene, OR, 97401, USA. jpfeifer@uoregon.edu (J.H. Pfeifer).

[1]Indicates joint first authorship.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2018.06.058.

Introduction

Adolescence is a period of optimal physical health, yet mortality rates are disproportionately high (Mulye et al., 2009; Patton et al., 2009). This "paradox" is thought to be the consequence of adolescents' heightened tendency to engage in various health-risking behaviors (Dahl, 2004). According to existing neurodevelopmental models, this heightened risk-taking tendency reflects normative changes in the adolescent brain (Casey et al., 2011; Crone and Dahl, 2012; Ernst, 2014; Steinberg, 2008), particularly in the reward system (Galván, 2010, 2014). These models have been tested using various laboratory paradigms that measure developmental changes in risk taking and reward processes thought to underlie risky decisions (reviewed in Defoe et al., 2015; Richards et al., 2013; and Van Duijvenvoorde et al., 2016). While these paradigms have provided insight into the neurocognitive fprocesses implicated in different stages of reward processing, such as during decision-making and/or outcome presentation (Richards et al., 2013), as of yet they have largely neglected to consider the associations between decision and outcome-related processes. Because outcomes often depend on decisions during real-life risk taking, characterizing how outcome-related processes may be impacted by different types of decisions is essential to our understanding of risk taking and associated reward processes.

Existing risk-taking and reward paradigms

Laboratory paradigms that have been used to assess reward processes commonly present participants with probabilistic rewards, although these paradigms differ in the action required to obtain the reward. Some paradigms require no action (i.e., passive reward tasks) or an action that has an objectively correct response (i.e., instrumental reward tasks), whereas other paradigms require a decision between alternatives associated with different expected values (i.e., reward decision-making tasks; for an overview of these different types of paradigms, see Richards et al., 2013). While knowledge from all three paradigm types can be applied to understanding the neural and psychological processes involved in adolescent risk taking, only reward decision-making tasks provide the opportunity to measure associations between decision and outcome processes.

There are two types of reward decision-making tasks that have been frequently employed in research involving adolescents: those that require participants to choose between two risky options with varying probabilities of obtaining a reward of a certain magnitude (i.e., the outcome is uncertain in both cases); and those that require participants to choose between a risky and a non-risky option (i.e., the outcome is certain in one case). Examples of each type of reward decision-making task that have been used in developmental neuroimaging studies are listed in Supplementary Table S1. Reward decision-making tasks that involve reward accrual, such as the Game of Chicken (Bjork et al., 2007), Columbia Card Task (CCT; Van Duijvenvoorde et al., 2015), and Balloon Analogue Risk Task (BART; Braams et al., 2015; Peper et al., 2018), were considered separately based on the premise that these kinds of tasks involve other decision-related processes (i.e., deciding when to stop risk taking instead of deciding when to take a risk; see Supplementary Table S1). Given the focus of this study on developmental differences in outcome-related brain activation, studies that reported only on

developmental differences in decision-related brain activation (Paulsen et al., 2012; Barkley-Levenson et al., 2013; Van Duijvenvoorde et al., 2015) are not further discussed.

Results from the studies using these reward decision-making paradigms described in Table S1 showed that adolescents generally tend to more strongly activate limbic regions - particularly the nucleus accum- bens - during positive outcome processing, especially when the laboratory paradigm involved deciding between two risky options (Braams et al., 2015; Ernst et al., 2005; Van Leijenhorst et al., 2010). However, one study reported an absence of developmental differences (May et al., 2004), another study reported decreases with pubertal maturation (Forbes et al., 2010), and yet another reported decreased limbic activation in adolescents compared to adults (Bjork et al., 2007). However, two of these three findings were from studies that used one particular paradigm (the Card Guessing Game), suggesting that these results might be driven by the paradigm used. The general pattern of increased limbic brain activation during adolescence also corresponded with an adolescent peak in risk-taking behavior, as measured using the BART, which was administered outside the MRI scanner (Braams et al., 2015; Peper et al., 2018). Taken together, the qualitative review suggests that adolescents frequently show heightened limbic activation during positive outcome processing during reward decision-making. This is in line with results from a recent meta-analysis on adolescent reward processing, which showed that ventral striatum (including nucleus accumbens) is more likely to be active in adolescents than in adults during positive > negative outcomes (Silverman et al., 2015).

Interestingly, developmental differences in outcome-related brain activation seem to be less consistently found when a safe alternative is present (Op de Macks et al., 2011, 2016; Van Duijvenvoorde et al., 2014), which suggests in part that outcome processes may be impacted by the types of decisions (risky or safe) that precede them. However, studies that used the tasks with a risky and a non-risky option have not been able to test this question, as safe decisions have a fixed outcome in these paradigms (e.g., in the Jackpot task, the decision to ‘pass’ is always followed by a neutral outcome; Op de Macks et al., 2011, 2016; Van Duijven-voorde et al., 2014). Thus, current reward decision-making tasks constrain our understanding of how risky versus safe decisions and outcome processing are associated, and potentially prevent our understanding of why different types of reward decision-making tasks obtain inconsistent results.

Introducing a novel risk-taking paradigm

One paradigm that allows for the comparison between outcomes following risky versus safe decisions is the Stoplight Task (Chein et al., 2011), a driving simulation during which participants are presented with probabilistic decisions with unknown outcomes that approximate real-life traffic scenarios. While the Stoplight Task has been used to measure risk taking and reward processing in adolescents and is suitable for the MRI scanner, only one study has used the Stoplight Task to assess developmental differences in brain activation. Importantly, within this study the researchers focused on results obtained by collapsing across the different types of decisions, as the focus of this study was on age-related differences in contextual effects on brain activity associated with decision-making (Chein et al., 2011). Two other studies that administered the Stoplight Task in adolescents

(14–17 years) reported on outcome-related brain activation *preceding* risky versus safe decisions (Kahn et al., 2015) and effects of social exclusion on risk-taking behavior and associated brain activation (Peake et al., 2013).

While together the aforementioned studies provide some information about the neural processes associated with decisions or outcomes during the Stoplight Task and the impact of different social contexts on these processes, differential influences of risky and safe decisions on outcome processes during adolescence have yet to be explored in this literature. Furthermore, in the default configuration of the Stoplight Task, the measure of risk taking (i.e., the proportion of decisions to continue through the intersection) is confounded with task performance (i.e., getting a faster time). This is because the choice to stop is associated with half as long a delay as crashing, but the probability of crashing is less than 50%. A faster time in the game can be obtained by choosing to 'go' rather than 'stop' at every intersection, giving 'go' decisions greater expected value than 'stop' decisions and making risk taking somewhat adaptive. Since adaptive and maladaptive risk taking might elicit different neural and psychological processes, there may be limitations or biases in what we can ascertain from prior studies using the Stoplight Task in its default configuration about the risky behaviors that contribute to adolescent mortality (e.g., health-risking behaviors).

The present study

For the present study, we adapted the Stoplight Task in an effort to address these and other limitations of prior work in this area. Our adapted paradigm, called the Yellow Light Game (YLG; https://dsn.uoregon.edu/research/yellow-light-game/), provides an opportunity to gain novel insights into the development of risky decision-making and associated neural and psychological processes during adolescence. The YLG allows assessment of how outcome processes vary according to the type of preceding decision (risky versus safe), which was a primary aim of this study. The YLG also allows the creation of conditions in which risky and safe decisions have unequal expected values (i.e., intersections vary based on crash probability), which not only more closely resembles real-life driving decisions, but also allows for the distinction between adaptive and maladaptive risk taking. Importantly, to prevent the task from promoting risk taking overall, in this configuration of the YLG the average crash probability across the task was set for this study at 50%. As such, there was no overall advantage to either taking risks or playing it safe. Finally, we administered the YLG in 11-to-17-year-olds to allow for exploration of developmental (i.e., age and puberty-related) effects in neural and behavioral patterns associated with decisions and outcome processing.

Our hypotheses for the current study were based on previous findings from (1) developmental neuroimaging studies that utilized reward decision-making tasks (see Supplementary Table S1), (2) studies that used the Stoplight Task (Chein et al., 2011) and its predecessor called Chicken (Gardner and Steinberg, 2005), as well as (3) a recent meta-analysis that investigated laboratory risk taking (Defoe et al., 2015). Specifically, we hypothesized that there would be no developmental differences in risk-taking behavior. However, we expected to find developmental differences in reward-related brain activation associated with the processing of outcomes, but not decisions, particularly in the nucleus

accumbens (NAcc; Braams et al., 2015; Silverman et al., 2015). In particular, we expected to find a linear age-related increase in NAcc activation during early-to mid-adolescence (Braams et al., 2015). Finally, given the idea that risk taking is motivated by rewards (Galvan, 2010, 2014; Van Duijvenvoorde et al., 2016), we hypothesized that the NAcc would more strongly differentiate between good versus bad outcomes following risky rather than safe decisions.

Materials and methods

Participants

For this study, we recruited adolescents aged 11–17 years from the community through flyers, outreach events, and online advertising. Adolescents were screened during a phone interview with their parent using the following exclusion criteria: (1) the presence of any known neurological disorders, such as seizure disorders, central nervous system infection (e.g. meningitis), brain tumors, muscular or myotonic dystrophy, and significant visual impairments, including color blindness; and (2) the presence of certain diagnosed psychiatric, developmental, or conduct disorders, such as autism spectrum disorders - including Asperger syndrome and pervasive developmental disorder, not otherwise specified (PDD-NOS) - schizophrenia, Tourette's syndrome, and obsessive compulsive disorder. Because the sample was intended to serve as a community control group to compare with a future sample of adolescents with a history of involvement in foster care and/or juvenile justice systems, we did not exclude for current or previous diagnoses of anxiety, depression, attention deficit (hyperactivity) disorder (ADD/ADHD), oppositional defiant disorder, or conduct disorder. Any medications that participants were taking for behavioral health concerns were documented, including their type and dose, as well as the frequency and duration of use. To determine eligibility for the MRI portion of the study, adolescents were additionally screened for MRI contraindications.

A total of 97 participants (51 females) enrolled in the study, of which 80.4% identified as Caucasian, 8.2% as mixed race, and 1% as African American (the remaining 10.4% did not report their ethnicity or race); 12.4% of enrolled participants identified as Hispanic. Furthermore, 22.4% of the parents who reported on their highest education level (N = 85) indicated they received a Bachelor's degree, 19.3% completed some college credit but no degree, 14.9% received a Master's degree, 13.0% completed high school or received a GED, 11.8% received an associate degree, 7.5% completed trade/technical/vocational training, 5.6% received a doctorate degree, 2.5% completed elementary school to 8th grade, and 1.2% completed a professional degree. As for reported household income (N = 82), 22.0% indicated an annual gross income between $20,000 and $40,000, 20.7% reported an income between $40,000 and $60,000, 19.5% reported an income between $80,000 and $100,000, 17.1% reported an income below $20,000,11.0% reported an income between $60,000 and $80,000, and 9.8% reported an annual income above $100,000. Reports on the mental health of the initial 97 enrollments demonstrated that three participants were diagnosed with anxiety disorders (social anxiety, generalized anxiety disorder, post-traumatic stress disorder) and two participants were diagnosed with ADD/ADHD; eight participants reported taking medications for mood-, attention-, anxiety-, and/or sleep-related symptoms.

Of the initial 97 enrollments, 88 participants (46 females) completed the Yellow Light Game (YLG). Reasons for dropout were: (1) the family never replied to schedule the second laboratory visit during which the task took place (n = 3); (2) the participant had braces and opted out of completing a behavioral rather than an MRI session (n = 3); (3) the participant did not show up (n = 1); (4) the family did not want to continue participation in the study (n = 1); and (5) unknown (n = 1). Of the 88 participants who completed the YLG, 81 participants (42 females) completed the task in the MRI scanner. Two additional participants (1 female) were excluded from analyses due to invalid imaging data (see Section 2.6 for details).

Cognitive functioning was measured based the Vocabulary and Matrix Reasoning subtests of Wechsler's Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), which provide a composite estimated full-scale intelligence quotient (FSIQ-2). Pubertal development was measured through self-report, using the Pubertal Development Scale (PDS; Petersen et al., 1988) See Table 1 for means, standard deviations, and ranges of age, IQ, and mean PDS score, calculated by sex, for the 79 participants who completed the task in the MRI scanner and had valid task-related imaging data. Note that while there were no sex differences in age or IQ, there was a sex difference in mean PDS score; on average, girls reported more advanced pubertal development compared to boys, consistent with the common observation that girls tend to mature sooner than boys (Shirtcliff et al., 2009). Furthermore, given the moderate to strong correlation between age and PDS score (for boys: $r = 0.7$, $t(34) = 5.7$, $p < 0.001$; for girls: $r = 0.62$, $t(39) = 4.9$, $p < 0.001$; and across all participants: $r = 0.61$, t (75) = 6.7, p < 0.001), we conducted separate analyses for each predictor.

Study procedure

Each participant visited the lab on two occasions. During the first visit, written informed consent for both visits was obtained from the parent. Written assent was obtained from the adolescent at each visit. After consent/assent, the adolescent was screened, in the presence of the parent, for eligibility for the MRI portion of the study using an MRI contraindications form provided by the Lewis Center for Neuroimaging at the University of Oregon. After MRI screening, the parent and the adolescent were seated in separate rooms where they filled out online questionnaires, which were administered through Qualtrics http://www.qualtrics.com/support/see Supplementary Table S2 for an overview of the questionnaires that were administered). The duration of the first visit was approximately 1.5 h and the adolescent was compensated $35. During the second visit, the adolescent completed more online questionnaires (see Supplementary Table S2). Upon completion of the questionnaires, the adolescent went into the mock scanner, where s/he practiced the fMRI task. After practicing the task, the adolescent was weighed, measured (height), and screened using a metal detector, before entering the MRI scanner. If the adolescent was not eligible for the MRI portion of the study, s/he completed the task in the same room that s/he completed the questionnaires in. After the scan/task completion, the adolescent could take a break during which snacks and drinks were provided. Subsequently, the adolescent completed a final set of online questionnaires and tasks. The experimenter concluded the visit by administering two subtests of the WASI (see Section 2.1) as well as a task experience interview, and debriefed both the adolescent and the parent. The duration of this

second visit ranged from 3 to 4 h and the adolescent was compensated $105. All materials and procedures were approved by the Institutional Review Board at the University of Oregon.

fMRI paradigm

Risk taking was assessed using the Yellow Light Game (YLG; Fig. 1), a computerized driving simulation that was adapted from the Stoplight Task (Chein et al., 2011). Each run in the YLG involves participants driving on a straight road with 20 intersections, each of which is controlled by a traffic light. Although a run thus consists of 20 trials, it is experienced as continuous. Participants are instructed that the goal of the game is to get the fastest time. At each intersection, when the traffic light turns yellow, participants choose to either continue through the inter-section (Go decision), or to stop the car (Stop decision); they have 1.5 s to make a decision and are not able to accelerate or steer. Failing to make a choice results in a steep time penalty (i.e., a 7.5-s delay), to motivate participants to respond in time. Participants are instructed that Go decisions will result in the fastest time, unless another car is present on the cross street, in which case the participant will crash. Note that the crossing car is not visible until after participants press a button to indicate their decision. Crashes double the time spent at an intersection compared to if the participant had decided to stop. Thus, Go decisions are considered ‘risky’ because this decision can either result in no delay or a 5-s delay (i.e., task performance is uncertain), whereas Stop decisions are considered ‘safe’ because this decision always results in a 2.5-s delay (i.e., task performance is certain). Participants were not informed about the exact delay times, but were told that the duration of the delay was longer for Crashes than for Stop decisions, and longest for Penalties. Upon completion of a run, participants are presented with their completion time and the number of crashes during that run.

A unique feature of the YLG as configured for this study is that there are three different types of intersections, which vary based on the timing of yellow light onset and the presence or absence of a car on the cross street. In this study, within each run of 20 intersections, 4 intersections have a 0.75 probability of crashing (i.e., 3 out of 4 intersections had cars approaching on the cross street); another 4 intersections have a 0.25 probability of crashing; and the remaining 12 intersections have a 0.50 probability of crashing. Thus, the cumulative probability of crashing, across all intersections of each run, is 0.50 (i.e., 10 out of the 20 intersections have cars approaching on the cross street, resulting in a crash if the participant makes a Go decision). Of note, the majority of trials had an equal expected value for Go versus Stop decisions (i.e., trials with a crash probability of 0.50), whereas only a few trials had differing expected values for Go versus Stop decisions (i.e., trials with a crash probability of 0.25 or 0.75); this configuration was chosen to optimize our ability to use the proportion of Go decisions as an index of individual differences in risk-taking behavior without overly encouraging participants to behave in a particular manner (e.g., promoting Go decisions by including more trials with low crash probabilities). This task feature was not explicitly communicated to participants, although implicit learning of this information was possible based on the differential timing of the yellow light onset associated with each type of intersection. That is, intersections at which the light turns yellow earlier (i.e., when the participant is further away from the intersection) signal a greater crash probability. By

including the different types of intersections, it may be possible to distinguish between adaptive and maladaptive risk taking, without promoting risk taking overall, as the cumulative probability of crashing is 0.50. Furthermore, rather than instructing participants about this feature, providing a visual cue allowed for the tracking of participants' ability to learn and attend to subtle contextual cues during risk decisions. For this study however, we collapsed across the different trial types, as our main aim was to investigate differential effects of safe versus risky decisions on outcome processing and we did not have sufficient trials to investigate the differential effects of trial type on behavior and associated brain activation. However, findings with regard to implicit learning based on all six runs that were administered in the MRI scanner will be reported elsewhere (for preliminary findings, see Flournoy et al., 2017). For access to the task, please visit: https://dsn.uoregon.edu/research/yellow-light-game/.

The different types of intersections were randomized to create eight different run configurations. The order of runs was randomized for each participant. Each participant completed 8 runs: 2 runs in a mock scanner, to allow the participant to learn and practice the YLG in a scanner-like environment, and 6 runs in the MRI scanner. The 2 practice runs were added to allow behavioral performance to stabilize before entering the MRI scanner, as previous studies demonstrated increased risk taking and behavioral variability during initial rounds of the Stoplight Task (Kahn et al., 2015; Peake et al., 2013). This study focuses on the first two runs in the MRI scanner, during all of which the participant completed the task while alone (see Table 1 for the summary statistics of run completion time and the number of trials per condition for the runs that were completed in the MRI scanner). Four more runs of the YLG were completed in the MRI scanner while participants were ostensibly being observed by two peers through the use of remote screen sharing; effects of this social manipulation on task behavior and associated brain activation will be reported elsewhere (Flannery et al. in prep).

Behavioral analyses

Overall risk taking was calculated as the probability of Go decisions across a run, with greater probabilities of Go decisions indicating greater risk-taking tendencies. The probability that these Go decisions resulted in Bad outcomes (i.e., crashes) provided an index of *maladaptive* risk taking, with greater crash probabilities indicating less adaptive risk taking. This measure of maladaptive risk taking is based on the expectation that participants who (implicitly) learned the crash probabilities associated with the intersection types would choose to Go on trials with relatively low crash probabilities and would therefore, on average, crash on fewer than 50% of the Go trials in a given run.

To assess for developmental differences in risk taking and maladaptive risk taking, we fit multilevel models using the function 'glmer' with family set to 'binomial' from the package 'lme4' version 1.1–12 (Bates et al., 2015) that was run in R (R Core Team, 2017). Specifically, effects of age and puberty (i.e., mean PDS score) - while controlling for sex and IQ (i.e., FSIQ score) - were assessed in separate models. Each of these models was compared with a null model that included a fixed effect of run (1 and 2), in addition to fixed effects of the control variables (sex and IQ), and a random intercept that accounted for

between-subject differences (see Supplementary Table S3). To facilitate model convergence, we used a mean-centered score for age (i.e., [age −14.4]) and a standardized score for FSIQ (i.e., [(FSIQ - 100)/15]). Because we were missing PDS scores for 2 male participants (one 13-year-old, one 15-year-old), the models including puberty are based on a dataset of 77, instead of 79 participants.

Given the unequal distributions of participants across the age and puberty ranges (Supplementary Fig. S1), we conducted additional analyses based on a median split (see Supplementary Materials) to test whether continuous age and puberty effects on task behavior were driven by biases in the sample distributions (DeCoster et al., 2011; Farrington and Loeber, 2000). Results from the multilevel models that included dichotomous variables based on the median split (age group, PDS group) were similar to the results from the multilevel models that included the continuous variables (age, PDS score; see Supplementary Materials).

fMRI data acquisition

Data were collected using a 3T Siemens Skyra MRI scanner at the Lewis Center for Neuroimaging at the University of Oregon. High-resolution T1-weighted structural images were acquired with the MP- RAGE sequence (TE = 3.41 ms, TR = 2500 ms, flip angle = 7°, 1.0 mm slice thickness, matrix size = 256 × 256, FOV = 256 mm, 176 slices, bandwidth = 190 Hz/pixel). Two functional runs of T2*-weighted BOLD- EPI images were acquired with a gradient echo sequence (TE = 27 ms, TR = 2000 ms, flip angle = 90°, 2.0 mm slice thickness, matrix size = 100 × 100, FOV = 200 mm, 72 slices, bandwidth = 1786 Hz/ pixel). There were 67–90 images per run, depending on participants' performance on the YLG. To correct for local magnetic field in-homogeneities, field maps were also collected (TE = 4.37 ms, TR = 639.0 ms, flip angle = 60°, 2.0 mm slice thickness, matrix size = 100 × 100, FOV = 200 mm, 72 slices, bandwidth = 1515 Hz/pixel).

fMRI preprocessing

Raw imaging files (i.e., DICOMs) were converted to Nifti files using MRIConvert (https://lcni.uoregon.edu/downloads/mriconvert/mriconvert-and-mcverter). Data were then preprocessed in SPM12 (Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) using scripts that can be accessed on GitHub: https://github.com/dsnlab/TDS_scripts/tree/master/fMRI/ppc/spm/tds2. Specifically, anatomical images were coregistered to the 152 Montreal Neurological Institute (MNI) stereotaxic template, segmented into six tissue types, and then used to create a group anatomical template using DARTEL. Functional images were unwarped using the field maps and corrected for head motion through image realignment. A group averaged field map was created (https://github.com/dsnlab/TDS_scripts/tree/master/fMRI/ppc/shell/create_avg_fmap) and used for one participant without a field map, and one participant with an invalid field map (collected in a different space than the functional images). The unwarped and realigned mean functional image was coregistered to the coregistered anatomical image (i.e., the MPRAGE that had been coregistered to the MNI template). Transformations were applied to warp the functional data to the DARTEL group template. Finally, the normalized images were smoothed using a 6-mm FWHM Gaussian kernel. Of

the 81 participants, one 16-year-old male participant was excluded after DICOM conversion due to excessive dropout in the functional images. One volume in one of the runs failed to convert for a 13-year-old male participant; for this participant, we modified the onsets and durations to reflect this missing volume to remain included in the analyses.

Motion artifacts were identified using an automated script that was developed in our lab and assesses changes in mean volume intensity as well as in translation (x, y, z) and rotation (pitch, yaw, roll) in Euclidean distance (https://github.com/dsnlab/auto-motion). Volumes that exceeded 0.3 mm movement in Euclidian distance; volumes that were greater than 3 standard deviations (SDs) above or 1.5 SDs below the mean voxel intensity across subjects; and volumes that were greater than 3 SDs above or below the mean SD of voxel intensity across subjects were marked. Volumes that were immediately preceded by and followed by the marked volumes were also flagged (i.e., “sandwich volumes”). The script flagged volumes for excessive head motion in 50 out of 160 total runs across the remaining 80 participants, with an average of 5.2±5.5 (range: 1–30) volumes identified in each run. One 12-year-old female participant was excluded because >25% of her volumes in one of the runs were flagged for excessive head motion. Thus, imaging analyses were conducted based on 79 participants. Importantly, the DARTEL template that was created during preprocessing also only included these 79 participants.

fMRI analysis

For each participant’s fixed-effects analysis, a general linear model (GLM) was created with three regressors of interest modeled as zero-duration events: Decisions > Driving baseline, Good out-comes > Driving baseline, and Bad outcomes > Driving baseline. In addition, a parametric modulator representing the associated decision type (Stop or Go) was included for each of these primary conditions of interest. Note that for fMRI analysis the runs completed in the MRI scanner were concatenated.

Modeling approach.—There were several rationales for our modeling strategy. First, modeling our primary regressors of interest as Decisions, Good outcomes, and Bad outcomes maximized the number of trials provided by each participant. Second, in our prior work (Kahn et al.,2015), we observed that Stop and Go decisions were more similar than expected, which we attributed to the shared underlying process of making a decision. Our modeling approach allowed us to characterize these primary conditions of interest at the average across Stop and Go decision types, while also allowing further inquiry into the ways in which these conditions varied by decision type via parametric modulators.

We grouped by valence for outcomes based on the idea that both Good outcomes following Go decisions and Good outcomes following Stop decisions result in a *gain* of 2.5 s compared to the alternative scenario. That is, if participants chose to go, they experienced no delay instead of a 2.5-s delay (i.e., the alternative scenario would have been that the participant chose to stop unnecessarily). If participants chose to stop, they experienced a 2.5-s delay instead of a 5-s delay (i.e., the alternative scenario would have been that the participant chose to go while there was a crossing car, which would have resulted in a crash). Similarly, both Bad outcomes following Go decisions and Bad outcomes following Stop decisions

result in a *loss* of 2.5 s compared to the alternative scenario. That is, if participants chose to go, they experienced a 5-s delay instead of a 2.5-s delay (i.e., the alternative scenario would have been that the participant chose to stop and would have prevented a crash). If participants chose to stop, they experienced a 2.5-s delay instead of no delay (i.e., the alternative scenario would have been that the participant chose to go across a safe intersection). Thus, Good (or Bad) outcomes have similar *relative* consequences for task performance, regardless of the type of decision (Stop or Go) that preceded it.

However, parametric modulators were also added to each of the three primary regressors of interest to assess the differences between (1) Go decisions versus Stop decisions; (2) Good outcomes after Go decisions versus Good outcomes after Stop decisions; and (3) Bad outcomes after Go decisions versus Bad outcomes after Stop decisions. To accomplish this, each parametric modulator was mean-centered relative to the average effect of Decisions, Good outcomes, and Bad outcomes, respectively (contrast value for Go = 0.5; for Stop = −0.5).

In addition to the three primary regressors of interest and the three complementary parametric modulators, an additional zero-duration event regressor marked when the game ended (i.e., game over). For participants who did not respond on one or more trials (i.e., penalties), two additional regressors were included: No decision and Penalty outcome. Four motion parameters (Euclidean distance of the translational motion parameters, Euclidean distance of the rotational motion parameters, and the first derivatives of each of these values) were added as regressors of no interest, and an additional motion regressor that flagged volumes with motion artifacts was added using the automated script described above. The fixed-effects GLM created for each participant convolved events with the canonical hemodynamic response function, used a high-pass filter of 128 s to remove low-frequency drift, and estimated serial autocorrelations with a restricted maximum likelihood algorithm using an autoregressive structure (order 1).

Contrasts.—The parameter estimates produced by the GLM were used to create several linear contrast images, including (i) Decisions; (ii) Good outcomes; and (iii) Bad outcomes, each of which was relative to a Driving baseline (i.e., brain activation measured while the participant was driving between intersections); as well as (iv) Good outcomes > Bad outcomes. Additionally, contrast images were created to examine the parametric modulations for each regressor type (Decisions, Good out-comes, and Bad outcomes) to assess (v) Go decisions > Stop decisions; (vi) Good outcomes after Go decisions > Good outcomes after Stop decisions; (vii) Bad outcomes after Go decisions > Bad outcomes after Stop decisions; and (vii) the difference between the outcome-related parametric modulators. Good outcomes after Go versus Stop decisions > Bad outcomes after Go versus Stop decisions.

Whole brain analysis.—Subsequent random-effects analyses (i.e., paired-samples t-tests) were based on these fixed-effects contrast images. Participants were excluded if they had no events of a given type, which was the case for outcome-related parametric modulations only ($N = 3$). Therefore, decision-related results are based on 79 participants, whereas outcome-related results are based on 76–78 participants, depending on the contrast. In addition, we

confirmed that the observed effects held when only including participants that had at least 4 trials of each type (which lowered the Ns by 1–9); these results are available at https://neurovault.org/collections/GHIJXHLP to compare to those presented in this manuscript.

Unless otherwise specified, reported results exceed the minimum cluster extent threshold needed for a 0.05 family-wise error (FWE) rate given a voxel-wise threshold of $p = 0.001$ for each contrast, as determined by AFNI 3dClustSim, version AFNI_16.1.06 (Mar 6 2016). This version of AFNI was updated to resolve recently identified software bugs (Eklund et al., 2016). Smoothness estimates entered into 3dClustSim were an average of subject level spatial autocorrelation function (acf) parameters based on individual subjects' residuals from each group level model, as calculated by 3dFWHMx using the -acf flag. For each contrast, the minimum cluster size at alpha = 0.05 (k_t) is included in the left column of the tables (see Tables 2–4) and ranged from $k_t = 80–109$, depending on the contrast.

ROI analysis.—Given the well-established role of NAcc in reward processing (e.g., Wang et al., 2016), we used the ImCalc function in SPM to create a binary mask of an anatomical ROI for bilateral NAcc from the Harvard-Oxford probabilistic atlas (set at a 75% probability of being in the NAcc). The resulting NAcc masks were 456 mm^3, which given our voxel size of 2 × 2 × 2 mm corresponds to 57 voxels. We used MarsBaR (Brett et al., 2002) to extract mean and median parameter estimates for each of the main contrasts of interest (Go > Stop decisions and Good > Bad outcomes). Developmental effects on NAcc activation were tested by regressing individual parameter estimates from these regions on age as well as puberty, while controlling for sex and IQ, using the function 'lm' from the R package 'lme4' version 1.1–12 (Bates et al.,2015). To test for developmental effects on brain activation outside the NAcc, we conducted additional exploratory whole-brain analyses, to compliment the region-of-interest analysis (see also: Pfeifer and Allen,2016). Specifically, we conducted whole-brain regressions with (1) age and (2) mean PDS score as covariates of interest, while controlling for sex and IQ, to test whether regions outside the NAcc showed age and/or puberty-related differences.

Behavioral results

Age differences in risk-taking behavior—Compared to the null model, additional variance in the probability of Go decisions could not be explained by a main effect of age (X^2 (1) = 0.70, $p = 0.40$). The lack of an age effect was consistent across runs, as indicated by a nonsignificant difference in model fit between the null model and the model with an interaction between age and run (X^2 (2) = 0.80, $p = 0.67$). For the proportion of Go decisions that resulted in Bad outcomes (i.e., crashes), the results were similar: neither the model including a main effect of age (in addition to a main effect of run) nor the model including an interaction between age and run explained additional variance compared with the model including run only (main model: X^2 (1) = 0.03, $p = 0.86$; interaction model: X^2 (2) = 0.26, $p = 0.88$).

Pubertal differences in risk-taking behavior—Compared to the null model, additional variance in the proportion of Go decisions could not be explained by a main effect of puberty (X^2 (1) < 0.001, $p = 0.98$). The lack of a main effect of puberty was consistent

across runs, as indicated by a nonsignificant difference in model fit between the null model and the model that included an interaction between puberty and run (X^2 (2) =2.3, p = 0.32). For the proportion of Go decisions that resulted in crashes, the results were similar: neither the model including a main effect of puberty (in addition to a main effect of run) nor the model including an interaction between puberty and run explained additional variance compared with the model including run only (main model: X^2 (1) = 0.01, p = 0.92; interaction model: X^2 (2) = 1.2, p = 0.55).

Summary of behavioral findings—Together, these findings suggest that individual differences in risk taking during the 'alone' runs (as measured by the proportion of Go decisions, and the proportion of Go decisions that resulted in Bad out-comes) were independent of age and pubertal maturation. Of note, the null models - which provided the best fit - showed no differences in the proportion of Go decisions (age models: z = 0.08, p = 0.94; puberty models: z = −0.08, p = 0.94) or the proportion of Go decisions that resulted in Bad outcomes (age models: z = 1.3, p = 0.18; puberty models: z = 1.3, p = 0.20) during the first versus second 'alone' run. See Supplementary Fig. S2 for the average proportion of Go decisions and the average proportion of Go decisions that resulted in Bad outcomes for each run. Furthermore, results showed that none of the control variables that were added into the null models explained additional variance in the proportion of Go decisions (for the age models, sex: z = −0.7, p = 0.48; IQ: z =−1.1, p = 0.29; and for the puberty models, sex: z = −0.4, p = 0.69; IQ: z = −1.2, p = 0.23), or in the proportion of Go decisions that resulted in Bad outcomes (for the age models, sex: z = −0.9, p = 0.36; IQ: z = −0.7, p = 0.46; and for the puberty models, sex: z = −0.7, p = 0.50; IQ: z = −0.8, p = 0.42), indicating that decision-making did not differ depending on sex or IQ.

Imaging results

All whole-brain results reported in sections 4.1 and 4.2 have been uploaded to NeuroVault (http://neurovault.org/collections/EVANWRAK/). In these NeuroVault images, warm colors indicate the results for the contrasts in *italics* (in sections 4.1 and 4.2), whereas cold colors indicate the results for the reverse contrasts. The whole-brain regressions reported in sections 4.4 have also been uploaded to NeuroVault (https://neurovault.org/collections/3395/). In these NeuroVault images, warm colors indicate brain activation increases, whereas cold colors indicate brain activation decreases with increases in age or pubertal maturation.

Whole-brain results: main effects of decisions and outcomes

Decisions.: Whole-brain results for the contrast *Decisions > Driving baseLine* showed relatively greater bilateral activation in dorsolateral prefrontal cortex (DLPFC), anterior insula, posterior parietal cortex (PPC), cerebellum, thalamus and striatum, in addition to the motor and visual cortices. Results for the reverse contrast (i.e., Driving base-line > Decisions) showed relatively greater bilateral activation in inferior frontal gyrus (IFG) and medial prefrontal cortex (mPFC), in addition to motor, somatosensory, visual, and auditory cortices. Whole-brain results for the contrast *Go decisions > Stop decisions* showed relatively greater bilateral activation in dorsal and ventral striatum (including NAcc), supplemental motor area (SMA), left primary motor and somatosensory cortices, and the

brainstem. Results for the reverse contrast (i.e., Stop decisions > Go decisions) showed relatively greater activation in right IFG and dorsomedial prefrontal cortex (DMPFC), in addition to visual and auditory cortices. For the contrasts above, the regions of activation that survived 3dClustSim correction are reported in Table 2.

Outcomes.: Whole-brain results for the contrast *Outcomes > Driving baseLine* showed relatively greater activation in bilateral anterior insula, right DLPFC, right DMPFC, bilateral precuneus/superior parietal, auditory and visual cortices. Results for the reverse contrast (i.e., Driving baseline > Outcomes) showed relatively greater activation in left DLPFC, posterior cingulate cortex and PPC, ventromedial prefrontal cortex, bilateral putamen, in addition to left primary motor and visual cortices. Whole-brain results for the contrast *Good outcomes > Driving baseline* showed relatively greater activation in bilateral anterior insula, DMPFC, right DLPFC, and bilateral precuneus, in addition to auditory and visual cortices. Results for the reverse contrast (i.e., Driving baseline > Good outcomes) showed relatively greater activation in bilateral putamen and hippocampus, left posterior cingulate, bilateral SMA and left primary motor cortex. Whole-brain results for the contrast *Bad outcomes > Driving baseline* showed relatively greater activation in right lateral orbitofrontal cortex, medial cingulate cortex/SMA, bilateral anterior insula and superior parietal cortex, in addition to bilateral auditory and visual cortices. Results for the reverse contrast (i.e., Driving baseline > Bad outcomes) showed relatively greater activation in bilateral VMPFC, left DLPFC, bilateral dorsal and ventral striatum - including putamen and NAcc, bilateral SMA, and left posterior cingulate cortex. Whole-brain results for the contrast *Good outcomes > Bad outcomes* showed relatively greater activation in bilateral dorsal and ventral striatum - including putamen and NAcc, and bilateral auditory cortex. Results for the reverse contrast (i.e., Bad outcomes > Good outcomes) showed relatively greater activation in the visual cortex. Overlapping regions within the ventral striatum, including NAcc, were activated for both Go decisions > Stop decisions and Good outcomes > Bad outcomes (Fig. 2a). For the contrasts above, the regions of activation that survived 3dClustSim correction are reported in Table 3.

Whole-brain results: outcome processes depend on preceding decisions

For both Good Stop outcomes > Good Go outcomes and Bad Stop out-comes > Bad Go outcomes, whole-brain results showed relatively greater activation in ventral striatum, including putamen and NAcc (Fig. 2b). The reverse contrasts, Good Go outcomes > Good Stop outcomes and Bad Go outcomes > Bad Stop outcomes, showed relatively greater activation in anterior cingulate cortex, mPFC, IFG, and the visual cortex. Furthermore, direct comparison of these outcome-related parametric modulators showed relatively greater activation for Good Stop outcomes vs. Good Go outcomes > Bad Stop outcomes vs. Bad Go outcomes in bilateral caudate, right auditory and primary motor cortices, and cerebellum. For the reverse contrast (i.e., Bad Stop outcomes vs. Bad Go outcomes > Good Stop outcomes vs. Good Go outcomes), results showed relatively greater activation in left DMPFC, middle and posterior cingulate cortices, and bilateral visual cortex (V1 and V5) (Fig. 2c). For the contrasts above, the regions of activation that survived 3dClustSim correction are reported in Table 4.

ROI results: developmental effects on NAcc activation

Significant activation in NAcc was found for four contrasts: Go decisions > Stop decisions, Good outcomes > Bad outcomes, Good Stop outcomes > Good Go outcomes, and Bad Stop outcomes > Bad Go outcomes. We queried whether there were additional contributions of chronological age or pubertal development to NAcc activation, by testing age and puberty in separate regression analyses for all four contrasts.

Age.—Compared with the null model that included main effects of sex and IQ, as well as a random intercept, the model including an additional main effect of age did not explain additional variance in NAcc activation during Go decisions > Stop decisions ($F(1, 75) = 1.6$, $p = 0.21$), during Good outcomes > Bad outcomes ($F(1, 75) = 0.14$,$p = 0.71$), during Good Stop outcomes > Good Go outcomes ($F(1, 73) = 1.1$, $p = 0.29$), or during Bad Stop outcomes > Bad Go outcomes ($F(1, 74) = 1.0$, $p = 0.32$) (Fig. 3a). The same analyses with age group, instead of age, showed similar patterns for a lack of significant differences in NAcc activation between older and younger adolescents during Go decisions > Stop decisions ($F(1, 75) = 0.25$, $p = 0.62$), during Good outcomes > Bad outcomes ($F(1, 75) = 0.20$, $p = 0.66$), during Good Stop outcomes > Good Go outcomes ($F(1, 73) = 1.2$, $p = 0.28$), and during Bad Stop outcomes > Bad Go outcomes ($F(1, 74) = 0.06$, $p = 0.80$). Of note, for each of the best fitting models (i.e., the null models), there were no main effects of sex ($0.19 \leq p \leq 0.79$), or IQ ($0.14 \leq p \leq 0.89$).

Puberty.—Unlike the nonsignificant effects of age, compared with the null model that included main effects of sex and IQ, as well as a random intercept, the model including an additional main effect of mean PDS score explained additional variance in NAcc activation during Good outcomes > Bad outcomes ($F(1, 73) = 8.0$, $p = 0.006$). Note that this result remains significant even at a bonferroni-corrected alpha of 0.00625, which corrects for all eight reported tests (i.e., alpha = 0.05/8 = 0.00625). Puberty did not explain additional variance in NAcc activation during Go decisions > Stop decisions ($F(1, 73) = 2.1$, $p = 0.15$), during Good Stop outcomes > Good Go outcomes ($F(1, 71) = 1.0$, $p = 0.32$), or during Bad Stop outcomes > Bad Go outcomes ($F(1, 72) = 2.7$, $p = 0.11$) (Fig. 3b). The same analyses with PDS group, instead of mean PDS score, confirmed these results for NAcc activation during Good outcomes > Bad outcomes ($F(1, 73) = 4.7$, $p = 0.033$) and for NAcc activation during Go decisions > Stop decisions ($F(1, 73) = 0.06$, $p = 0.81$), during Good Stop outcomes > Good Go outcomes ($F(1, 71) = 0.44$, $p = 0.51$), and during Bad Stop outcomes > Bad Go outcomes ($F(1, 72) = 0.43$, $p = 0.51$). Of note, for each of the null models (for Go decisions > Stop decisions; Good Stop outcomes > Good Go outcomes; Bad Stop outcomes > Bad Go outcomes), there were no main effects of sex ($0.22 \leq p \leq 0.30$), or IQ ($0.51 \leq p \leq 0.89$). In the best fitting models for Good outcomes > Bad outcomes, there were no main effects of sex (PDS model: $p = 0.19$; PDS group model: $p = 0.36$), or IQ (PDS model: $p = 0.09$; PDS group model: $p = 0.18$).

Together, these findings suggest that NAcc activation, particularly during outcome processing, is associated with pubertal maturation. Specifically, adolescents who were more mature were more likely to differentially activate their NAcc in response to positive outcomes (i.e., safely crossing an intersection, or having stopped in the presence of crossing

car), as opposed to negative outcomes (i.e., crashing, or having stopped at a safe intersection). These findings were confirmed by the median split analysis and replicated with regression analyses with the median rather than mean parameter estimates extracted from the ROI (see Supplemental Materials).

Exploratory analyses

Whole-brain results with age as a regressor-of-interest, while controlling for sex and IQ, showed activation decreases with age in the right cerebellum (peak voxel: $x = 34$, $y = -82$, $z = -42$; $t = 3.2$, $p = 0.001$, $k = 87$; Fig. 4a) for Outcomes > Driving baseline and Bad outcomes > Driving baseline. Whole-brain results with mean PDS score as a regressor-of-interest, while controlling for sex and IQ, showed activation decreases with pubertal maturation in the right temporal parietal junction (peak voxel: $x = 52$, $y = -40$, $z = 40$; $t = 3.2$, $p = 0.001$, $k = 91$; Fig. 4b) for Good outcomes > Driving baseline and in left cerebellum (peak voxel: $x = -44$, $y = -54$, $z = -34$; $t = 3.2$, $p = 0.001$, $k = 88$; Fig. 4c) for Bad outcomes > Driving baseline. Together, these findings indicate that while younger adolescents (age effects) as well as less mature adolescents (pubertal effects) tend to activate cerebellum more while processing negative outcomes, less mature adolescents tend to activate right TPJ more while processing positive outcomes (compared to while driving). There were no significant age or puberty effects on brain activation for any of the other contrasts.

Discussion

The goal of this study was to investigate neural and psychological processes associated with decision making and outcome processing that might contribute to risky behavior in adolescence, and explore developmental differences in these processes across early and middle adolescence. To this end, we adapted the Stoplight Task (Chein et al., 2011). This new paradigm, referred to as the Yellow Light Game (YLG), allowed us to (1) measure brain activation during decisions and outcomes, (2) assess the extent to which decisions impact outcome-related brain activation, and (3) possibly distinguish maladaptive risk taking (i.e., the proportion of Go decisions that resulted in Bad outcomes) from a more general tendency to choose a risky option (i.e., the proportion of Go decisions).

The absence of developmental differences in laboratory risk taking

As expected, results showed no age or puberty effects on risk-taking behavior - i.e., there were no age or puberty-related differences in the probability of Go decisions and the probability of Go decisions that resulted in Bad outcomes (i.e., Crashes). These findings are consistent with a recent meta-analysis that investigated developmental differences in risk-taking behavior using laboratory tasks (Defoe et al., 2015). Specifically, while results from this meta-analysis showed that early adolescents (11–13 years) took more risks compared to mid-late adolescents (14–19 years), the reported effect size was very small (Hedges' g = 0.15), which supports the idea that developmental differences in risk-taking behavior are largely absent when measured using laboratory paradigms.

Interestingly, results from the present study were similar for overall risk taking (i.e., the probability of Go decisions) and maladaptive risk taking (i.e., the probability of Go

decisions that resulted in Crashes), suggesting that individual differences in risky decisions, regardless of whether it is advantageous or disadvantageous to take a risk, may be consistent across adolescence. Although researchers have started to acknowledge the importance of distinguishing between adaptive and maladaptive risk taking in a laboratory setting (e.g., Humphreys et al.,2013), there is limited research on the development of these different types of risk taking during adolescence, and on how laboratory measures of adaptive and maladaptive risk taking generalize to real-world risky behaviors (c.f., Rao et al., 2011). On the other hand, because there is a relative dearth of information about how many of these laboratory tasks correlate with risk taking in the real world, individual differences in adaptive and maladaptive risk taking may not be consistent across adolescence if measured in the real world instead of the laboratory.

A novel perspective on the role of reward processes

Previous studies that separately investigated decision and outcome-related processes within a single paradigm have reported increased NAcc activation during riskier decisions and during the processing of positive versus negative outcomes experienced upon risky decisions (e.g., Op de Macks et al., 2016; Van Leijenhorst et al., 2010). Replicating these previous findings, we found overlapping regions of activation, including activation in the nucleus accumbens, for Go > Stop decisions and for Good > Bad outcomes. In contrast to previously used paradigms that compared positive to negative outcomes following risky decisions only, the YLG allowed us to compare positive and negative outcomes experienced upon safe decisions, in addition to risky decisions. As such, we were able to assess the extent to which outcome-related brain processes depended on different types of decisions (i.e., safe > risky). Results of the present study demonstrated elevated NAcc activation during outcomes experienced upon Stop decisions > Go decisions, regardless of the type of outcome. That is, NAcc activation was elevated for both Good outcomes and Bad outcomes following Stop decisions compared to Go decisions. In other words, NAcc was significantly more active when the participant's decision prevented crashing into a crossing car compared to when it resulted in going through an intersection without crashing; and NAcc was significantly less active when the participant's decision led to crashing compared to when having stopped unnecessarily.

Importantly, NAcc activation during Good outcomes and Bad outcomes was similarly impacted by safe and risky decisions, as indicated by the lack of significant differences in NAcc activation for the contrast that compared the outcome-related parametric modulators (i.e., when contrasting Good outcomes following Stop versus Go decisions with Bad outcomes following Stop versus Go decisions). While prior studies that have directly tested the relation between NAcc activation and subjective experience have reported that increased NAcc activation corresponded with more positive value ratings (Galván and McGlennen, 2013), it is unclear why adolescents might experience outcomes following safe decisions as more rewarding than outcomes following risky decisions. However, it is consistent with their decision preferences, as on average adolescents chose to stop more often than go. In other words, adolescents may prefer a missed go opportunity (bad stop) versus a decidedly nasty go mistake (bad go), as well as a correct decision that avoids a crash (good stop) versus a correct decision that wins them a bit of time (good go). Future studies that use the YLG or

other similarly flexible paradigms are needed to directly assess the subjective experience of these outcomes.

Although existing theories about adolescent risk taking have proposed that neural responses to rewards, often operationalized as NAcc activation, are what motivate adolescents to take risks (Galvan, 2010, 2014; Van Duijvenvoorde et al., 2016), it remains to be determined how outcome-related NAcc activation contributes to subsequent decision-making. Interestingly, one study that administered the Stoplight Task in a sample of 14–16 year-olds showed that increased activation in ventral striatum, including NAcc, preceded risky decisions (i.e., Go decisions; Kahn et al., 2015), suggesting that NAcc activation might indicate the motivation to take a risk. A potential future direction might be to test whether adolescents are most inclined to take a risk after successfully avoiding a bad outcome during the YLG, and least inclined to take a risk after crashing.

Developmental differences in NAcc activation during outcomes, but not decisions

In contrast with previously reported findings (Braams et al., 2015; Silverman et al., 2015), results of this study revealed no age differences in NAcc activation during outcome processing (i.e., Good > Bad outcomes). Although the age range in this study (11–17 years) was chosen to capture a period in adolescent development during which pubertal changes occur, a possible explanation for the lack of age differences could be that the age range of our sample might have been too close to the age at which the peak in NAcc activation occurs (i.e., 16 years; Braams et al., 2015) to detect any age differences. Alternatively, the age range may have been too narrow to detect age differences, as the age range in Braams et al. (2015) started from 8 years, rather than from 11 years.

Rather than age-related differences, however, the results did reveal puberty-related differences in NAcc activation during outcome processing (i.e., Good outcomes > Bad outcomes), but not during decision-making (i.e., Go decisions > Stop decisions). Specifically, more mature adolescents were more likely to differentially activate NAcc during positive versus negative outcomes, regardless of the type of decision that preceded it. This finding suggests that positive outcomes might become more salient than negative outcomes as adolescents progress through puberty. While this idea is in line with existing neurobiological models that propose puberty plays a role in elevating subcortical reward-related brain activation (Crone and Dahl, 2012; Nelson et al., 2005, 2016; Peper and Dahl, 2013), longitudinal studies that utilize the YLG or other similarly flexible paradigms are necessary to confirm whether pubertal maturation or chronological age is a more important predictor of outcome-related NAcc activation. Note that puberty usually starts between 8 and 11 years, especially in girls (Biro et al., 2014), which suggests that puberty-related brain changes might already be occurring before age 11 years.

It is interesting that puberty was associated with NAcc activation in the absence of complementary associations with behavior. It is frequently considered useful when behavioral performance of functional tasks is matched across a given variable (puberty, age, diagnosis, etc.) so that observed neural differences can be interpreted straightforwardly as such, rather than as consequences of ‘doing the task differently.’ In other words, neural patterns may vary in the absence of behavioral differences on the same task. Nevertheless,

we expect that the observed neural differences are meaningful and that in other contexts (whether in the laboratory or real world), the increases in NAcc activation with pubertal development would have behavioral consequences (e.g., Braams et al., 2015). Alternatively, future studies with additional trials, 'practice' runs, and/or contexts for the YLG may reveal a behavioral association with puberty.

The role of cortical midline structures in risk taking

While outcome processes following Stop decisions were associated with relatively greater activation in NAcc compared to outcome processes following Go decisions, the reverse contrast showed relatively greater activation in cortical midline structures, including medial pre-frontal cortex, anterior cingulate cortex and precuneus. This network of cortical midline structures has been strongly implicated in self-referential processing across development (Denny et al., 2012; Pfeifer et al., 2013; Pfeifer and Berkman, 2018), as has the NAcc - leading some to propose an intrinsic overlap between self-relevance and reward (Northoff and Hayes, 2011; D'Argembeau, 2013). As such, these findings may suggest that outcomes following a risky decision might be experienced as more self-relevant compared to a safe decision. In other words, outcomes that are the result of an active decision to engage in a risky situation (i.e., approach behavior) might be experienced as more self-relevant than outcomes that result from a more passive decision to stop (i.e., avoidance behavior). Interestingly, one prior study showed that heightened activation of, and functional connectivity between, regions involved in self-referential and reward processing during positive social evaluations was associated with increased risky sexual behavior among adolescents (Eckstrand et al., 2017). While these findings suggest that risk taking might be related to a combination of heightened reward and self-referential processing, further research is needed to investigate the extent to which these processes are involved in risk-taking behavior as opposed to other phenomena also associated with activity in cortical midline structures, including approach/avoidance behavior, performance monitoring, or self-regulation (Tamnes et al., 2013; van Lei-jenhorst et al., 2006; van Noordt and Segalowitz, 2012).

Limitations and future directions

This study is part of a larger project in which decision-making among adolescents from the community will be compared with decision-making among adolescents with a history of early adversity (i.e., those with involvement in the child welfare and/or juvenile justice system). To accomplish this ultimate goal, it was necessary to constrain the age range of the current sample, which limits the extent of cross-sectional age-related analyses we were able to conduct. The age range of this sample was somewhat better tailored to explore pubertal differences, which may account for our observation of pubertal but not age-related effects. However, puberty usually starts between 8 and 11 years, especially in girls (Biro et al., 2014), which suggests that puberty-related brain changes might already be occurring before the age of our youngest participants (11 years) and should be assessed in future studies using the YLG.

In addition, there are aspects of the YLG that could be adjusted to increase the paradigm's ecological validity. For example, the speed of the vehicle was constant throughout the game

and could not be modulated by the participant. Including a gas pedal option to the game would not only increase the game's ecological validity, but it would also allow for the measurement of risk-taking behavior as a function of its use. Alternatively, risk-taking behavior could be measured as a function of response times (i.e., the amount of time it takes participants to decide), providing a potentially more nuanced measure of risk taking. For example, in a study that used an earlier version of the Stoplight Task, called Chicken (Gardner and Steinberg, 2005), risk-taking behavior was measured based on the amount of time that the car was in motion (starting from when the light turned yellow) as well as the number of restarts (participants could restart the car to move it further, if they had not crashed yet). In the current version of the YLG, we decided not to operationalize 'risk taking' based on the response times because we did not instruct our participants to respond as quickly as possible. Thus, the validity of using response times as an indication risk preference is uncertain, as hesitation by the participant could be reflective of risk avoidance but could also be the consequence of other processes (e.g., waiting for a cue). For future versions of the YLG, instructing participants to respond as quickly as possible could reduce the prevalence of 'waiting behavior' and would thus provide perhaps a more accurate measure of risk preference. By allowing for the measurement of different *types* of risk taking (e.g., running a yellow light, speeding), future studies could help us gain a better understanding of which types of risky decisions correspond with the risky behaviors that increase adolescent mortality - the ultimate test of ecological validity.

Conclusions

To advance our understanding of adolescent decision-making, we developed a new decision-making paradigm that built on a previous measure. Our goal was to gain novel insights into the role of reward processes in adolescent decision-making while providing adolescents with the opportunity to take risks independent from task performance. With this modified paradigm, we demonstrated that across adolescence outcome-related NAcc activation is modulated by the type of decision that precedes it. Importantly, results suggested that adolescents might experience avoiding a larger loss by incurring a smaller loss (i.e., stopping at a yellow light and successfully avoiding a crash) as more rewarding than successfully taking a risk (i.e., driving through a yellow light without crashing), which calls into question the idea that adolescents engage in risk taking because they experience positive outcomes following risky decisions as particularly rewarding. Together, these novel insights highlight the importance of continuing to increase the precision of how we measure reward-related brain activation across different types of decisions and outcomes (Pfeifer and Allen, 2016).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by grants from the National Institute on Drug Abuse of the National Institutes of Health [P50 DA035763 to P.A.F. and J.H.P., and R21 DA043015 to J.H.P.] and a National Science Foundation Graduate Research Fellowship [2015172132] to J.E.F.; its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors would like to thank Maureen Durnin and

Garrett Ross for their assistance with data acquisition and data management, as well as Justus Post and Timothy Heider of Organic Workflow for assistance with developing the Yellow Light Game application. The authors declare no competing financial interests.

References

Barkley-Levenson EE, Van Leijenhorst L, Galván A, 2013 Behavioral and neural correlates of loss aversion and risk avoidance in adolescents and adults. Dev. Cogn. Neurosci. 3, 72–83. 10.1016/j.dcn.2012.09.007. [PubMed: 23245222]

Bates D, Maechler M, Bolker B, Walker S, 2015 Fitting linear mixed-effects models using lme4. J. Stat. Software 67 (1), 1–48. 10.18637/jss.v067.i01.

Biro FM, Pinney SM, Huang B, Baker ER, Walt Chandler D, Dorn LD, 2014 Hormone changes in peripubertal girls. J. Clin. Endocrinol. Metab. 99 (10), 3829–3835. 10.1210/jc.2013-4528. [PubMed: 25029416]

Bjork JM, Smith AR, Danube CL, Hommer DW, 2007 Developmental differences in posterior mesofrontal cortex recruitment by risky rewards. J. Neurosci. 27 (18), 4839–4849. 10.1523/JNEUR0SCI.5469-06.2007. [PubMed: 17475792]

Braams BR, van Duijvenvoorde AC, Peper JS, Crone EA, 2015 Longitudinal changes in adolescent risk-taking: a comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. J. Neurosci. 35 (18), 7226–7238. 10.1523/JNEUROSCI.4764-14.2015. [PubMed: 25948271]

Brett M, Anton JL, Valabregue R, Poline JB, 2002 Region of interest analysis using an SPM toolbox. Neuroimage 16 (2).

Casey BJ, Jones RM, Somerville LH, 2011 Braking and accelerating of the adolescent brain. J. Res. Adolesc. 21 (1), 21–33. 10.1111/j.1532-7795.2010.00712.x. [PubMed: 21475613]

Chein J, Albert D, O'Brien L, Uckert K, Steinberg L, 2011 Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. Dev. Sci. 14 (2), F1–F10. 10.1111/j.1467-7687.2010.01035.x. [PubMed: 21499511]

Crone EA, Dahl RE, 2012 Understanding adolescence as a period of social-affective engagement and goal flexibility. Nat. Rev. Neurosci. 13 (9), 636–650. 10.1038/nrn3313. [PubMed: 22903221]

Dahl RE, 2004 Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. Ann. N. Y. Acad. Sci. 1021, 1–22. 10.1196/annals.1308.001. [PubMed: 15251869]

D'Argembeau A, 2013 On the role of the ventromedial prefrontal cortex in self-processing: the valuation hypothesis. Front. Hum. Neurosci. 7, 372 10.3389/fnhum.2013.00372. [PubMed: 23847521]

DeCoster J, Gallucci M, Iselin AR, 2011 Best practices for using median splits, artificial categorization, and their continuous alternatives. J Exp Psychopathol. 2 (2), 197–209. 10.5127/jep.008310.

Defoe IN, Dubas JS, Figner B, van Aken MA, 2015 A meta-analysis on age differences in risky decision making: adolescents versus children and adults. Psychol. Bull. 141 (1), 48–84. 10.1037/a0038088. [PubMed: 25365761]

Denny BT, Kober H, Wager TD, Ochsner KN, 2012 A meta-analysis of functional neuroimaging studies of self and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. J. Cognit. Neurosci. 24 (8), 1742–1752. 10.1162/jocn_a_00233. [PubMed: 22452556]

Eckstrand KL, Choukas-Bradley S, Mohanty A, Cross M, Allen NB, Silk JS, Jones NP, Forbes EE, 2017 Heightened activity in social reward networks is associated with adolescents' risky sexual behaviors. Dev Cogn Neurosci. 27, 1–9. 10.1016/j.dcn.2017.07.004. [PubMed: 28755632]

Eklund A, Nichols TE, Knutsson H, 2016 Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. Proc. Natl. Acad. Sci. Unit. States Am. 113 (28), 7900–7905. 10.1073/pnas.1602413113.

Ernst M, 2014 The triadic model perspective for the study of adolescent motivated behavior. Brain Cognit. 89, 104–111. https://doi.org/10.10167j.bandc.2014.01.006. [PubMed: 24556507]

Ernst M, Nelson EE, Jazbec S, McClure EB, Monk CS, Leibenluft E, Blair J, Pine DS, 2005 Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and

adolescents. Neuroimage 25, 1279–1291. 10.1016/j.neuroimage.2004.12.038. [PubMed: 15850746]

Farrington DP, Loeber R, 2000 Some benefits of dichotomization in psychiatric and criminological research. Crim. Behav. Ment. Health 10 (2), 100–122. 10.1002/cbm.349.

Flournoy JC, Peake SJ, Alberti S, Flannery JE, Mobasser A, Fisher PA,Pfeifer JH, 2017 Evaluation of a Bayesian Cognitive Model for Adolescent Risky Decision Making in the Yellow Light Game. figshare. 10.6084/m9.figshare.5393329.v1.

Forbes EE, Ryan ND, Phillips ML, Manuck SB, Worthman CM, Moyles DL, Tarr JA, Sciarrillo SR, Dahl RE, 2010 Healthy adolescents' neural response to reward: associations with puberty, positive affect, and depressive symptoms. J. Am. Acad. Child Adolesc. Psychiatry 49 (2), 162–172. [PubMed: 20215938]

Galvan A, 2010 Adolescent development of the reward system. Front. Hum. Neurosci. 4 (6) 10.3389/neuro.09.006.2010.

Galván A, 2014 Neural systems underlying reward and approach behaviors in childhood and adolescence. Curr Top Behav. Neurosci. 16, 167–188. 10.1007/7854_2013_240. [PubMed: 23959429]

Galván A, McGlennen KM, 2013 Enhanced striatal sensitivity to aversive reinforcement in adolescents versus adults. J. Cognit. Neurosci. 25 (2), 284–296. 10.1162/jocn_a_00326. [PubMed: 23163417]

Gardner M, Steinberg L, 2005 Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. Dev. Psychol. 41 (4), 625–635. 10.1037/0012-1649.41.4.625. [PubMed: 16060809]

Humphreys KL, Lee SS, Tottenham N, 2013 Not all risk taking behavior is bad: associative sensitivity predicts learning during risk taking among high sensation seekers. Pers. Indiv. Differ. 54 (6), 709–715. 10.1016/j.paid.2012.11.031.

Kahn LE, Peake SJ, Dishion TJ, Stormshak EA, Pfeifer JH, 2015 Learning to play it safe (or not): stable and evolving neural responses during adolescent risky decision-making. J. Cognit. Neurosci. 27 (1), 13–25. 10.1162/jocn_a_00694. [PubMed: 25100220]

May JC, Delgado MR, Dahl RE, Stenger VA, Ryan ND, Fiez JA, Carter CS, 2004 Event-related functional magnetic resonance imaging of reward-related brain circuitry in children and adolescents. Biol. Psychiatr. 55, 359–366. 10.1016/j.biopsych.2003.11.008.

Mulye TP, Park MJ, Nelson CD, Adams SH, Irwin CE Jr., Brindis CD, 2009 Trends in adolescent and young adult health in the United States. J. Adolesc. Health 45, 8–24. 10.1016/j.jadohealth.2009.03.013. [PubMed: 19541245]

Nelson EE, Jarcho JM, Guyer AE, 2016 Social re-orientation and brain development: an expanded and updated view. Dev Cogn Neurosci. 17, 118–127. 10.1016/j.dcn.2015.12.008. [PubMed: 26777136]

Nelson EE, Leibenluft E, McClure EB, Pine DS, 2005 The social re-orientation of adolescence: a neuroscience perspective on the process and its relation to psychopathology. Psychol. Med. 35 (2), 163–174. [PubMed: 15841674]

Northoff G, Hayes DJ, 2011 Is our self nothing but reward? Biol. Psychiatr. 69 (11), 1019–1025. 10.1016/j.biopsych.2010.12.014.

Op de Macks ZA, Bunge SA, Bell ON, Wilbrecht L, Kriegsfeld LJ, Kayser AS, Dahl RE, 2016 Risky decision-making in adolescent girls: the role of pubertal hormones and reward circuitry. Psychoneuroendocrinology 74, 77–91. 10.1016/j.psyneuen.2016.08.013. [PubMed: 27591399]

Op de Macks ZA, Gunther Moor B, Overgaauw S, Guroglu B, Dahl RE, Crone EA, 2011 Testosterone levels correspond with increased ventral striatum activation in response to monetary rewards in adolescents. Dev Cogn Neurosci. 1 (4), 506–516. 10.1016/j.dcn.2011.06.003. [PubMed: 22436568]

Patton GC, Sawyer SM, Viner RM, Haller DM, Bose K, Vos T, Ferguson J, Mathers CD, 2009 Global patterns of mortality in young people: a systematic analysis of population health data. Lancet 374 (9693), 881–892. 10.1016/S0140-6736(09)60741-8. [PubMed: 19748397]

Paulsen DJ, Carter RM, Platt ML, Huettel SA, Brannon EM, 2012 Neurocognitive development of risk aversion from early childhood to adulthood. Front. Hum. Neurosci. 5, 178 10.3389/fnhum.2011.00178. [PubMed: 22291627]

Peake SJ, Dishion TJ, Stormshak EA, Moore WE, Pfeifer JH, 2013 Risk-taking and social exclusion in adolescence: neural mechanisms underlying peer influences on decision-making. Neuroimage 82, 23–34. 10.1016/j.neuroimage.2013.05.061. [PubMed: 23707590]

Peper JS, Braams BR, Blankenstein NE, Bos MGN, Crone EA, 2018 Development of multiple-faceted risk-taking and the relations to sex steroid hormones: A longitudinal study. Child Dev 10.1111/cdev.13063.

Peper JS, Dahl RE, 2013 Surging hormones: brain-behavior interactions during puberty. Curr. Dir. Psychol. Sci. 22 (2), 134–139. 10.1177/0963721412473755. [PubMed: 26290625]

Petersen AC, Crockett L, Richards M, Boxer A, 1988 A self-report measure of pubertal status: reliability, validity, and initial norms. J. Youth Adolesc. 17 (2), 117–133. 10.1007/BF01537962. [PubMed: 24277579]

Pfeifer JH, Allen NB, 2016 The audacity of specificity: moving adolescent developmental neuroscience towards more powerful scientific paradigms and translatable models. Dev Cogn Neurosci. 17, 131–137. 10.1016/j.dcn.2015.12.012. [PubMed: 26754460]

Pfeifer JH, Berkman ET, 2018 The development of self and identity in adolescence: neural evidence and implications for a value-based choice perspective on motivated behavior. Child Development Perspectives. 10.1111/cdep.12279.

Pfeifer JH, Kahn LE, Merchant JS, Peake SJ, Veroude K, Masten CL, Lieberman MD, Mazziotta JC, Dapretto M, 2013 Longitudinal change in the neural bases of adolescent social self-evaluations: effects of age and pubertal development. J. Neurosci. 33 (17), 7415–7419. 10.1523/JNEUROSCI.4074-12.2013. [PubMed: 23616547]

Rao U, Sidhartha T, Harker KR, Bidesi AS, Chen LA, Ernst M, 2011 Relationship between adolescent risk preferences on a laboratory task and behavioral measures of risk-taking. J. Adolesc. Health 48 (2), 151–158. 10.1016/j.jadohealth.2010.06.008. [PubMed: 21257113]

R Core Team, 2017 R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.

Richards JM, Plate RC, Ernst M, 2013 A systematic review of fMRI reward paradigms used in studies of adolescents vs. adults: the impact of task design and implications for understanding neurodevelopment. Neurosci. Biobehav. Rev. 37 (5), 976–991. 10.1016/j.neubiorev.2013.03.004. [PubMed: 23518270]

Shirtcliff EA, Dahl RE, Pollak SD, 2009 Pubertal development: correspondence between hormonal and physical development. Child Dev. 80 (2), 327–337. 10.1111/j.1467-8624.2009.01263.x. [PubMed: 19466995]

Silverman MH, Jedd K, Luciana M, 2015 Neural networks involved in adolescent reward processing: an activation likelihood estimation meta-analysis of functional neuroimaging studies. Neuroimage 122, 427–439. 10.1016/j.neuroimage.2015.07.083. [PubMed: 26254587]

Steinberg L, 2008 A social neuroscience perspective on adolescent risk-taking. Dev. Rev. 28 (1), 78–106. 10.1016/j.dr.2007.08.002. [PubMed: 18509515]

Tamnes CK, Walhovd KB, Torstveit M, Sells VT, Fjell AM, 2013 Performance monitoring in children and adolescents: a review of developmental changes in the error-related negativity and brain maturation. Dev. Cognit. Neurosci. 6, 1–13. 10.1016/j.dcn.2013.05.001. [PubMed: 23777674]

Van Duijvenvoorde AC, Huizenga HM, Somerville LH, Delgado MR, Powers A, Weeda WD, Casey BJ, Weber EU, Figner B, 2015 Neural correlates of expected risks and returns in risky choice across development. J. Neurosci. 35 (4), 1549–1560. 10.1523/JNEUROSCI.1924-14.2015. [PubMed: 25632132]

Van Duijvenvoorde AC, Op de Macks ZA, Overgaauw S, Gunther Moor B, Dahl RE, Crone EA, 2014 A cross-sectional and longitudinal analysis of reward-related brain activation: effects of age, pubertal stage, and reward sensitivity. Brain Cognit. 89, 3–14. 10.1016/j.bandc.2013.10.005. [PubMed: 24512819]

Van Duijvenvoorde AC, Peters S, Braams BR, Crone EA, 2016 What motivates adolescents? Neural responses to rewards and their influence on adolescents' risk taking, learning, and cognitive control. Neurosci. Biobehav. Rev. 70, 135–147. 10.1016/j.neubiorev.2016.06.037. [PubMed: 27353570]

van Leijenhorst L, Crone EA, Bunge SA, 2006 Neural correlates of developmental differences in risk estimation and feedback processing. Neuropsychologia 44 (11), 2158–2170. 10.1016/j.neuropsychologia.2006.02.002. [PubMed: 16574168]

Van Leijenhorst L, Gunther Moor B, Op de Macks ZA, Rombouts SARB, Westenberg PM, Crone EA, 2010 Adolescent risky decision-making: neurocognitive development of reward and control regions. Neuroimage 51, 345–355. 10.1016/j.neuroimage.2010.02.038. [PubMed: 20188198]

van Noordt S, Segalowitz SJ, 2012 Performance monitoring and the medial prefrontal cortex: a review of individual differences and context effects as a window on self-regulation. Front. Hum. Neurosci. 6, 197 10.3389/fnhum.2012.00197. [PubMed: 22798949]

Wang KS, Smith DV, Delgado MR, 2016 Using fMRI to study reward processing in humans: past, present, and future. J. Neurophysiol. 115 (3), 1664–1678. 10.1152/jn.00333.2015. [PubMed: 26740530]

Wechsler D, 1999 Wechsler Abbreviated Scale of Intelligence The Psychological Corporation. Harcourt Brace & Company, New York, NY.

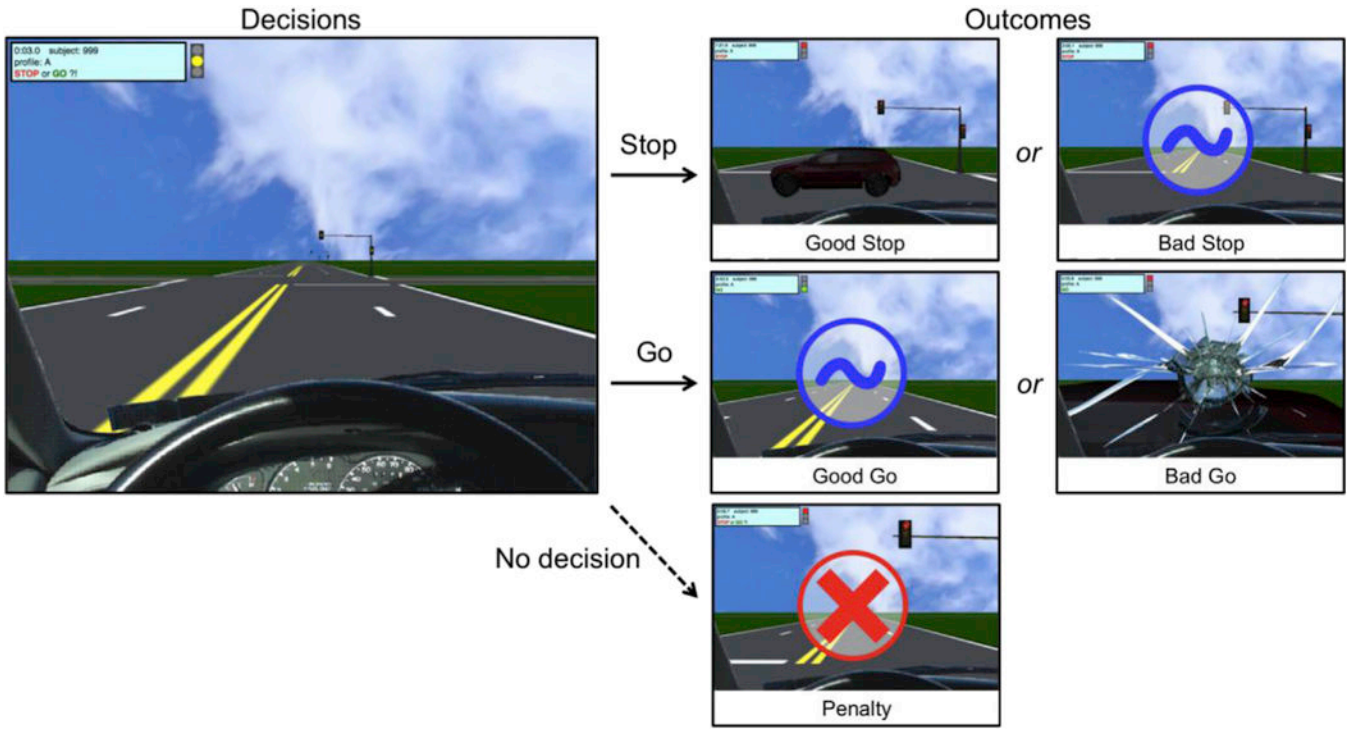

Fig. 1.

The Yellow Light Game. Brain activation was modeled for decisions (i.e., at yellow-light onset) and outcomes. Both Stop and Go decisions result in either Good or Bad outcomes; if the participant fails to press a button (i.e., no decision), they receive a penalty. For Stop decisions, Good outcomes involve seeing a crossing car, indicating that the participant prevented a crash (and incurred a shorter delay); Bad outcomes involve seeing a safe intersection, indicating that the participant could have gone through the intersection safely (and prevented a delay). For Go decisions, Good outcomes involve seeing a safe intersection, indicating that the participant crossed without crashing (and prevented a delay); Bad outcomes involve crashing into a crossing car (and incurring a delay). Note that Bad outcomes for Stop decisions and Good outcomes for Go decisions are visually similar (i.e., blue circle).

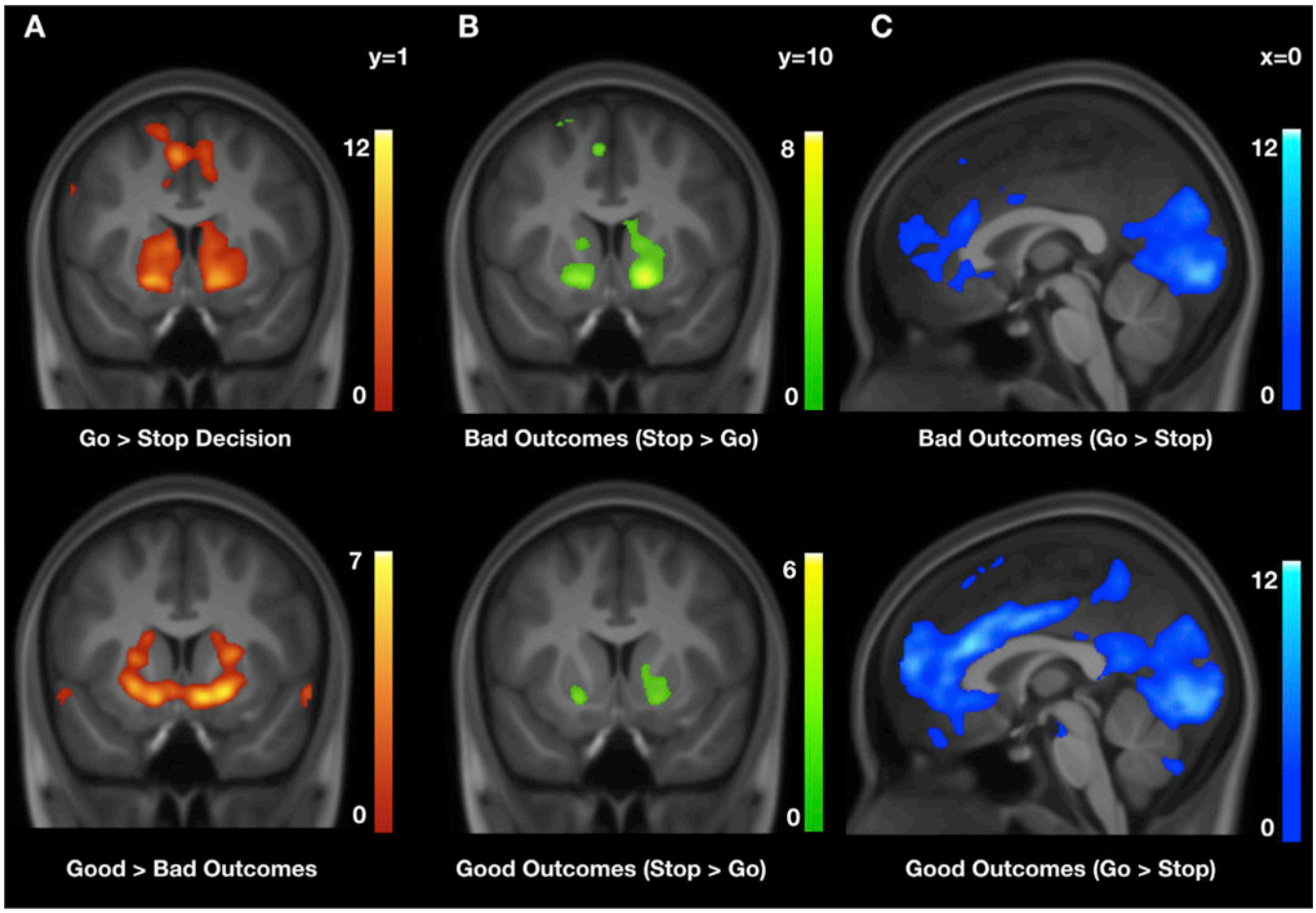

Fig. 2.

Whole-brain results. Overlapping regions of activation in the ventral striatum, including nucleus accumbens, for (a) Go decisions > Stop decisions and Good outcomes > Bad outcomes; (b) Bad Stop outcomes > Bad Go outcomes and Good Stop outcomes > Good Go outcomes; and (c) Bad Go outcomes > Bad Stop outcomes and Good Go outcomes > Good Stop outcomes. Results displayed here exceeded the minimum cluster extent threshold needed for a 0.05 family-wise error (FWE) rate given a voxel-wise threshold of $p = 0.001$. The cluster extent threshold (k_t) for each contrast was calculated using 3dClustSim in AFNI and can be found in Tables 2–4. Heatmaps refer to t values.

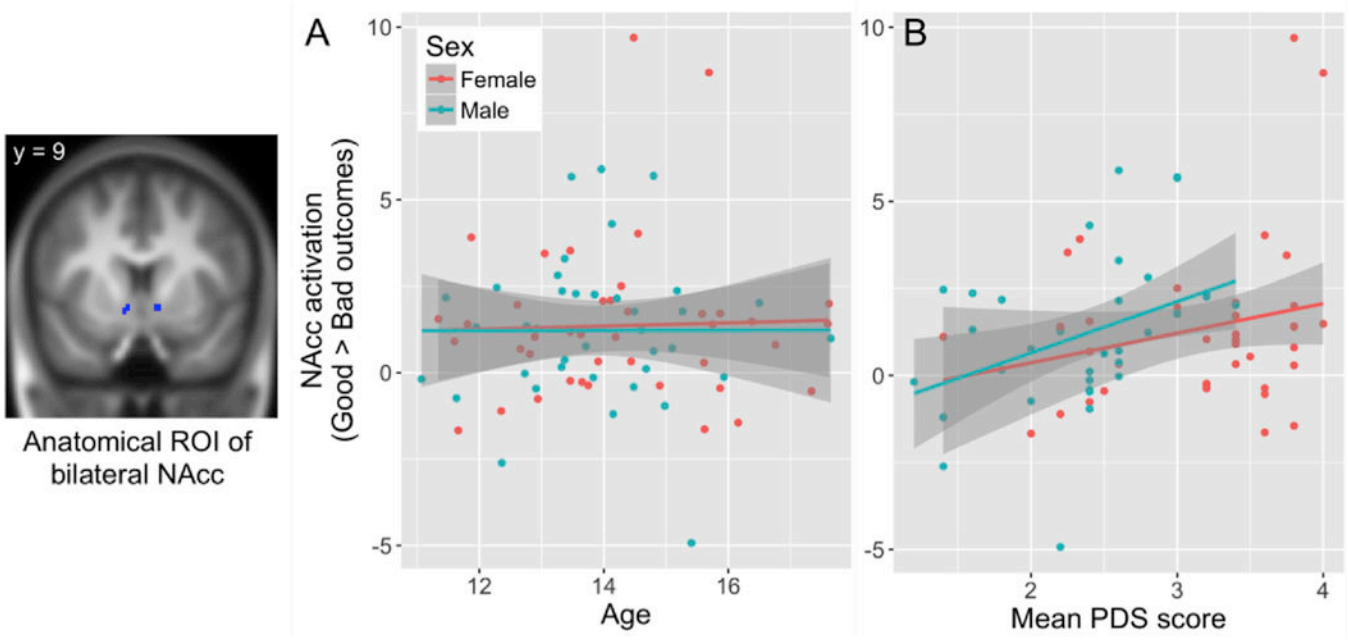

Fig. 3.

ROI results. Outcome-related nucleus accumbens activation plotted against (a) age and (b) puberty (PDS = Pubertal Development Scale). Note: the two female participants with the highest activation parameters (i.e., parameter estimate > 8) were not outliers according to their Mahalanobis distance based on their age, IQ, mean PDS score, and task behavior. However, results remained significant even after removal of these two participants ($F(1, 71) = 4.7$, $p = 0.034$).

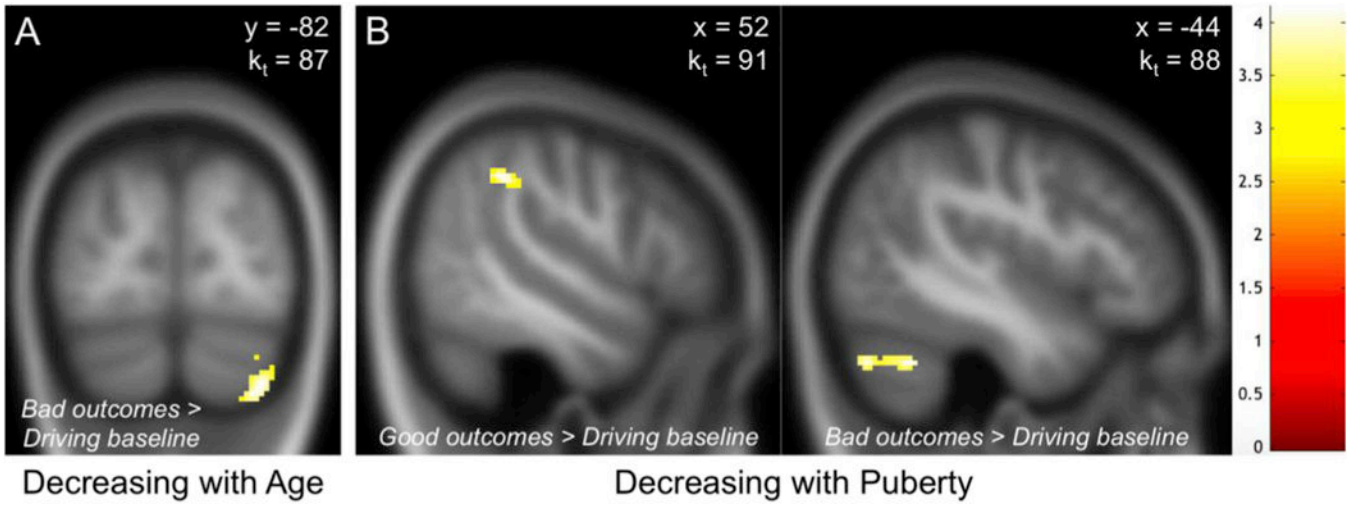

Fig. 4.

Exploratory whole-brain regression analyses. Results of (a) age and (b) puberty (i.e., mean PDS score) effects on outcome processing displayed here exceeded the minimum cluster extent threshold needed for a 0.05 family-wise error (FWE) rate given a voxel-wise threshold of $p = 0.001$. The cluster extent threshold (k_t) for each contrast was calculated using 3dClustSim in AFNI. Heatmap refers to *t* values.

Table 1

Demographics and summary statistics for task-related behavior by sex: Means ± standard deviations (and ranges).

	Boys	**Girls**	***Sex differences***
Demographics	(*n* = 38)	(*n* = 41)	
Age (in years)	13.9 ± 1.4	14.2 ± 1.7	*t* (75.9) = −1.1,
	(11.1–17.7)	(11.3–17.6)	*p* = 0.30
IQ (i.e., FSIQ-2 scores)	(*n* = 38)	(*n* = 41)	*t* (76.4) = −0.04,
	107.4 ± 10.8	107.6 ± 12.7	*p* = 0.97
	(81–127)	(79–132)	
Puberty (i.e., mean PDS score)	(*n* = 36)	(*n* = 41)	*t* (74.8) = −5.60,
	2.4 ± 0.6	3.2 ± 0.6	*p*<.001
	(1.2–3.4)	(1.4–4.0)	
Task-related behavior	(*n* = 39)	(*n* = 40)	
Run completion time (in minutes) (Number of) Go decisions	2.3 ± 0.13	2.3 ± 0.14	*t* (3147.6) = −0.65,
	(2.0−2.6)	(2.0–2.7)	*p* = 0.51
	(*n* = 39)	(*n* = 40)	*t* (76.6) = 0.50,
	15.0 ± 6.8	14.3 ± 6.5	*p* = 0.62
	(2–27)	(1–32)	
Stop decisions	(*n* = 39)	(*n* = 40)	*t* (76.8) = −0.38,
	23.5 ± 7.0	24.1 ± 6.8	*p* = 0.71
	(11–37)	(8–37)	
Go decisions with Good outcomes	(*n* = 39)	(*n* = 40)	*t* (76.9) = 0.16,
	7.1 ± 3.4	7.0 ± 3.6	*p* = 0.87
	(0–14)	(0–17)	
Stop decisions with Good outcomes	(*n* = 39)	(*n* = 40)	*t* (76.6) = −0.53,
	11.7 ± 4.1	12.2 ± 3.9	*p* = 0.60
	(5–20)	(4–19)	
Go decisions with Bad outcomes	(*n* = 39)	(*n* = 40)	*t* (75.1) = 0.73,
	7.9 ± 4.0	7.3 ± 3.5	*p* = 0.47
	(0–15)	(1–15)	
Stop decisions with Bad outcomes	(*n* = 39)	(*n* = 40)	*t* (77.0) = −0.16,
	11.8 ± 3.5	11.9 ± 3.6	*p* = 0.87
	(5–19)	(3–18)	

FSIQ-2, Full-Scale Intelligence Quotient based on 2 subscales - Vocabulary and Matrix Reasoning - of the Wechsler's Abbreviated Scale of Intelligence (WASI); PDS, Pubertal Development Scale

Table 2

Regions of activation for the main effects of Decisions.

Contrast	Region (Brodmann Area)	MNI coordinates (x, y, z)	Cluster size (k_E)	Statistic (t)
Decisions > Driving baseline $k_t = 109$	L Precentral Gyrus (BA 4)	−34, −20, 56	36,092	15.6
	L Middle Frontal Gyrus (BA 10)	−38, 52, 18	1049	7.7
	R Inferior Occipital Gyrus	32, −92, −8	258	7.3
	R Middle Frontal Gyrus	38, 40, 32	1187	7.2
	L Middle Occipital Gyrus	−40, −72, 6	224	6.8
	L Inferior Occipital Gyrus (BA 18)	−26, −96, −10	172	6.0
	R Inferior Temporal Gyrus	68, −26, −22	113	5.1
Driving baseline > Decisions $k_t = 109$	L Orbital Frontal Gyrus	−44, 28, −6	34,477	11.6
	R Superior Temporal Gyrus	60, −4, −12	13,128	9.5
	L Cerebellum	−24, −80, −38	303	4.9
Go > Stop decisions $k_t = 98$	L Postcentral Gyrus (BA 3)	−40, −24, 52	2470	13.8
	L Putamen	−16, 8, −10	1030	9.4
	L Supplementary Motor Area (BA 6)	−4, −6, 54	1401	9.0
	R Caudate	14, 10, −10	1052	7.8
	L Precentral Gyrus	−56, 4, 36	300	6.9
	L Midbrain	−6, −28, −10	457	5.8
	R Postcentral Gyrus	44, −22, 42	210	5.2
	R Superior Frontal Gyrus (BA 6)	30, −10, 62	185	4.8
Stop > Go decisions $k_t = 97$	R Cuneus	14, −90, 18	30, 198	9.1
	R Middle Frontal Gyrus (BA 6)	48, 2, 56	155	6.3
	L Cerebellum	−12, −76, −44	271	5.4
	L Inferior Frontal Gyrus (BA 46)	−56, 28, 16	240	5.1
	R Superior Frontal Gyrus	10, 52, 46	700	4.8
	R Supplementary Motor Area (BA 6)	6, 18, 68	104	4.5

Note: Results reported here exceeded the minimum cluster extent threshold needed for a 0.05 family-wise error (FWE) rate given a voxel-wise threshold of $p = 0.001$. The cluster extent threshold (k_t) for each contrast was calculated using 3dClustSim in AFNI. Number of participants = 79 for all contrasts.

Table 3

Regions of activation for the main effects of Outcomes.

Contrast	Region	MNI coordinates (x, y, z)	Cluster size (k_E)	Statistic (t)
Outcomes > Driving baseline $k_t = 107$	R Superior Temporal Gyrus	66, −22, 8	52,257	17.1
	Midbrain	0, −32, −4	174	6.5
	L Precentral Gyrus	−36, −2, 48	351	5.5
	R Corpus Callosum/Posterior Cingulate Gyrus (BA 23)	4, −24, 26	121	5.3
Driving baseline > Outcomes $k_t = 107$	L Postcentral Gyrus	−36, −22, 46	29,139	14.7
	L Inferior Parietal Gyrus (BA 19)	−38, −76, 42	2118	14.2
	L Cuneus (BA 18)	−26, −102, −6	616	11.4
	R Inferior Occipital Gyrus (BA 18)	24, −102, −4	406	11.0
	R Angular Gyrus	50, −68, 44	548	10.1
	L Precentral Gyrus (BA 6)	−56, 2, 32	257	9.3
	R Cerebellum	16, −86, −38	1595	8.7
	R Inferior Temporal Gyrus	64, −8, −28	675	8.1
	R Cerebellum	12, −48, −16	607	7.3
	R Cerebellum	6, −48, −46	263	6.9
	R Lateral Ventricle/Hippocampus	36, −42, 0	394	6.2
	R Insula	46, −20, 22	108	5.6
Good outcomes > Driving baseline $k_t = 100$	R Superior Temporal Gyrus	66, −22, 8	49,912	16.5
	R Cerebellum	16, −50, −50	319	7.0
	L Midbrain	−2, −32, −4	157	6.0
	L Precentral Gyrus	−36, −2, 48	436	5.4
Driving baseline > Good outcomes $k_t = 100$	L Precentral Gyrus	−38, −14, 54	10,439	13.7
	L Inferior Parietal Gyrus	−38, −76, 44	1792	12.9
	R Inferior Occipital Gyrus (BA 18)	24, −102, −4	381	10.1
	L Middle Occipital Gyrus (BA 18)	−24, −102, −4	525	9.5
	R Angular Gyrus	48, −68, 44	395	8.9
	L Medial Orbitofrontal Gyrus	−2, 60, −14	3582	8.8
	R Postcentral Gyrus (BA 3)	46, 22, 40	2204	8.8

Contrast	Region	MNI coordinates (x, y, z)	Cluster size (k_E)	Statistic (t)
	L Precentral Gyrus (BA 9)	−56, 4, 30	198	8.7
	L Inferior Temporal Gyrus (BA 37)	−58, −50, −12	2559	8.6
	R Cerebellum	16, −86, −38	1005	7.7
	R Parahippocampal Gyrus (BA 28)	20, −14, −20	321	7.2
	L Hippocampus	−22, −20, −18	465	7.1
	R Putamen	26, 12, −2	844	7.1
	R Inferior Temporal Gyrus	64, −10, −28	519	6.7
Bad outcomes > Driving baseline $k_t = 93$	R Superior Temporal Gyrus (BA 42)	66, −22, 10	46,711	15.9
	R Cerebellum	28, −40, −46	358	7.2
	R Corpus Callosum	4, −28, 24	102	5.3
	L Precentral Gyrus	−36, −2, 48	181	4.6
Driving baseline > Bad outcomes $k_t = 93$	L Postcentral Gyrus	−34, −22, 46	32,896	14.9
	L Precuneus (BA 19)	−40, −78, 42	2111	12.7
	L Cuneus (BA 18)	−24, −102, −6	654	11.4
	R Angular Gyrus	48, −72, 40	588	10.3
	R Inferior Occipital Lobe (BA 18)	24, −102, −6	397	9.6
	R Middle Temporal Gyrus	64, −8, 26	548	8.6
	R Cerebellum	14, −86, −38	1848	8.2
	R Cerebellum	20, −50, −24	569	7.5
	R Cerebellum	6, −48, −46	319	7.4
	R Insula	46, −20, 22	110	5.3
	R Inferior Temporal Gyrus	68, −46, −12	104	4.4
Good > Bad outcomes $k_t = 94$	R Putamen	20, 8, −8	2151	7.0
	L Middle Temporal Gyrus	−62, −14, −2	1091	6.6
	R Superior Temporal Gyrus	60, −6, −10	623	5.4
	R Postcentral Gyrus	28, −38, 72	257	4.7
	L Paracentral Lobule	−10, −30, 64	176	4.4
	R Middle Frontal Gyrus	22, 28, 38	138	4.4
	R Precuneus	10, −52, 12	94	4.3
Bad > Good outcomes $k_t = 94$	R Lingual Gyrus	18, −84, −6	276	5.4

Contrast	Region	MNI coordinates (x, y, z)	Cluster size (k_E)	Statistic (t)
	L Lingual Gyrus	−18, −82, −6	163	4.1

Note: Results reported here exceeded the minimum cluster extent threshold needed for a 0.05 family-wise error (FWE) rate given a voxel-wise threshold of $p = 0.001$.

The cluster extent threshold (k_t) for each contrast was calculated using 3dClustSim in AFNI. Number of participants = 79 for all contrasts.

Table 4

Regions of activation for Outcomes modulated by Decisions.

Contrast	Region	MNI coordinates (x, y, z)	Cluster size (k_E)	Statistic (t)
Good outcomes (Stop > Go)	L Postcentral Gyrus	−38, −22, 52	761	6.7
N = 77	L Supplementary	−4, −8, 52	100	5.0
k_t = 80	Motor Area (BA 6)			
	R Putamen	18, 8, −10	141	4.5
	R Inferior	58, −56, 44	86	4.3
	Parietal Gyrus			
Good outcomes (Go > Stop)	R Lingual Gyrus	8,−84, −8	22,852	12.9
	L Inferior Frontal	−50, 36, 10	403	6.5
N = 77	Gyrus			
k_t = 89	R Inferior Frontal Gyrus (BA 46)	52, 40, 4	2202	6.5
	R Anterior Cingulum	2, 28, 22	1670	6.2
	L Superior Temporal Gyrus	−60, −26, 14	886	5.8
	L Precentral Gyrus (BA 6)	−44, 2, 58	259	5.7
	L Inferior Frontal Gyrus	−42, 12, 20	92	5.5
	L Inferior Frontal Gyrus	−24, 36, −10	225	5.3
	L Superior Temporal Pole (BA 38)	−48, 16, −10	912	5.1
	L Cingulate Gyrus	−8, 0, 30	89	5.1
	R Inferior Temporal Gyrus (BA 20)	40, −2, −44	128	5.1
	L Superior Parietal Gyrus	−20, −46, 72	162	4.6
	R Superior Orbitofrontal Gyrus	24, 34, −14	184	4.6
	R Superior Parietal Gyrus	26, −54, 66	193	4.0
Bad outcomes (Stop > Go)	L Postcentral Gyrus (BA 3)	−40, −24, 52	3970	8.7
N = 78	R Putamen	16, 8, −10	675	8.7
k_t = 90	L Putamen	−18, 6, −10	572	7.1
	L Sub-Gyral	−28, −50, 2	164	6.9
	L Supplementary Motor Area (BA 6)	−4, −8, 52	223	6.2
	L Middle Frontal Gyrus (BA 9)	−44, 30, 34	194	5.9
	R Inferior Parietal Gyrus (BA 40)	56, −60, 38	499	5.3
	R Cerebellum	24, −90, −38	255	5.1
	L Cuneus	−18, −90, −42	211	5.0
	L Superior Frontal Gyrus (BA 6)	−22, 18, 62	114	4.6
	L Middle Frontal Gyrus	−44, 54, 4	141	3.7
Bad outcomes (Go > Stop)	R Lingual Gyrus	12, −88, −10	33,663	13.3
N = 78	Anterior	0, 28, 22	7951	9.9
k_t = 102	Cingulate			
	L Insula (BA 13)	−38, 8, −8	2155	7.3
	R Insula	38, −8, −8	1306	6.8

Contrast	Region	MNI coordinates (x, y, z)	Cluster size (k_E)	Statistic (t)
	L Precentral Gyrus	−44, 0, 58	181	6.8
	L Supplementary Motor Area (BA 6)	−8, 8, 72	537	6.5
	R Inferior Frontal Gyrus (BA 45)	56, 38, 6	488	6.5
	R Middle Frontal Gyrus (BA 6)	46, 2, 56	532	6.1
	L Fusiform Gyrus	−28, 2, −44	122	5.0
Good outcomes (Stop > Go) > Bad outcomes (Stop > Go)	L Middle Temporal Gyrus (BA 39)	−50, −72, 10	1191	9.0
	R Middle	44, −60, 10	902	7.9
$N = 76$	Temporal Gyrus			
$k_t = 80$	L Calcarine Gyrus	−10, −92, 2	1205	6.6
	R Precuneus (BA 7)	10, −52, 64	866	6.5
	L Supramarginal Gyrus (BA 40)	−56, −30, 24	727	6.2
	R Supramarginal Gyrus (BA 40)	54, −30, 32	223	5.4
	R Cuneus (BA 7)	20, −78, 42	102	5.2
	R Superior Occipital Gyrus (BA 18)	16, −88, 18	179	5.0
	L Middle Cingulate Gyrus	−8, −22, 44	107	4.6
	R Calcarine Gyrus	20, −98, 2	93	4.3
Bad outcomes (Stop > Go) > Good outcomes (Stop > Go)	R Lingual Gyrus (BA 18)	6,−82,−8	251	5.1
	L Inferior Parietal Gyrus	−52, −60, 48	230	5.0
$N = 76$				
$k_t = 90$	R Middle Temporal Gyrus	56, −28, 0	382	4.9
	R Precentral Gyrus	28, −18, 70	104	4.5
	L Middle Frontal Gyrus	−28, 24, 54	138	4.1

Note: Results reported here exceeded the minimum cluster extent threshold needed for a 0.05 family-wise error (FWE) rate given a voxel-wise threshold of $p = 0.001$. The cluster extent threshold (k_t) for each contrast was calculated using 3dClustSim in AFNI. Number (N) of participants varies by contrast.