The assessment of affective decision-making: Exploring alternative scoring methods for the Balloon Analog Risk Task and Columbia Card Task

Received: 29 January 2023 | Revised: 29 November 2023 | Accepted: 3 December 2023

DOI: 10.1002/bdm.2367

RESEARCH ARTICLE

WILEY

The assessment of affective decision-making: Exploring alternative scoring methods for the Balloon Analog Risk Task and Columbia Card Task

Stjepan Sambol | Emra Suleyman | Michelle Ball

Institute for Health and Sport (IHES), Victoria University, Melbourne, Victoria, Australia

Correspondence
Stjepan Sambol, Institute for Health and Sport (IHES), Victoria University, Melbourne, Victoria, Australia.
Email: stefan.sambol@vu.edu.au

Funding information
No grants or fellowships were used to fund this work.

Abstract

Affective decision-making (ADM) is recognized as the ability to effectively reappraise stimuli during these decisions to make choices that maximize long-term outcomes. Currently, the Iowa Gambling Task (IGT) is the gold-standard measure of ADM. Previous research has shown that other commonly used decision-making tasks such as the Balloon Analog Risk Task (BART) and Columbia Card Task (CCT) are unrelated to the IGT and may assess distinct decision-making constructs from ADM. Yet the exact decision-making constructs that these tasks assess may be dependent on the scoring method utilized. One-hundred and eight-four participants (18–58 years; $M = 26.29$, $SD = 7.79$) completed the IGT, BART, and CCT. The relationships between these tasks while utilizing both traditional and novel scoring methods for the BART and CCT were investigated. Results showed that whether using the novel or traditional scoring methods, the BART failed to produce any meaningful relationships with the IGT or CCT. The BART may capture unique decision-making processes involved during conditions of uncertainty, whereas the other tasks involve decision-making processes under conditions of known risk. Alternatively, the lack of meaningful relationships may be due to the stochastic design of the BART. Conversely, the novel and traditional scoring methods for the CCT, which were not significantly correlated with each other, were both related to the IGT. Ultimately, this study showed that the CCT can capture different decision-making constructs depending on the scoring methods used. The traditional scoring method, the total number of cards flipped, assesses risk propensity, whereas the newly developed optimal–suboptimal difference score assesses ADM.

KEYWORDS
affective decision-making, Balloon Analog Risk Task, Columbia Card Task, Iowa Gambling Task, risk propensity, risk-taking

1 | INTRODUCTION

Decision-making is defined as the process of selecting an option or action out of two or more competing choices (Buelow, 2015). Crucially, these decisions are not reflexive and involve deliberate goal-directed consideration to achieve a specific outcome (Wang, 2008). It has been suggested that all decision-making contains an emotional ("hot") component and a cognitive ("cool/cold") component (Buelow & Blaine, 2015). The hot component refers to the emotions experienced as a function of uncertainty regarding the potential consequences or the risk of attaining negative outcomes inherent within the decision-making process (Wang, 2008). Conversely, the cool component purely represents the cognitive information processing that occurs during the decision-making process, independent of emotional processing (Peters et al., 2006).

Cool decision-making processes are purported to be associated with the attainment of logical outcomes (Timmer et al., 2021). However, such characterizations are problematic since hot and cool decision-making processes cannot be disentangled (Figner et al., 2009). Emotions, as described by Lerner et al. (2015), have a persistent influence on decision-making, albeit to a varying degree. Our current emotional state due to the decision or incidental influences, perception of consequences or expected outcomes, and level of motivation are all factors that influence decision-making (Lerner et al., 2015), alongside any deliberate cost–benefit analyses (i.e., cool component). Given the persistent presence of emotions during decision-making there is considerable individual variability in how people approach decisions, particularly those imbued with highly emotional circumstances (e.g., financial, or health-related decisions). Subsequently, researchers are interested in how individuals approach affective decision-making (ADM)—that is the ability to flexibly appraise whether to approach or avoid an option or action, with the intention of maximizing optimal outcomes (Zelazo & Carlson, 2012).

The work of Edmund Rolls (2019) has demonstrated that ADM is dependent on an intact orbitofrontal cortex (OFC). The OFC (comprised Brodmann areas 11, 12, and 13) receives inputs from cortical sensory regions that process information independent of emotional valence (e.g., primary taste cortex, primary olfactory cortex, inferior temporal visual cortex, temporal cortical auditory areas, and somatosensory cortex; Dixon et al., 2017; Rolls, 2019). These inputs projected to the OFC are then rapidly appraised, which involves determining the expected outcome of a perceived stimulus/action and whether it is a reward or punishment, pleasant or unpleasant (Rolls, 2019). Once emotional valence has been ascribed to a stimulus/action the OFC projects this information to the ventromedial prefrontal cortex (VMPFC) and cingulate cortex to guide decision-making (Rolls, 2019). Crucially, the OFC can rapidly override learned stimuli-reward associations should choice-relevant stimuli be incorrectly interpreted (Rolls, 2019), which allows for prudent behavioral change—the central component of successful ADM (Zelazo & Carlson, 2012). Supportively, increased activation of the OFC has been directly linked to the magnitude of anticipated and received rewards or punishments (Kahnt et al., 2010; Roesch & Olson, 2007; Schnider et al., 2005). Additionally, OFC lesion patients have repeatedly demonstrated insensitivity to rewards or punishments, and displayed poorer decision-making, often selecting unfavorable choices that do not serve long-term goals and show no sign of reverse learning (i.e., adjusting from irrational choice selection; Bechara et al., 1994; Fellows, 2007; Howard & Kahnt, 2021). Beyond, the neuroanatomical context, it is important to acknowledge that there are tangible differences in ADM among neurotypical individuals. Variations in ADM can have considerable real-world consequences, manifesting in behaviors such as problem gambling (Brevers et al., 2013) and substance abuse (Buelow & Suhr, 2009). Therefore, a thorough understanding and precise assessment of ADM is essential to provide insight into how people may vary in their responses to rewards, punishments, and their overall decision-making strategies.

To date, the Iowa Gambling Task (IGT) is the most frequently used measure of ADM (Buelow & Blaine, 2015). Bechara et al. (1994) developed the IGT to create a laboratory task that mimics real-life decision-making. Participants are required to select a card from one of four decks where they could win or lose money; however, the probability associated with each deck is initially unknown. Successful performance is demonstrated through participants learning to reappraise the motivational significance of the two initially desirable high-reward decks, which, if selected consistently over time, result in greater losses and switch to selecting from the low-reward decks as these produce an overall net benefit (Bechara et al., 1994). The total number of times participants select from the low-reward decks in the last 60 trials is used as a measure of ADM. The initial 40 trials are discarded from the score since the risks associated with each deck are initially unknown to participants; thus, their choices reflect decision-making under conditions of uncertainty. However, the subsequent 60 trials, where participants have had the opportunity to learn the risks associated with each deck, capture their decision-making under known risks. Crucially, this scoring method successfully captures the core component of ADM, that is, a person's ability to reappraise stimuli and change their behavior to maximize long-term rewards (Zelazo & Carlson, 2012). Critically, the IGT was one of the key tasks used to help identify the role of the OFC in ADM (Bechara et al., 1994). Patients with OFC lesions display stagnated decision-making strategies on the IGT, inflexibly shifting away from the high-risk deck resulting in more disadvantageous choices and an overall net loss (Bechara et al., 1994; Fellows, 2007; Ouerchefani et al., 2017). Additionally, greater activation of the OFC during the IGT has been linked to shifting from the high-risk to the low-risk deck resulting in more advantageous choices and an overall net gain (i.e., superior ADM; Li et al., 2010; Zha et al., 2022).

Since the inception of the IGT, the task has been considered the gold-standard ADM measure and has demonstrated sound ecological validity (Brevers et al., 2013). It has been widely used across various populations with significant real-world implications. For example, problem-gamblers (Brevers et al., 2013), patients with substance abuse (Buelow & Suhr, 2009), violent criminals (Umbach et al., 2019), and children with foetal alcohol syndrome (Kully-Martens et al., 2013), have all showed impaired decision-making on the IGT when compared to control groups. Furthermore, IGT performance has been shown to be predictive of both gambling behaviors and diagnostic criteria for gambling disorders (Brevers et al., 2013). More recent meta-analytic evidence has further illustrated the importance of ADM assessment. Medium to large effect sizes were observed when comparing performance on the IGT between healthy controls to obese individuals (Rotge et al., 2017), patients with alcohol use disorder or gambling disorder (Kovács et al., 2017), and patients with schizophrenia (Betz et al., 2019). Therefore, the assessment of ADM provides valuable information on potential vulnerabilities individuals may face with decision-making in their daily lives. Nonetheless, despite demonstrated validity, the IGT has been criticized for its poor test–retest reliability ($r = .51$; Buelow & Barnhart, 2018), although this could be attributed to learning effects which are central to task performance. Once participants realize particular choices involve greater risk, they typically modify their behavior and select more from the low-risk decks, improving their performance compared to previous attempts (Buelow & Barnhart, 2018). When test–retest reliability is assessed over a longer period (i.e., greater than a year) reliability properties have been shown to improve ($r = .74$; Tuvblad et al., 2013).

There are many gambling-based decision-making tasks that provide participants with choices that could result in reward or punishment (Weber & Johnson, 2009), although few accurately assess ADM. These tasks produce greater levels of emotional salience compared to other cognitive tasks by emphasizing choices with the potential for gain or loss, increasing the engagement of cortical regions such as the OFC (Løvstad et al., 2012). However, for a task to accurately assess ADM the decision-making process should involve conditions of risk, where the outcomes and probabilities are apparent through experience, rather than conditions of uncertainty, where the outcome probabilities are unknown. This dynamic allows researchers to evaluate whether participants can flexibly appraise choices and pursue more prudent decisions (Bechara et al., 1994; Zelazo & Carlson, 2012). Few gambling-based tasks have scoring protocols that capture participants' ability to reappraise stimuli. Thus, ADM has been predominately measured through the IGT. This limits our understanding of construct validity and the factors that influence ADM as assessment has been mostly constrained to the task-specific context of the IGT.

The Balloon Analog Risk Task (BART) is another widely used decision-making task that has sometimes been viewed as analogous to the IGT (Buelow & Blaine, 2015), although it is predominately used as a measure of risk-propensity (i.e., an individual's orientation towards taking or avoiding risk; Canning et al., 2022; Wang et al., 2022). During the BART participants are given the dichotomous choice to either pump a virtual balloon using a keyboard space bar, accruing virtual money with each pump, or stop and collect the amount accumulated (Lejuez et al., 2002). Communicated to the participants is that the chances of the balloon popping increase with each subsequent pump, which would result in losing their accumulated money for the current trial. Previous research has demonstrated convergent validity for the BART as performance has been associated with self-reported, real-world risk-taking behavior (e.g., substance abuse and gambling; Hopko et al., 2006; Hunt et al., 2005; Lejuez, Aklin, Jones, et al., 2003; Lejuez, Aklin, Zvolensky, et al., 2023), and self-reported sensation-seeking and impulsivity (Buelow & Blaine, 2015; Xu et al., 2013). The BART has also shown sound test–retest reliability across a 2-week period ($r = .77$; White et al., 2008). Conflicting results, however, call the validity of the BART into question. Firstly, an analysis of the effect size across 22 studies found the relationship between the BART and sensation-seeking was small–medium, and the relationship between the BART and impulsivity was small (Lauriola et al., 2014). Furthermore, performance on the BART has been inconsistently related to alcohol use (Canning et al., 2022) and was not found to be positively related to problematic social media usage (Meshi et al., 2020).

The BART is typically scored using the adjusted average number of pumps, representing the average number of pumps on trials where money was collected (Lejuez et al., 2002). While this scoring method excludes trials where balloons popped, under the rationale that popped balloons prematurely limit the amount of risk participants can take, it may not capture ADM. This is partly because decisions within the BART are made under conditions of uncertainty due to the unknown and randomized pop probability, differing from decision-making under known risk. It is crucial to recognize that with the risk of balloon pops being unknown, participants' decision-making is primarily informed by their experiences of both successful and unsuccessful trials. Since these experiences vary and are not standardized, individuals who encountered an early balloon pop might adapt their approach to the task differently compared to those who experienced a series of successful trials (Kessler et al., 2017). Subsequently, the learning conditions within the BART are inherently based on these random experiences, leading each participant to develop a unique understanding of the risk–reward balance. This individualized perception, in turn, shapes their subsequent decision-making strategy. Supportively, Di Plinio et al. (2022) showed that once randomness was controlled for by administering fixed-probability trials, the BART more accurately predicted participants' risk-taking profiles. This finding suggests that the task's stochastic design may be one possible explanation for the conflicting results regarding its validity. Furthermore, while adaptation during the BART can occur following both balloon pops and successful trials, the standard measure of the adjusted average number of pumps remains a problematic measure of ADM. This scoring method may not adequately capture how participants reappraise and adjust their strategies throughout the task (De Groot, 2020; Di Plinio et al., 2022).

An alternative scoring measure for the BART that may more accurately reflect ADM is "post-pump loss," originally proposed by Schmitz et al. (2016). This scoring method for the BART seeks to circumvent the task's stochastic design by providing participants a score that is relative to their experience instead of a score influenced by uncertainty surrounding the balloon pop probability. Post-pump loss is calculated by averaging the difference between the number of pumps on a loss trial and the immediate subsequent trial where participants elected to collect money prior to a balloon pop. Rather than providing the magnitude of risk-taking, similar to the IGT, this scoring method captures reappraisal following a negative consequence (i.e., balloon pop) by quantifying whether a person responds by engaging in impulsive compensatory behavior (signified by an increased number of pumps) or displays superior ADM through reducing the number of pumps. Crucially, by only utilizing trials immediately after a balloon pop, this scoring method seeks to evaluate if participants incorporate their experiences from prior trials as known risks. This contrasts with the conventional scoring method of the BART, which utilizes all successful collect trials and consequently is more influenced by decisions made under conditions of uncertainty. Neuroimaging studies have shown the greatest level of OFC activation followed a balloon pop (Schonberg et al., 2012; Wang et al., 2022), which likely represents participants reappraisal of the task following negative consequences. Post-loss pumps are arguably the ideal scoring method to assess ADM, as the score attempts to capture behavioral changes following a balloon pop.

One of the newest gambling-based decision-making tasks is the Columbia Card Task (CCT; Figner et al., 2009), which has both hot and cool versions. This computerized task comprises 32 face-down cards which participants flip over to reveal a potential gain or loss. In the cool version participants are asked to make a singular decision at the start of each trial on the number of cards they will flip over, whereas the hot version involves a stepwise decision-making process where the participants flip over one card at a time (Figner et al., 2009). The novelty of the CCT is the fluctuating levels of gain (e.g., \$10 or \$30), loss (e.g., −\$250 or −\$750), and probability (e.g., 1 or 3 loss cards) presented to participants within each trial. This allows researchers to gain a deeper understanding of participants' considerations when making decisions under conditions of known risk (Haffke & Hübner, 2020; Kluwe-Schiavon et al., 2020). The CCT has not been as extensively utilized in research compared to the IGT and BART (Buelow, 2015), although recent studies have demonstrated that cocaine-dependent users (Kluwe-Schiavon et al., 2020), VMPFC lesion patients (Spaniol et al., 2019), and sleep-deprived individuals (Salfi et al., 2020) all displayed significantly greater risk propensity on the CCT compared to control counterparts.

The primary scoring method for the CCT is the total number of cards flipped (Figner et al., 2009). While this scoring method provides insightful information into participants' risk propensity, the total number of cards flipped is problematic for assessing ADM as it is not necessarily a sensitive measure of a person's ability to reappraise stimuli and decide whether to approach or avoid decisions with the intention of maximizing reward. The total number of cards flipped fails to capture nuances in decision-making across the CCT's varying levels of gain, loss, and probability. For example, this scoring method cannot distinguish between circumstances when a person flipped a consistent number of cards across all trials compared to a person who elects to flip a few cards during trials when the gain amount was low and more cards during trials when the gain amount was high. To compensate for this limitation, Figner et al. (2009) proposed the information use score, which identifies how many of the fluctuating trial criteria participants responded to. The score is calculated by conducting an ANOVA at an individual/subject-level to determine if there is a significant difference in the number of cards flipped when a trial criterion (i.e., gain, loss, or probability) is low compared to high. For each significant difference identified participants receive one point. Therefore, information use scores range from 0 to 3, with the minimum score indicating a participant accounted for none of the trial criteria in their decision-making, and the maximum score indicating they factored in all trial criteria (Figner et al., 2009). Unfortunately, by relying on identifying all-or-nothing statistical differences, this scoring system does not capture the magnitude to which individuals utilized the gain, loss, or probability criteria.

Potentially, the difference between the average number of cards flipped on optimal trials (i.e., trials with at least two of the following criteria: low loss, high gain, low probability) compared to suboptimal trials (i.e., trials with at least two of the following criteria: high loss, low gain, high probability) could provide a superior measure of ADM for the CCT. The proposed "optimal–suboptimal difference," rather than capturing participants' risk propensity, captures the degree to which participants modify their decision-making when presented with favorable or unfavorable conditions across the varying levels of gain, loss, and probability. Higher scores would indicate superior ADM as more cards were selected during optimal trials, when odds were favorable, compared to suboptimal trials where odds are unfavorable. While the impact of randomness cannot be completely removed from the CCT, by creating an average across all optimal/suboptimal trials the effect of randomness will potentially be distributed across the two trial conditions, and therefore mitigated.

Despite the fact that Buelow and Blaine (2015) theorized the IGT, BART, and CCT assess ADM, the results of their exploratory factor analysis found limited evidence that these tasks measure the same construct. Their study found that the BART and CCT loaded together weakly, whereas the IGT loaded separately. Buelow and Blaine (2015) therefore suggested these measures assess unique components of decision-making; however, there were two methodological issues that could have impacted the results of the factor analysis. Firstly, performance on the IGT and the BART were divided into five 20-block trials, and three 10-block trials, respectively, which were then input as separate variables into the factor analysis. This unconventional scoring method is problematic as the varimax rotation used attempts to associate each variable with a single factor (Floyd & Widaman, 1995). This could be the reason why the IGT blocks loaded together on a single factor separate from the BART blocks, since performance on one block of a task is closely associated with subsequent blocks of that same task (e.g., block 1 of the BART would likely share a stronger relationship with other blocks from the BART compared to blocks from the IGT due to overlapping methodology). Thus, the relationships between tasks were likely suppressed by the stronger relationships within the blocks of a task.

Secondly, Buelow and Blaine (2015) tested the relationship between the tasks using conventional scoring methods. As previously discussed, unlike the IGT, the original scoring methods used for the BART and CCT fail to capture the core component of ADM; the capacity to reappraise emotionally salient stimuli and adapt to make more advantageous decisions. As a result, these scoring methods which measure the frequency of specific behaviors (e.g., number of pumps/cards flipped) likely assess risk propensity. In the case of the BART, risk propensity is assessed within a static setting as participants are presented with indistinguishable trials where the probability of the balloon popping is unknown and they are given a dichotomous decision to either pump the balloon or collect the money (Weber & Johnson, 2009). Conversely, in the case of the CCT, risk propensity is assessed within a dynamic setting as the probability of loss fluctuates across trials and is known by participants (Haffke & Hübner, 2020). Utilizing the alternative scoring methods previously discussed (i.e., post-loss pumps for the BART and optimal–suboptimal difference for the CCT) could allow these tasks to more appropriately assess ADM.

Decision-making is a pervasive component of everyday life, with adults estimated to make 35,000 decisions a day (Pignatiello et al., 2020). ADM in particular is a vital construct as it directly relates to our capacity to behave in a prudent manner during emotional circumstances (Zelazo & Carlson, 2012). The IGT is currently the only task with a theoretically grounded scoring method for ADM, while other tasks such as the BART and CCT have been used as presumed measures of ADM. Therefore, it is important to delineate the precise decision-making construct assessed by these three prevalent laboratory decision-making tasks. Identifying whether the IGT, BART, and CCT assess distinct or separable decision-making constructs will provide clarity for future studies aiming to investigate different decision-making mechanisms and how these impact real-world judgments and choices.

2 | THE PRESENT STUDY

With modified scoring methods the BART and CCT may also be capable of measuring ADM, allowing these tasks to provide further information on construct validity and the factors that influence ADM, beyond the task-specific context of the IGT. Therefore, this study aimed to investigate the relationships between the IGT, BART, and CCT using both traditional and alternative scoring methods for the BART and CCT designed to more effectively capture ADM. It was hypothesized that when traditional scoring methods were used there would be a significant relationship between the BART and CCT, but not between these measures and the IGT. It was further hypothesized that the IGT, BART, and CCT, would be significantly related when alternative scoring methods were used on the BART (i.e., post-loss pumps) and CCT (i.e., optimal–suboptimal difference).

3 | METHOD

3.1 | Participants

A sample of 184 participants aged between 18 and 58 years were recruited from the Greater Melbourne (Victoria, Australia) area through convenience and snowball sampling methods via electronic advertisements posted on social media (e.g., Facebook) and on university notice boards. An a priori G-power analysis, with alpha set at .05, statistical power at .80, and anticipating a small to medium effect size (i.e., 0.2–0.3), recommended a minimum sample of 84 to 193 participants to conduct two-tailed correlation analyses. Participants were recruited between December 2020 and December 2021, and inclusion criteria for the current study required participants to be aged between 18 and 60 years, with normal or corrected to normal vision, and no currently diagnosed psychological disorder (e.g., mood disorders and anxiety disorders) or pre-existing neurological disorder (e.g., Alzheimer's and Dementia and ADHD). Participants' drug and alcohol use was not assessed during this study. Demographic information is presented in Table 1. For taking part in the study each participant was compensated with a $20 gift card and an optional brief personality report (that was produced as part of a larger study). To mitigate potential fatigue effects, all tasks were administered in a counterbalanced order. Participants were randomly assigned to complete the tasks in one of three predetermined sequences (see Table 2).

3.2 | Materials

3.2.1 | Demographics questionnaire

A demographics questionnaire was administered via the online platform Inquisit Version 6 (Millisecond Software, 2022), collecting information on participants' eligibility to participate, age, gender, and

TABLE 1 Participant demographic information ($N = 184$).

Variable	M (SD)	n	%
Gender			
Men		78	42.4
Women		105	57.1
Non-binary		1	0.5
Age	26.29 (7.79)		
Education			
High school		60	32.6
Certificate 3 or 4		19	10.3
Diploma or advanced diploma		16	8.7
Bachelor's or honors degree		85	46.2
Master's degree		3	1.6
Doctorate		1	0.5

TABLE 2 Counterbalanced test orders used in the current study ($N = 184$).

	Order A $n = 61$	Order B $n = 63$	Order C $n = 60$
1)	IGT	BART	CCT
2)	CCT	IGT	BART
3)	BART	CCT	IGT

Abbreviations: BART, Balloon Analog Risk Task; CCT, Columbia Card Task; IGT, Iowa Gambling Task.

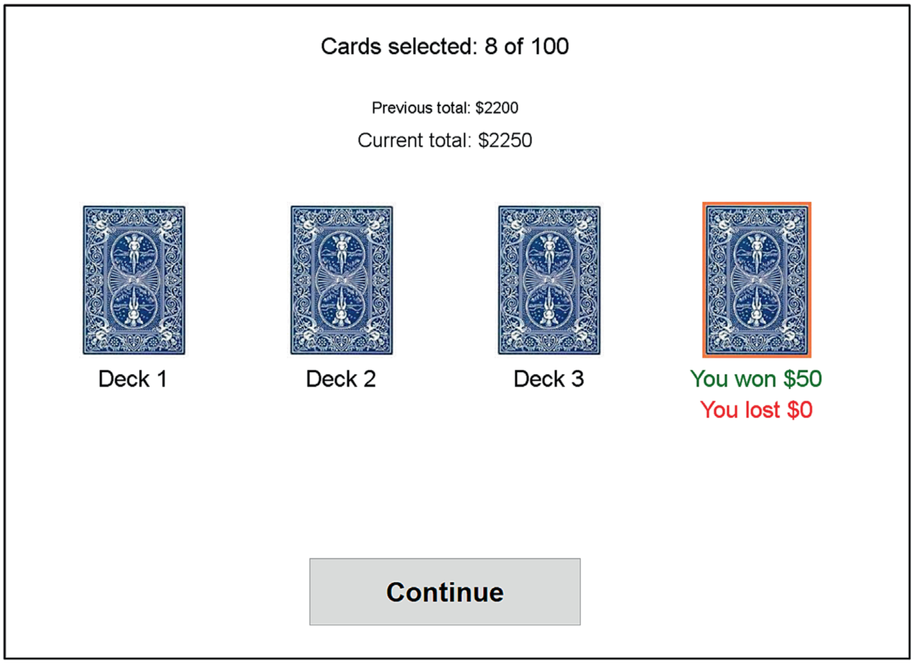

FIGURE 1 Screenshot of the IGT used in the current study.

education. Participants who self-reported a current psychological or pre-existing neuropsychological condition or aged below 18 or above 60, were informed they were ineligible to participate in the current study and thanked for their time.

3.2.2 | IGT

The IGT consists of five blocks of 20 trials that present participants with four decks of cards. In each trial, participants are asked to select a card from one of the four decks (see Figure 1). Two of the decks provide larger gains (e.g., $100); however, there is a great risk of larger losses (e.g. −$1250), whereas the other two decks provide smaller gains (e.g., $50) and smaller losses (e.g. −$250). The cards within each deck are randomized without replacement to bias outcomes. Repeatedly selecting the decks with larger gains will result in an overall net loss, while selecting from the decks with smaller gains will result in an overall net gain. Each participant starts with a hypothetical $2000, and they are told that the goal is to maximize profit. The total number of times participants selected from the two advantageous decks in the final three blocks (i.e., last 60 trials) provides a measure of ADM, with higher scores demonstrating superior ADM as they show adjustment to repeated significant negative consequences from the disadvantageous decks (Bechara et al., 1994).

3.2.3 | BART

The BART consists of three blocks of 10 trials where participants are tasked with maximizing profit across three blocks of 10 trials. In each trial, participants are presented with a simulated balloon and choose to either pump the balloon, accruing $1 for each pump, or at any point collect the hypothetical money accumulated for that trial contributing to their overall money earned (see Figure 2). However, with each consecutive pump, excluding the initial two pumps, there is a chance the balloon will pop (i.e., 1/18, 1/17, and 1/16) resulting in no money being earned for that trial. The algorithm, unbeknownst to participants, places the average explosion point at 11 pumps, guaranteeing the balloon will pop at the twentieth pump. This study adopted a modified version of the BART (Éltető et al., 2019), which increased the reward for each pump from $.05 to $1 and shortened the maximum possible pumps for each balloon from 128 to 20 to minimize potential participant fatigue across ADM tasks.

Two sores were obtained from the BART. First, the traditional score adjusted the average number of pumps (i.e., the average number of pumps on successful trials where money was collected) reflecting risk propensity. The adjusted number of pumps removes the constraint in participants' scores from trials where the balloon pops, as these trials end prior to assessing the maximum risk participants were willing to take (Lejuez, Aklin, Jones, et al., 2003; Lejuez, Aklin, Zvolensky, et al., 2023). Second, post-loss pumps (i.e., the average difference between the number of pumps on balloon pop trials and immediately subsequent successful trials where money was collected; Schmitz et al., 2016) assessing ADM. The score captures the relative change in the number of pumps after experiencing a loss. Trials with consecutive balloon pops are discounted from post-loss pump scoring. Participants who on average reduce the number of pumps on trials following a balloon explosion trial receive a negative post-loss pump score. Lower post-loss pump scores are considered to demonstrate superior ADM as participants showed adjustment to negative consequences (i.e., balloon pop) and avoided compensatory risk-taking behavior in the following trials (i.e., increased number of pumps; Schmitz et al., 2016).

FIGURE 2 Screenshot of the BART used in the current study.

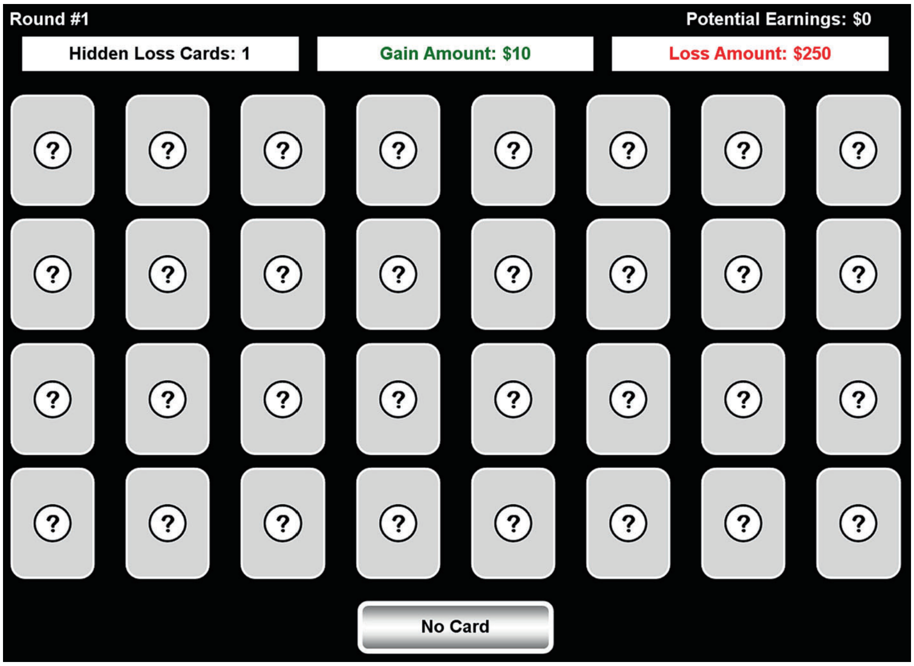

FIGURE 3 Screenshot of the CCT used in the current study.

FIGURE 4 The facedown, gain card, and loss card stimuli for the CCT. *Note*: Question mark indicates facedown card, smiley face indicates gain card, and sad face indicates loss card.

3.2.4 | Columbia Card Task (CCT)

This study utilized a shortened CCT "hot version" (Penolazzi et al., 2012). In each trial, participants are presented with a series of 32 simulated facedown cards in four rows of eight cards (see Figure 3). Each card is either a gain card providing participants with money or a loss card where participants lose money (see Figure 4), with the trial immediately ending when a loss card is turned over. Hypothetical money was used instead of the original points to maintain task salience with the BART and IGT. Participants are instructed to maximize money across the 24 trials by flipping over one card at a time; however, at any point, participants could choose to end the trial without incurring a loss and accumulate the money earned in that trial to their overall total. Trials are presented in a fixed order with each combination of a number of loss cards (1 or 3), amount per gain card ($10 or $30), and loss amount ($250 or $750), presented three times (Penolazzi et al., 2012). Participants are notified of the current trial criteria at the top of their screen.

Three scores were obtained from the CCT. Firstly, the typical scoring method tallies the total number of cards turned over and provides a measure of risk propensity. Second, the information use score, which determines how many trial criteria participants were sensitive to (i.e., gain, loss, probability; see Figner et al., 2009). Lastly, the "optimal–suboptimal difference" novel scoring method was implemented to capture ADM. This score represents the difference between the average number of cards flipped during optimal trials (i.e., trials with the following criteria: 1 loss card, \$30 per gain card, and −\$250 per loss card) compared to the average number of cards flipped during suboptimal trials (i.e., trials with the following criteria: three loss cards, \$10 per gain card, and −\$750 points per loss card). A higher score indicates that on average the participant flipped over more cards during the most advantageous trials compared to the least advantageous trials, thus demonstrating superior ADM.

3.3 | Procedure

The current study received ethics approval from the Victoria University Human Research Ethics Committee. Individuals who expressed interest in participating in the study were emailed a participant information form, detailing the goals and procedure of the study, and sent a hyperlink. The link was to Inquisit Version 6 (Millisecond Software, 2022), which ascertained participants' informed consent and contained the demographic questionnaire and computerized decision-making tasks (IGT, BART, CCT). Participants were instructed to access the link at a time and place most convenient to them and complete the tasks with no enforced time limit.

3.4 | Statistical design

Test performance data was imputed, cleaned, and analyzed using IBM® SPSS® Statistics Version 27. The primary analysis involved investigating the relationship between the three ADM measures (IGT, BART, CCT). First, Pearson's correlation analysis explored the relationship between IGT scores (i.e., adjusted advantageous deck selected), the two BART scores (i.e., adjusted average pumps, and post-loss pumps), and the three CCT scores (i.e., total cards selected, information use score, and optimal–suboptimal difference). These variables were normally distributed with skewness and kurtosis values between −3 and 3, and no standardized values were beyond −3 and 3 indicating no outliers (Field, 2018). However, one participant was missing a post-loss pump score on the BART as they did not experience a balloon pop.

Second, three principal axis factor analyses were performed with Varimax rotation, which attempts to extrapolate the smallest number of factors with the highest correlations between variables and their respective factors whilst minimizing relationships between factors (Floyd & Widaman, 1995). The first factor analysis replicated Buelow and Blaine (2015) and included IGT advantageous minus disadvantageous card selection for blocks 1–5, BART adjusted average pumps for blocks 1–3, and CCT average number of cards selected. The second factor analysis only utilized a single score for each task, following the original scoring method: IGT-adjusted advantageous deck selected, BART-adjusted average pumps, and CCT total cards selected. The third factor analysis utilized the IGT-adjusted advantageous deck selected, and the novel scoring methods for the BART (i.e., post-loss pumps) and CCT (i.e., optimal–suboptimal difference). The KMO statistic was below the acceptable limit of .5 for the second factor analysis (KMO = .457) but was above the limit for the first and third factor analysis (KMO = .717 and KMO = .504, respectively; Kaiser, 1974). Furthermore, Bartlett's test of sphericity was significant for the first $\chi^2(36) = 679.949$, $p < .001$ and the second factor analysis $\chi^2(3) = 23.058$, $p < .001$, supporting the factorability of the correlation matrix, however, was not significant for the third factor analysis $\chi^2(3) = 4.897$, $p = .180$, suggesting that there may be insufficient correlations for factor analysis.

Supplementary analyses were performed to explore the psychometric properties of the newly proposed BART and CCT scoring methods, utilizing publicly available datasets. While the full details of these analyses extend beyond the primary focus of this study, their results offer valuable insights for future research and are therefore included in the supplementary materials for reference and further exploration.

4 | RESULTS

4.1 | Correlation analysis

Results from Pearson's correlation analysis and descriptive statistics for variables assessed in the current study are presented in Table 3. Results from the Person's correlation revealed a significant relationship between the IGT-adjusted advantage and both CCT scoring methods, offering preliminary evidence of convergent validity between these measures. Specifically, a weak negative relationship was observed with total cards flipped, while a weak positive relationship was observed with the optimal–suboptimal difference. Therefore, participants who selected exhibit prudent decision-making on the IGT, by selecting more from advantageous decks, tended to display similar caution on the CCT by flipping fewer cards overall, and flipping more cards during the optimal trials compared to suboptimal trials on the CCT.

When examining the blocks of the IGT individually, performance on blocks 3, 4, and 5 showed weak, negative correlations with the total number of cards flipped on the CCT. Only performance on block 4 demonstrated a weak positive relationship to both the information use and optimal–suboptimal difference score from the CCT. This suggests that participants who are cautious in the latter states of the IGT tend to be similarly cautious on the CCT by flipping fewer cards and being more receptive to suboptimal trial conditions. Conceptually, the relationships between the later trials of the IGT and all three scoring methods used for the CCT suggest that both tasks may be capturing similar aspects of decision-making under conditions of known risk as opposed to under conditions of uncertainty.

TABLE 3 Descriptive statistics and correlations for decision-making tasks ($N = 184$).

Variable	Range	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. IGT A-D1	−16–20	−2.35	6.96	–	.460**	.408**	.287**	.230**	−.368**	.057	.026	.038	.046	−.005	−.098	−.098	.074	−.020
2. IGT A-D2	−20–20	−2.10	8.87		–	.437**	.303**	.342**	−.433**	−.043	−.061	−.048	−.053	−.133	−.076	−.076	.069	−.017
3. IGT A-D3	−20–20	0.33	9.44			–	.609**	.421**	−.806**	.050	.030	−.006	.029	−.044	−.192**	−.192**	.085	.067
4. IGT A-D4	−20–20	2.54	9.12				–	.593**	−.874**	.085	.108	.127	.112	.112	−.250**	−.250**	.186*	.204**
5. IGT A-D5	−20–20	3.80	10.23					–	−.819**	.001	−.092	−.025	−.044	−.004	−.259**	−.259**	.072	.139
6. IGT AA	0–54	26.66	11.98						–	−.053	−.014	−.035	−.035	−.023	.281**	.281**	−.135	−.164*
7. BART 1	1.70–15.50	6.09	2.41							–	.814**	.689**	.905**	.472**	.149*	.149*	.078	.054
8. BART 2	2.20–13.29	6.44	2.36								–	.789**	.946**	.500**	.183*	.183*	.115	.109
9. BART 3	2.00–12.33	6.50	2.22									–	.895**	.580**	.182*	.182*	.059	.081
10. BART AA	1.97–12.08	6.34	2.14										–	.552**	.188*	.188*	.089	.087
11. BART P-LP[a]	−4.00–5.00	0.35	1.52											–	.059	.059	.066	.023
12. CCT Total	72–321	206.20	51.60												–	1.000**	.047	.099
13. CCT Avrg	3.00–13.38	8.59	2.15													–	.047	.099
14. CCT IU	0–3	0.90	0.61														–	.467**
15. CCT O-SD	−8.33–22.67	5.66	5.75															-

Abbreviations: BART, Balloon Analog Risk Task, the average number of balloon pumps within a 10-trial block; BART AA, BalloonAnalog Risk Task, overall adjusted average number of balloon pumps; BART P-LP, BalloonAnalog Risk Task, post-loss pumps; CCT Total, Columbia Card Task, the total number of cards flipped; CCT Avrg, Columbia Card Task, the average number of cards flipped; CCT IU, Columbia Card Task, information use; CCT O-SD, Columbia Card Task, optimal–suboptimal difference. IGT A-D, Iowa Gambling Task, advantageous minus disadvantageous card selection with a 20-trial block; IGT AA, Iowa Gambling Task, adjusted advantageous cards selected. *M*, mean; *SD*, standard deviation.

[a]Correlations with BART P-LP are at $n = 183$.

*Significant at .05 level.

**Significant at .01 level.

There was no significant relationship between either of the BART scoring methods and the IGT, suggesting discriminant validity between the constructs measured by these tasks. While the BART-adjusted average demonstrated a weak positive relationship with the CCT total cards flipped, this likely indicates a general tendency toward risk propensity (i.e., participants with a higher number of pumps on the BART also flipped over more cards on the CCT), and does not necessarily translate to a robust measure of ADM. Notably, the absence of a significant relationship between the BART, particularly when using the "post-loss pumps" scoring method, and both the IGT and the CCT's optimal–suboptimal difference score challenges our initial assertions about the BART's efficacy in capturing ADM in the same vein as the IGT or CCT.

4.1.1 | Exploratory factor analysis

Three EFAs were conducted to further investigate the underlying relationships between the IGT, BART, and CCT, and determine whether these tasks assessed the same decision-making construct. Table 4 presents the factor loadings after varimax rotation for all three EFAs. Following the method ascribed by Buelow and Blaine (2015), the first factor analysis produced a three-factor solution and accounted for 71.150% of the variance. Factor 1 comprised the adjusted average number of pumps from blocks 1, 2, and 3 of the BART. Factor 2 comprised advantageous minus disadvantageous card selection from blocks 4 and 5 of the IGT, and the average number of cards flipped from the CCT. Lastly, factor 3 comprised the remaining blocks from the IGT (i.e., blocks 1, 2, and 3).

The second factor analysis utilizing only the IGT-adjusted advantage, BART-adjusted average, and CCT total cards produced a two-factor solution accounting for a total of 78.513% of the variance. Factor 1 comprised the IGT and CCT, whereas Factor 2 comprised predominately of the BART and accounted for an additional 34.42% of the variance. The third factor analysis substituting novel scoring methods for the BART (post-pump loss) and CCT (optimal–suboptimal difference) produced a one-factor solution accounting for a total of 38.914% of the variance. Varimax rotation could not be conducted due to the single-factor solution, which saw strong loadings from the IGT and CCT, and a weak loading from the BART.

TABLE 4 Component matrix loading of factors extracted by EFAs.

	Rotated loadings		
Factor analysis 1 ($n = 184$)	**Factor 1**	**Factor 2**	**Factor 3**
IGT A-D1	0.037	0.076	**0.798**
IGT A-D2	−0.079	0.134	**0.813**
IGT A-D3	0.058	0.545	**0.588**
IGT A-D4	0.173	**0.772**	0.337
IGT A-D5	−0.018	**0.739**	0.283
BART 1	**0.902**	−0.003	0.010
BART 2	**0.943**	−0.044	−0.010
BART 3	**0.897**	−0.020	−0.006
CCT-Avrg	0.224	−**0.689**	0.172
Eigenvalue	2.745	2.606	1.053
Total variance	30.501%	28.953%	11.897%
	Rotated loadings		
Factor analysis 2 ($n = 184$)	**Factor 1**	**Factor 2**	
IGT AA	**0.857**	0.205	
BART AA	0.003	**0.941**	
CCT Total	−**0.729**	0.402	
Eigenvalue	1.323	1.033	
Total variance	44.093%	34.420%	
Factor analysis 3 ($n = 183$)	**Loadings Factor 1**		
IGT AA	**0.750**		
BART P-LP	0.209		
CCT O-SD	**0.750**		
Eigenvalue	1.167		
Total variance	38.914%		

Abbreviations: BART, Balloon Analog Risk Task, the average number of balloon pumps within a 10-trial block; BART AA, Balloon Analog Risk Task, overall adjusted average number of balloon pumps; BART P-LP, Balloon Analog Risk Task, post-loss pumps; CCT Total, Columbia Card Task, the total number of cards flipped; CCT Avrg, Columbia Card Task, the average number of cards flipped; CCT O-SD, Columbia Card Task, optimal–suboptimal difference; IGT A-D, Iowa Gambling Task, advantageous minus disadvantageous card selection with a 20-trial block; IGT AA, Iowa Gambling Task, adjusted advantageous cards selected.

5 | DISCUSSION

The current study aimed to evaluate the relationships between three decision-making measures the IGT, BART, and CCT. It was further aimed to investigate whether alternative scoring methods designed to capture affective decision-making (ADM), post-loss pumps on the BART, and optimal–suboptimal difference on the CCT would improve relationships with the IGT. The first hypothesis that there would be a significant relationship between the traditional scoring methods for the BART and CCT, but not between these measures and the IGT was partially supported. The traditional scoring methods for the BART and CCT were weakly positively related, however, contrary to this hypothesis a weak negative relationship was found between the total number of cards flipped on the CCT and the adjusted advantage score on the IGT. The second hypothesis also found partial support. The alternative CCT score (optimal–suboptimal difference) was weakly correlated with the IGT, although the alternative BART score (post-loss pumps) did not significantly correlate with either the IGT or CCT.

The significant relationship between adjusted average pumps on the BART and total cards flipped on the CCT aligned with researchers' expectations. However, the weak strength of this relationship was unexpected. Both scores were theorized to assess risk propensity as they reflect the amount of risk participants were willing to take, through the frequency of pumps or cards flipped (Canning et al., 2022; Figner et al., 2009). Therefore, it stands to reason if an individual was willing to take a substantial risk on one of the tasks, this would also translate to the other. One possibility for the weak relationship between the two tasks could be the static properties of the BART and the dynamic properties of the CCT (Weber & Johnson, 2009). All trials of the BART are identical with no new information emerging throughout the task. Participants are only given the choice to pump the balloon or collect the accumulated reward and are never notified of the probability that a balloon will pop (Lejuez et al., 2002). Conversely, participants are notified of the changing level of gain, loss, and probability across the trials of the CCT. Thus, participants may have employed difference decision-making strategies for the two tasks, as they could use deliberate decision-making processes on the CCT based on known trial criteria (Figner et al., 2009), whereas decisions for the BART are made under conditions of uncertainty (Schonberg et al., 2012). Lastly, the stochastic design of the BART and CCT results in randomly experiencing balloon pops or loss cards, which may have further diminished the relationship between the two tasks (De Groot, 2020).

Surprisingly, CCT's total number of cards flipped was significantly, albeit weakly, related to IGT-adjusted advantage. Buelow and Blaine (2015) suggested that the IGT, BART, and CCT assess different decision-making constructs as they found each task loaded on a distinct factor. However, when we replicated their EFA we showed that the CCT hot version strongly loaded with the last three blocks of the IGT. Initially, the consequences of each deck on the IGT are unknown to the participant and it is not until the third block that researchers suggest that the level of risk becomes apparent (Bechara et al., 1994). For the CCT the level of risk (i.e., probability of encountering a loss card) is known at the start of each trial (Figner et al., 2009). Therefore, the strong loading of the CCT with the latter trials of the IGT suggests that once the consequences of each IGT deck were learned, those who were more likely to select from the advantageous decks flipped fewer cards on the CCT. In other words, those who understood the calculated risk associated with the different IGT decks also displayed a greater understanding of the calculated risk for the CCT. This first EFA was limited by including the individual blocks of the BART and IGT as separate variables. This increases the likelihood of identified factors representing methodological overlap due to stronger associations between blocks of the same task, rather than an underlying cognitive construct.

Subsequently, the second EFA only included a single score from each decision-making task. The results supported a two-factor solution with a similar variance accounted for (78.513%) as the first EFA (71.150%). Factor 1 exclusively contained the BART-adjusted average, and factor 2 contained strong loadings from the IGT-adjusted advantage and CCT total cards flipped. The discovery of consistent relationships between CCT total cards flipped and IGT performance is presumably due to risk-aversive tendencies. Previous research has shown that people typically present as risk-averse (i.e., prefer to avoid a loss than achieve a gain of similar value; Harrison & Rutström, 2008), as cortical activity was greater following a punishment compared to a reward of the same magnitude (Sokol-Hessner et al., 2013). In both the CCT and IGT participants are given choices where there is a risk of losing money and receiving a negative (in debt) score. Crucially the probability of obtaining a negative outcome is known by participants or in the case of the IGT becomes apparent in later blocks due to the fixed deck outcomes. Thus, an innate drive to avoid loss would have resulted in participants flipping fewer cards on the CCT and selecting more from advantageous decks on the IGT once they learned the associate risk of each deck. Risk aversion would not motivate performance on the BART to the same degree as the IGT and CCT, since the probability of negative outcomes is unknown (De Groot, 2020) and participants never receive a negative (in debt) score, but instead lose a potential reward (Lejuez et al., 2002).

When investigating the relationships between the three decision-making tasks while utilizing alternative scoring methods, the IGT and CCT were still significantly correlated and shared strong factor loadings. Both scores are likely to capture the same ADM construct—the flexible ability to whether approach or avoid a salient stimulus, with the intention of maximizing beneficial outcomes (Zelazo & Carlson, 2012). Participants who recognized the need to transition from the high-reward and high-risk decks to low-reward but low-risk decks to produce a greater reward also recognize the practicality of reducing risk-taking during unfavorable trials (i.e., trials with at least two of the following criteria: increased probability of encountering loss cards, reduced gain, or increased loss) relative to favorable trials (i.e., trials with at least two of the following criteria: reduced probability of encountering loss cards, increased gain, or reduced loss). Importantly, the two scoring methods from the CCT (total cards flipped and the optimal–suboptimal difference) were not significantly correlated, despite both sharing a significant relationship with the IGT. This reinforces the notion that the novel optimal–suboptimal difference score captures a distinct decision-making construct (i.e., ADM) compared to the original CCT scoring method.

The relationship detected between optimal–suboptimal difference and adjusted advantage needs to be interpreted tentatively given the weak effect size. Two methodological characteristics of the CCT could have constrained the effect of observed relationships. Firstly, similar to the BART, randomness is embedded into the design of the CCT. It is possible individuals recognized the need to maximize reward during favorable trials; however, they may have encountered a loss card prematurely by chance, thereby limiting their total cards flipped during optimal trials relative to suboptimal trials. Secondly, a shortened hot version of the CCT (24-trials instead of 48; Penolazzi et al., 2012) was used in the current study. The additional trials included in the longer version may have more adequately diluted the effect of randomness across the task and the sample. Therefore, the shortened, stochastic design of the CCT could have minimized the effect of the association between optimal–suboptimal difference and IGT-adjusted advantage.

Unfortunately, the BART post-loss pump scoring method did not significantly correlate with any score from the CCT or IGT. The third EFA did produce a one-factor solution, although the BART post-loss pump only weakly loaded onto this factor. Furthermore, the identified model only accounted for 38.914% of the variance, most of which is attributed to the CCT and IGT, which was substantially less than the two-factor solution identified by the second EFA (78.513%) where the BART loaded onto an entirely separate factor. Ultimately, the current study found no evidence that the BART can be used to assess ADM.

Repeatedly, the BART demonstrated negligible relationships with the other decision-making tasks and only presented strong relationships with itself (e.g., adjusted average pumps and post-loss pumps). Scores derived from the BART are likely to represent decision-making under uncertainty—situations where the probability of outcomes is entirely unknown (Steiner & Frey, 2021). In contrast, the CCT and IGT can be considered decision-making under risk—situations where the probability of outcomes is known. This distinction may be further conceptualized through the lens of the “description vs. experience gap” in risk-taking behavior (Hertwig & Wulff, 2022). In the BART, the uncertainty of outcomes precludes the ability to predict or employ deliberative decision-making strategies, leading to decisions largely based on immediate experiences during the task (Johnson & Busemeyer, 2010). Conversely, the CCT and IGT present scenarios where risks are either initially described or become apparent after initial trials. This allows participants to anticipate negative outcomes and apply more strategic, deliberative decision-making processes, consistent with decisions made on the basis of known risk descriptions (Johnson & Busemeyer, 2010).

Research indicates that when decisions are based on experience, rather than a defined level of risk, individuals tend to make riskier choices (Hertwig & Wulff, 2022). This suggests that performance on the BART, a task emblematic of decision-making under uncertainty, is likely driven by factors distinct from those influencing the IGT or CCT. For instance, while risk aversion might impact the choices made in the IGT and CCT, where risks are known or become discernible, it may not play as significant a role in the BART, where the associated risk with each choice remains obscure (De Groot, 2020). Supporting this theory, Di Plinio et al. (2022) found that the original BART was not effective in predicting participants' risk-taking profiles, likely due to its reliance on immediate experiences rather than deliberate decision-making strategies. However, decisions on the BART are likely influenced by participants' sensitivity to punishment (i.e., the degree to which an individual is motivated to avoid negative outcomes). Hevey et al. (2017) demonstrated that clinically depressed individuals significantly reduced their number of pumps on the BART after experiencing a balloon loss compared to healthy controls. The authors posited that this reduced risk-taking behavior was attributable to heightened punishment sensitivity, a characteristic feature of major depressive disorder, thus offering a potential explanation for the unique performance patterns observed on the BART (Hevey et al., 2017).

The distinction between decision-making under risk (i.e., CCT and IGT) and decision-making under uncertainty (i.e., BART) does not only align with the “description vs. experience” framework in risk-taking behavior but is also neurologically grounded. The orbitofrontal cortex (OFC) plays a pivotal role in stimuli appraisal, while the ventromedial prefrontal cortex (VMPFC) is key in regulating decision-making (Morawetz et al., 2019; Rolls, 2019; Sokol-Hessner et al., 2013). Lesion studies have shown that individuals with impairments in these regions exhibit increased risk-taking on the CCT (Spaniol et al., 2018) and disadvantageous selections on the IGT (Fellows, 2007; Ouerchefani et al., 2017), tasks that involve deliberative decision-making strategies. Correspondingly, OFC activity has been linked to reward processing and transitions from high-risk to low-risk choices in the IGT (Li et al., 2010; Zha et al., 2022). Conversely, Schonberg et al. (2012) reported a decrease in VMPFC activity with successive pumps in the BART, a task where risks are unknowable, and decisions are primarily guided by experience. Further research has shown that the greatest level of cortical activity in this region follows a balloon pop (Wang et al., 2022). These differential patterns of activation between the OFC and VMPFC may reflect the underlying cognitive processes specific to decision-making under risk or uncertainty. In tasks like the CCT and IGT, there is increased neural activity as potential risks, which are known or describable, are deliberated. In contrast, the BART, governed by uncertainty, shows reduced activity during the decision-making process when risks are unknown but heightened activity during the processing of experienced outcomes. Regardless, the inherent stochastic design of the BART limits its utility in certain respects (De Groot, 2020). A potential solution, such as a scripted BART that standardizes the experience of balloon pops among participants (Di Plinio et al., 2022; Steiner & Frey, 2021), could negate the task's randomness. This would allow for a learning component regarding the risks associated with each choice, shifting from decisions from experience to decisions from description, thereby potentially enabling the post-loss pump score to more accurately reflect ADM.

6 | LIMITATIONS

The findings of the current study need to be interpreted in light of its limitations. First, the decision-making tasks were administered using the online Inquisit platform, which participants completed in their own time. Subsequently, the length of breaks between tasks or environmental distractors was not controlled as the study was not conducted in a standardized laboratory. Second, the CCT cool version was not administered, and shortened versions of the BART and CCT hot versions were administered to minimize participant fatigue from consecutive decision-making tasks. The cool version of the CCT is theorized to encourage the use of deliberative decision-making processes (Figner et al., 2009) and could have provided further insight into the utility of the optimal–suboptimal scoring method. The reduced maximum number of pumps on the BART, and the reduced number of trials on the CCT could have constrained variability on these tasks impacting their relationships with other variables. Lastly, hypothetical money was used as “payment” across all three decision-making tasks. While studies have shown strong emotional responses are elicited from participants when gambling with hypothetical money, using real money has been shown to change participants' decision-making across gambling tasks (Locey et al., 2011; Xu et al., 2018).

7 | FUTURE DIRECTION

Given the findings from this study, several avenues emerge for future research. It is clear that the inherent randomness of the BART might not be addressed by modifying scoring techniques, and this randomness may undermine the task's psychometric properties. Future research could explore the use of a scripted BART, providing all participants with a standardized experience. Such a modification could enhance the reliability and validity of the BART. The relationship observed in the current study between the CCT and the IGT suggests a potential overlap in the constructs they measure. While several neuroimaging studies have been conducted with the IGT, the CCT remains under investigated. Future research is needed to explore the potential overlap between the two tasks by determining whether the CCT and IGT exhibit shared neural pathways or mechanisms. While the primary focus of this research was to devise more robust scoring methods for commonly used decision-making tasks to assess ADM, further research is needed to examine the ecological validity of these scoring methods. Specifically, research to identify whether these alternative scoring methods for the BART and CCT demonstrate predictive validity concerning real-world decision-making behaviors, such as problem gambling, substance misuse, and other consequential health or financial behaviors. Importantly, we suggest future research to adopt alternative statistical methods for analyzing CCT performance, such as those used by Panno et al. (2013). These methods could provide deeper insights into the nuanced interplay of CCT criteria with cognitive and emotional factors, thereby enriching our understanding of ADM across different contexts.

8 | CONCLUSION

The IGT is the gold standard measure of ADM, providing insight into participants' capacity to reappraise salient stimuli and make prudent long-term decisions. It has been suggested that the BART and CCT are distinct, yet potentially related decision-making tasks. Indeed, the current study showed that the BART is uniquely a measure of static decision-making under uncertainty that does not capture ADM. The randomness inbuilt into the task may be contributing to inconsistencies reported across the literature. The CCT and IGT can be categorized as measures of decision-making under risk, with the CCT uniquely incorporating a dynamic component due to its fluctuating trial criteria. Subsequently, performance on both tasks is likely susceptible to participants' risk aversion, with participants who have a low tolerance for risk likely opting for the advantageous decks on the IGT and flipping fewer cards on the CCT. The current study provided evidence using an alternative scoring method (i.e., optimal–suboptimal score) that the CCT can measure ADM similar to the IGT. Thus, there may be redundancy in administering both the CCT and IGT.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest in publishing this work.

ACKNOWLEDGEMENTS

Open access publishing facilitated by Victoria University, as part of the Wiley - Victoria University agreement via the Council of Australian University Librarians.

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed during the current study are not publicly available due to ethical constraints but are available from the corresponding author on reasonable request and permission from the appropriate human research ethics committee.

CONSENT TO PARTICIPATE

All participants provided consent prior to participation, and the study was conducted with adherence to all relevant ethical standards.

ORCID

Stjepan Sambol https://orcid.org/0000-0002-1969-8192

REFERENCES

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Betz, L. T., Brambilla, P., Ilankovic, A., Premkumar, P., Kim, M. S., Raffard, S., Bayard, S., Hori, H., Lee, K. U., Lee, S. J., Koutsouleris, N., & Kambeitz, J. (2019). Deciphering reward-based decision-making in schizophrenia: A meta-analysis and behavioral modelling of the Iowa gambling task. *Schizophrenia Research*, *204*, 7–15. https://doi.org/10.1016/j.schres.2018.09.009

Brevers, D., Bechara, A., Cleeremans, A., & Noël, X. (2013). Iowa gambling task (IGT): Twenty years after—Gambling disorder and IGT. *Frontiers in Psychology*, *4*, 1–12, 665. https://doi.org/10.3389/fpsyg.2013.00665

Buelow, M. T. (2015). Predicting performance on the Columbia Card Task: Effects of personality characteristics, mood, and executive functions. *Assessment*, *22*(2), 178–187. https://doi.org/10.1177/1073191114539383

Buelow, M. T., & Barnhart, W. R. (2018). Test-retest reliability of common behavioral decision making tasks. *Archives of Clinical Neuropsychology*, *33*(1), 125–129. https://doi.org/10.1093/arclin/acx038

Buelow, M. T., & Blaine, A. L. (2015). The assessment of risky decision making: A factor analysis of performance on the Iowa Gambling Task, Balloon Analogue Risk Task, and Columbia Card Task. *Psychological Assessment*, *27*(3), 777–785. https://doi.org/10.1037/a0038622

Buelow, M. T., & Suhr, J. A. (2009). Construct validity of the Iowa Gambling Task. *Neuropsychology Review*, *19*(1), 102–114. https://doi.org/10.1007/s11065-009-9083-4

Canning, J. R., Schallert, M. R., & Larimer, M. E. (2022). A systematic review of the Balloon Analogue Risk Task (BART) in alcohol research. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, *57*(1), 85–103. https://doi.org/10.1093/alcalc/agab004

De Groot, K. (2020). Burst beliefs—Methodological problems in the Balloon Analogue Risk Task and implications for its use. *Journal of Trial & Error*, *1*(1), 43–51. https://doi.org/10.36850/mr1

Di Plinio, S., Pettorruso, M., & Ebisch, S. J. H. (2022). Appropriately tuning stochastic-psychometric properties of the Balloon Analog Risk Task. *Frontiers in Psychology*, *13*, 881179. https://doi.org/10.3389/fpsyg.2022.881179

Dixon, M. L., Thiruchselvam, R., Todd, R., & Christoff, K. (2017). Emotion and the prefrontal cortex: An integrative review. *Psychological Bulletin*, *143*(10), 1033–1081. https://doi.org/10.1037/bul0000096

Élteto, N., Janacsek, K., Kóbor, A., Takács, Á., Tóth-Fáber, E., & Németh, D. (2019). Do adolescents take more risks? Not when facing a novel uncertain situation. *Cognitive Development*, *50*, 105–117. https://doi.org/10.1016/j.cogdev.2019.03.002

Fellows, L. K. (2007). The role of orbitofrontal cortex in decision making: A component process account. *Annals of the New York Academy of Sciences*, *1121*, 421–430. https://doi.org/10.1196/annals.1401.023

Field, A. (2018). *Discovering statistics using SPSS* (5th ed.). SAGE Publications.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: Age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 709–730. https://doi.org/10.1037/a0014983

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286–299. https://doi.org/10.1037/1040-3590.7.3.286

Haffke, P., & Hübner, R. (2020). Are choices based on conditional or conjunctive probabilities in a sequential risk-taking task? *Journal of Behavioral Decision Making*, *33*(3), 333–347. https://doi.org/10.1002/bdm.2161

Harrison, G. W., & Rutström, E. (2008). *Risk aversion in the laboratory*. Emerald Group Publishing Limited. https://doi.org/10.1016/S0193-2306(08)00003-3

Hertwig, R., & Wulff, D. U. (2022). A description–experience framework of the psychology of risk. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *17*(3), 631–651. https://doi.org/10.1177/17456916211026896

Hevey, D., Thomas, K., Laureano-Schelten, S., Looney, K., & Booth, R. (2017). Clinical depression and punishment sensitivity on the BART. *Frontiers in Psychology*, *8*, 670. https://doi.org/10.3389/fpsyg.2017.00670

Hopko, D. R., Lejuez, C. W., Daughters, S. B., Aklin, W. M., Osborne, A., Simmons, B. L., & Strong, D. R. (2006). Construct validity of the Balloon Analogue Risk Task (BART); relationship with MDMA use by inner-city drug users in residential treatment. *Journal of Psychopathology and Behavioral Assessment*, *28*(2), 95–101. https://doi.org/10.1007/s10862-006-7487-5

Howard, J. D., & Kahnt, T. (2021). Causal investigations into orbitofrontal control of human decision making. *Current Opinion in Behavioral Sciences*, *38*, 14–19. https://doi.org/10.1016/j.cobeha.2020.06.013

Hunt, M. K., Hopko, D. R., Bare, R., Lejuez, C. W., & Robinson, E. V. (2005). Construct validity of the balloon analog risk task (BART) associations with psychopathy and impulsivity. *Assessment*, *12*(4), 416–428. https://doi.org/10.1177/1073191105278740

Johnson, J. G., & Busemeyer, J. R. (2010). Decision making under risk and uncertainty. *WIREs Cognitive Science*, *1*(5), 736–749. https://doi.org/10.1002/wcs.76

Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J. D. (2010). The neural code of reward anticipation in human orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 6010–6015. https://doi.org/10.1073/pnas.0912838107

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31–36. https://doi.org/10.1007/BF02291575

Kessler, L., Hewig, J., Weichold, K., Silbereisen, R. K., & Miltner, W. H. (2017). Feedback negativity and decision-making behavior in the Balloon Analogue Risk Task (BART) in adolescents is modulated by peer presence. *Psychophysiology*, *54*(2), 260–269. https://doi.org/10.1111/psyp.12783

Kluwe-Schiavon, B., Kexel, A., Manenti, G., Cole, D. M., Baumgartner, M. R., Grassi-Oliveira, R., Tobler, P. N., & Quednow, B. B. (2020). Sensitivity to gains during risky decision-making differentiates chronic cocaine users from stimulant-naïve controls. *Behavioural Brain Research*, *379*, 112386. https://doi.org/10.1016/j.bbr.2019.112386

Kovács, I., Richman, M. J., Janka, Z., Maraz, A., & Andó, B. (2017). Decision making measured by the Iowa Gambling Task in alcohol use disorder and gambling disorder: A systematic review and meta-analysis. *Drug and Alcohol Dependence*, *181*, 152–161. https://doi.org/10.1016/j.drugalcdep.2017.09.023

Kully-Martens, K., Treit, S., Pei, J., & Rasmussen, C. (2013). Affective decision-making on the Iowa Gambling Task in children and adolescents with fetal alcohol spectrum disorders. *Journal of the International Neuropsychological Society*, *19*(2), 137–144. https://doi.org/10.1017/S1355617712001026

Lauriola, M., Panno, A., Levin, I. P., & Lejuez, C. W. (2014). Individual differences in risky decision making: A meta-analysis of sensation seeking and impulsivity with the balloon analogue risk task. *Journal of Behavioral Decision Making*, *27*(1), 20–36. https://doi.org/10.1002/bdm.1784

Lejuez, C. W., Aklin, W. M., Jones, H. A., Richards, J. B., Strong, D. R., Kahler, C. W., & Read, J. P. (2003). Balloon Analogue Risk Task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology*, *11*(1), 33–36. https://doi.org/10.1037/1064-1297.11.1.26

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

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84. https://doi.org/10.1037/1076-898X.8.2.75

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, *66*, 799–823. https://doi.org/10.1146/annurev-psych-010213-115043

Li, X., Lu, Z. L., D'Argembeau, A., Ng, M., & Bechara, A. (2010). The Iowa Gambling Task in fMRI images. *Human Brain Mapping*, *31*(3), 410–423. https://doi.org/10.1002/hbm.20875

Locey, M. L., Jones, B. A., & Rachlin, H. (2011). Real and hypothetical rewards in self-control and social discounting. *Judgment and Decision Making*, *6*(6), 522–564. https://doi.org/10.1017/S1930297500002515

Løvstad, M., Funderud, I., Endestad, T., Due-Tønnessen, P., Meling, T. R., Lindgren, M., Knight, R. T., & Solbakk, A. K. (2012). Executive functions after orbital or lateral prefrontal lesions: Neuropsychological profiles and self-reported executive functions in everyday living. *Brain Injury*, *26*(13–14), 1586–1598. https://doi.org/10.3109/02699052.2012.698787

Meshi, D., Ulusoy, E., Özdem-Mertens, C., Grady, S. M., Freestone, D. M., Eden, A., & Ellithorpe, M. E. (2020). Problematic social media use is associated with increased risk-aversion after negative outcomes in the Balloon Analogue Risk Task. *Psychology of Addictive Behaviors*, *34*(4), 549–555. https://doi.org/10.1037/adb0000558

Millisecond Software. (2022). *Inquisit 6 [computer software]*. Millisecond Software. Retrieved from https://www.millisecond.com

Morawetz, C., Mohr, P. N. C., Heekeren, H. R., & Bode, S. (2019). The effect of emotion regulation on risk-taking and decision-related activity in prefrontal cortex. *Social Cognitive and Affective Neuroscience*, *14*(10), 1109–1118. https://doi.org/10.1093/scan/nsz078

Ouerchefani, R., Ouerchefani, N., Allain, P., Ben Rejeb, M. R., & Le Gall, D. (2017). Contribution of different regions of the prefrontal cortex and lesion laterality to deficit of decision-making on the Iowa Gambling Task. *Brain and Cognition*, *111*, 73–85. https://doi.org/10.1016/j.bandc.2016.06.010

Panno, A., Lauriola, M., & Figner, B. (2013). Emotion regulation and risk taking: Predicting risky choice in deliberative decision making. *Cognition & Emotion*, *27*(2), 326–334. https://doi.org/10.1080/02699931.2012.707642

Penolazzi, B., Gremigni, P., & Russo, P. M. (2012). Impulsivity and reward sensitivity differentially influence affective and deliberative risky decision making. *Personality and Individual Differences*, *53*(5), 655–659. https://doi.org/10.1016/j.paid.2012.05.018

Peters, E., Västfjäll, D., Gärling, T., & Slovic, P. (2006). Affect and decision making: A "hot" topic. *Journal of Behavioral Decision Making*, *19*(2), 79–85. https://doi.org/10.1002/bdm.528

Pignatiello, G. A., Martin, R. J., & Hickman, R. L. Jr. (2020). Decision fatigue: A conceptual analysis. *Journal of Health Psychology*, *25*(1), 123–135. https://doi.org/10.1177/1359105318763510

Roesch, M. R., & Olson, C. R. (2007). Neuronal activity related to anticipated reward in frontal cortex: Does it represent value or reflect motivation? *Annals of the New York Academy of Sciences*, *1121*, 431–446. https://doi.org/10.1196/annals.1401.004

Rolls, E. T. (2019). *The orbitofrontal cortex*. Oxford University Press. https://doi.org/10.1093/oso/9780198845997.001.0001

Rotge, J. Y., Poitou, C., Fossati, P., Aron-Wisnewsky, J., & Oppert, J. M. (2017). Decision-making in obesity without eating disorders: A systematic review and meta-analysis of Iowa gambling task performances. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity*, *18*(8), 936–942. https://doi.org/10.1111/obr.12549

Salfi, F., Lauriola, M., Tempesta, D., Calanna, P., Socci, V., De Gennaro, L., & Ferrara, M. (2020). Effects of total and partial sleep deprivation on reflection impulsivity and risk-taking in deliberative decision-making. *Nature and Science of Sleep*, *12*, 309–324. https://doi.org/10.2147/NSS.S250586

Schmitz, F., Manske, K., Preckel, F., & Wilhelm, O. (2016). The multiple faces of risk-taking: Scoring alternatives for the balloon-analogue risk task. *European Journal of Psychological Assessment*, *32*(1), 17–38. https://doi.org/10.1027/1015-5759/a000335

Schnider, A., Treyer, V., & Buck, A. (2005). The human orbitofrontal cortex monitors outcomes even when no reward is at stake. *Neuropsychologia*, *43*(3), 316–323. https://doi.org/10.1016/j.neuropsychologia.2004.07.003

Schonberg, T., Fox, C. R., Mumford, J. A., Congdon, E., Trepel, C., & Poldrack, R. A. (2012). Decreasing ventromedial prefrontal cortex activity during sequential risk-taking: An FMRI investigation of the balloon analog risk task. *Frontiers in Neuroscience*, *6*, 80. https://doi.org/10.3389/fnins.2012.00080

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience*, *8*(3), 341–350. https://doi.org/10.1093/scan/nss002

Spaniol, J., Di Muro, F., & Ciaramelli, E. (2019). Differential impact of ventromedial prefrontal cortex damage on "hot" and "cold" decisions under risk. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(3), 477–489. https://doi.org/10.3758/s13415-018-00680-1

Steiner, M. D., & Frey, R. (2021). Representative design in psychological assessment: A case study using the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: General*, *150*(10), 2117–2136. https://doi.org/10.1037/xge0001036

Timmer, A., Antonaccio, O., & French, M. T. (2021). Hot or cool processing? Adolescent decision-making and delinquency. *Justice Quarterly*, *38*(6), 961–994. https://doi.org/10.1080/07418825.2020.1730424

Tuvblad, C., Gao, Y., Wang, P., Raine, A., Botwick, T., & Baker, L. A. (2013). The genetic and environmental etiology of decision-making: A longitudinal twin study. *Journal of Adolescence*, *36*, 245–255. https://doi.org/10.1016/j.adolescence.2012.10.006

Umbach, R., Leonard, N. R., Luciana, M., Ling, S., & Laitner, C. (2019). The Iowa Gambling Task in violent and nonviolent incarcerated male adolescents. *Criminal Justice and Behavior*, *46*(11), 1611–1629. https://doi.org/10.1177/0093854819847707

Wang, M., Zhang, S., Suo, T., Mao, T., Wang, F., Deng, Y., Eickhoff, S., Pan, Y., Jiang, C., & Rao, H. (2022). Risk-taking in the human brain: An activation likelihood estimation meta-analysis of the balloon analog risk task (BART). *Human Brain Mapping*, *43*(18), 5643–5657. https://doi.org/10.1002/hbm.26041

Wang, X. J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, *60*(2), 215–234. https://doi.org/10.1016/j.neuron.2008.09.034

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, *60*, 53–85. https://doi.org/10.1146/annurev.psych.60.110707.163633

White, T. L., Lejuez, C. W., & de Wit, H. (2008). Test-retest characteristics of the Balloon Analogue Risk Task (BART). *Experimental and Clinical Psychopharmacology*, *16*(6), 565–570. https://doi.org/10.1037/a0014083

Xu, S., Korczykowski, M., Zhu, S., & Rao, H. (2013). Risk-taking and impulsive behaviors: A comparative assessment of three tasks. *Social Behavior and Personality: an International Journal*, *41*(3), 477–486. https://doi.org/10.2224/sbp.2013.41.3.477

Xu, S., Pan, Y., Fang, Z., Wang, F., Rao, H., Qu, Z., Yang, F., & Yang, Z. (2018). Differential effects of real versus hypothetical monetary reward magnitude on risk-taking behavior and brain activity. *Scientific Reports*, *8*(1), 3712. https://doi.org/10.1038/s41598-018-21820-0

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zha, R., Li, P., Liu, Y., Alarefi, A., Zhang, X., & Li, J. (2022). The orbitofrontal cortex represents advantageous choice in the Iowa gambling task. *Human Brain Mapping*, *43*(12), 3840–3856. https://doi.org/10.1002/hbm.25887

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sambol, S., Suleyman, E., & Ball, M. (2024). The assessment of affective decision-making: Exploring alternative scoring methods for the Balloon Analog Risk Task and Columbia Card Task. *Journal of Behavioral Decision Making*, *37*(1), e2367. https://doi.org/10.1002/bdm.2367