Hu, Tianqi; Leppänen, Ilkka; Franco, L. Alberto

Intuition in Decision Making : Insights From Drift Diffusion Modeling

Published in:
Journal of Behavioral Decision Making

DOI:
10.1002/bdm.70033

E-pub ahead of print: 18/07/2025

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

Journal of Behavioral Decision Making

WILEY

RESEARCH ARTICLE OPEN ACCESS

Intuition in Decision Making: Insights From Drift Diffusion Modeling

Tianqi Hu[1] | Ilkka Leppänen[1] | L. Alberto Franco[2,3]

[1]School of Business, Aalto University, Espoo, Finland | [2]University of Bristol Business School, Bristol, UK | [3]Universidad del Pacífico, Lima, Peru

Correspondence: Tianqi Hu (tianqi.hu@aalto.fi; tianqimark@gmail.com)

Received: 7 July 2024 | **Revised:** 9 June 2025 | **Accepted:** 7 July 2025

Funding: This work was supported by Loughborough Business School at Loughborough University.

Keywords: cognitive reflection test | domain-specific preference | drift diffusion model | intuition

ABSTRACT

Research on intuition often produces conflicting results and suffers from reliability issues due to the lack of tools that can conclusively evaluate a person's latent intuitive position. It has been recently proposed that a decision maker's intuitive position can be evaluated by estimating parameters with sequential sampling models (SSMs), which provide a biologically plausible framework to measure how intuition affects decisions. In two studies, we use the drift diffusion model (DDM), as a type of SSM, to investigate topics where intuition is difficult to evaluate. In Study 1, we used the DDM to examine how the cognitive reflection test (CRT) scores relate to intuition in risky decision making and found that individuals with high CRT scores had superior performance and relied more on intuition. These findings challenge the conventional view that high CRT scores imply less reliance on intuition and that intuition is detrimental to decision performance. In Study 2, we examined the cross-domain stability of the preference for intuition and found that decision makers rely more on intuition in the social decision domain than in the risky decision domain and that these measures are not correlated across the two domains. The evidence for this unstable preference has hitherto primarily resulted from self-reports, which have a questionable ability to assess the preference for intuition. In both studies, we demonstrate that the DDM can accurately simulate the decision outcome and decision time patterns that are affected by intuition, providing evidence for the usefulness of DDM analysis in the study of intuition.

1 | Introduction

Intuition (or intuitive thinking, intuitive judgment) in decision making is broadly defined as judgments that arise from rapid, holistic thinking with little apparent effort and conscious awareness (Dane and Pratt 2007; Hogarth 2010). The concept of intuition has a pivotal role in descriptive theories of thinking and reasoning. For example, it is popularly theorized that decisions are driven by the interaction between an intuitive thinking process and a deliberative thinking process that is reflective, effortful, and analytical (Epstein et al. 1996; Kahneman 2011). The prominence of intuition in decision theory has inspired a wide range of literature to investigate how the utilization or suppression of intuition affects various aspects of decision making, such as decision performance (e.g., Fellnhofer and Deng 2024; Phillips et al. 2016), propensity to cognitive biases (e.g., Boissin et al. 2021), and creativity (e.g., Baldacchino et al. 2023).

While there has been extensive literature on the role of intuition in decision making, it is recognized that research on intuition often meets methodological problems. Indeed, intuition is a latent cognitive process that cannot be directly observed, making its evaluation difficult. For instance, one widely used strategy for studying intuition involves examining behaviors under conditions manipulated to promote and suppress intuitive thinking. Nevertheless, a recent review shows that such manipulations often fail to affect the use of intuition, and different manipulation methods can yield inconsistent findings (Isler and Yilmaz 2023). Another popular approach for investigation is to ask respondents to self-report how they tend to use intuition in various situations. However, it is unclear how accurately decision makers can assess their unconscious use of intuition (Isler and Yilmaz 2023; Nisbett and Wilson 1977). Overall, it is acknowledged that tools that can conclusively evaluate a person's intuitive position are still lacking, preventing clear conclusions about how intuition affects behaviors and decisions (Diederich and Trueblood 2018; Fellnhofer 2022; Lufityanto et al. 2016).

One ongoing debate arising from the difficulty of evaluating intuition centers on the cognitive reflection test (CRT; Frederick 2005). It has been contested whether individuals scoring high on the CRT have superior decision performance because they tend to inhibit intuition in decision making. The CRT consists of questions designed to have incorrect but intuitive answers, and deducing the correct answer is thought to necessitate inhibition of error-prone intuition (Frederick 2005). Since it was first developed, the CRT has gained popularity due to its ease of administration and potency in predicting reasoning and job performances across a range of contexts (Brañas-Garza et al. 2019; Otero et al. 2021; Toplak et al. 2011). Importantly, given the construct conceptualization of the CRT, many have adopted the interpretation that individuals measured to have high judgment ability on the CRT (hereafter referred to as "high-ability individuals") achieve superior performances because they can better suppress intuition that tends to produce erroneous judgment (e.g., Moritz et al. 2013, 2014; Thoma et al. 2015). However, the validity of this interpretation has been contested. For example, it has been shown that CRT measurements tend to strongly correlate with measurements of numeracy, and some studies even report that CRT measurements provide no independent predictive power for decision performances from numeracy[1] (Patel et al. 2019; Sinayev and Peters 2015; Welsh et al. 2013). These findings give rise to the proposition that intuition is not relevant to the CRT's predictive power. Nonetheless, it should be noted that intuition is thought to play an important role in numeracy as well (Peters 2012). Both competent mathematical reasoning and common errors in mathematical problem solving have been attributed to intuitive thinking (Cokely et al. 2012; Jiang et al. 2020; Van Dooren et al. 2003). Therefore, while numeracy likely contributes to the CRT's predictive power, it is mistaken to assume that intuitive thinking is irrelevant. Another school of thought has proposed an opposite explanation about the CRT's predictive power by showing that CRT measurements can predict decision performance under conditions conducive to the use of intuition (De Neys and Pennycook 2019; Erceg et al. 2022; Šrol and De Neys 2021). It is thus argued that superior decision-making performance owes much to the use of, rather than the inhibition of, intuitive thinking and that consequently, the CRT measures a decision maker's intuitive ability. Overall, as the operation of intuition is unobservable, investigation on how intuition relates to the CRT's predictive power has relied on associating CRT measurements with personal traits and behavioral patterns that are only indicative of the involvement of intuition. Consequently, this approach has yielded conflicted findings, preventing a clear conclusion being drawn.

The cross-domain stability of the preference for intuition is another topic where understanding is hindered by the difficulty to measure the involvement of intuition. Conventionally, the preference for relying on intuition, as opposed to deliberation, for thinking and reasoning is typically assumed to be a chronic disposition (Betsch 2008). In other words, the extent to which decision makers prefer to use intuition in one decision domain, such as financial investment, should be consistent with that in another decision domain, such as choosing a romantic partner. Consistent with this assumption, studies on decision makers' preferences for intuition in specific domains commonly adopt psychometric questionnaires designed to assess the domain-general preference for intuition (Pachur and Spaar 2015). For an illustrative example, an item from such psychometric questionnaires states: "How true does the following statement describe you: Using my gut feelings usually works well for me in figuring out problems in my life." (Epstein et al. 1996). It should be noted that the use of these psychometric measures relies on the crucial yet unverified assumption that a decision maker's preference for intuition is stable across decision domains. Only when this assumption holds, measures of the domain-general preference for intuition can reliably indicate a decision maker's domain-specific preference in the situations under investigation. To date, studies that explore the cross-domain stability of the preference for intuition have heavily relied on self-report questionnaires for measuring the preference in specific decision domains. It has been documented that decision makers report varying degrees of reliance on intuition across decision domains due to characteristics of decision problems (Berman et al. 2018; Gallo et al. 2017; Inbar et al. 2010; Pachur and Spaar 2015; Oktar and Lombrozo 2022). However, there are reasons to assert that the subjective self-report approach used in the literature does not accurately capture the preference for intuition. It has been shown that individuals are often not consciously aware of their own judgment-forming processes, leading to biased self-assessment (Scopelliti et al. 2015). Considering that intuition involves little conscious awareness, it is unclear to what extent the respondents are sentient about their preferences for intuition, especially when they are asked to provide their preferences in hypothetical situations where no direct, personal experience is involved. In fact, Leach and Weick (2018) show that a person's belief and confidence in her intuitive ability are not reflective of her implicit learning performance where intuitive thinking is required, suggesting that people cannot accurately judge their own intuition. As such, it is questionable whether the self-report approach can truthfully capture preferences for intuition, hence leaving the question of the stability of preference for intuition unsettled.

The difficulty in measuring intuition has inspired continuing research in devising tools that can credibly assess the role of intuition in decision making. Recently, developments in computational decision science have provided new opportunities to understand how intuition affects decisions. A growing body of literature provides evidence that the cognitive process in decision making can be represented by a drift diffusion model (DDM) process (Bogacz et al. 2006; Forstmann et al. 2016; Ratcliff et al. 2016). The DDM, as a type of sequential sampling model, represents decision making as a process of evidence accumulation, i.e., as a process by which noisy information regarding the desirability of available options is sequentially sampled until sufficient evidence has accrued to favor one option over others. The DDM is composed of a set of model components that index various features of the cognitive process in decision making. Through the estimation of these model components, the DDM decomposes and quantifies the effects of latent factors that affect decisions (Clithero 2018). To note, the DDM can be formulated to have a biased *starting point* of the evidence accumulation process, a component that indexes a decision maker's predisposition in favor of one type of decision prior to evidence accumulation (Desai and Krajbich 2022). The DDM starting point has been interpreted in various ways in empirical studies. For instance, Frydman and Nave (2017) used the DDM starting point to represent the extrapolative beliefs of decision makers, and Enax et al. (2016) used the DDM starting point to infer the effect of stimulus–response association. Importantly, one body of literature has theorized that the DDM starting point represents the decision maker's intuitive position in decision making. Specifically, the concept of a predisposition modeled by the starting point in evidence accumulation is analogous to the concept of intuitive thinking that is often theorized to affect decisions by generating default responses prior to effortful reasoning (Evans and Stanovich 2013). The brain imaging study of Mulder et al. (2012) provides support for this conceptual overlap, which shows that the starting point in the DDM is related to neural activity underlying the formation of biased decisions due to prior knowledge regarding the desirability of available options. Therefore, the DDM is seen to provide a biologically plausible framework for how intuition affects decision making, and the starting point in the model has been used to represent and measure a person's intuitive position in various contexts, providing insights that are otherwise difficult to be obtained (Chen and Krajbich 2018; Montero-Porras et al. 2022; Zhao et al. 2019, 2022).

In the research reported here, we used the DDM to probe into the latent intuition in decision making in two studies. In Study 1, we investigated the effect that intuition has in forming the relationship between the CRT and expected value maximization in risky decisions. The results help to understand how intuition contributes to the CRT's predictive power for superior decision performance. In Study 2, we examined the association between the preferences for intuition in risky decision and social decision domains. The results provide evidence for the stability of the preference for intuition in decision making. For both studies, we show that the DDM can simulate observed patterns in decision making with high accuracy, including both decision outcome and decision time patterns. This result provides evidence for the validity of the DDM as a tool for evaluating intuition, which is thought to simultaneously affect decision outcomes and decision times by generating fast, default responses (Evans and Stanovich 2013). The ability of the DDM to accurately simulate key behavioral patterns, a feature not shared by other methods for measuring intuition, highlights the value of using the DDM to inform unresolved issues about the role of intuition in decision making.

2 | Study 1

To understand the role of intuition in forming the relationship between the CRT and superior decision performance, we study how CRT measurements can predict expected value maximization in risky decision making. The association between the CRT and risky decision making has attracted strong scholarly interest (for a review, see Lilleholt 2019). In behavioral studies, high-ability individuals are often observed to evaluate risky options close to their expected values (or expected returns), which are regarded as normatively optimal judgments (e.g., Ackert et al. 2020; Cokely and Kelley 2009; Frederick 2005; Oechssler et al. 2009; Sajid and Li 2019). A person's tendency to follow expected value maximization in experimental settings is found to associate with superior risk assessment in real-life domains, such as financial and health decision making (Anderson and Mellor 2008; Boyle et al. 2012). Broadly speaking, risk assessment is pervasive in life. From trivial decisions such as whether to try a new drink, to significant decisions requiring large investments, almost all life domains involve risk assessment of various levels (Figner and Weber 2011). Given the ubiquity of risk assessment, the insights learned about why the CRT can predict superior risky decisions can likely inform the CRT's predictive power for superior performances in a wide range of contexts.

We carried out an experimental study that administers the CRT to each participant. The decision-making task in the experiment includes a series of incentivized risky gambles, based on which the decision maker's degree of reliance on intuition and the effect of intuition on risky preference formation are estimated using the DDM. The following two research questions are proposed to guide the analysis:

> Research Question 1A (RQ1A): How does the CRT relate to the degree of reliance on intuition in risky decision making?
>
> Research Question 1B (RQ1B): Is the effect of intuition the mechanism underlying the CRT's predictive power for expected value maximizing responses?

If the CRT's predictive power can be explained by its original conceptualization, we can expect to observe CRT measurements to inversely correlate with the degree of reliance on intuition, and the predictive power of CRT measurements for expected value maximization to be mediated by the effect of intuition. Otherwise, alternative explanations need to be explored.

2.1 | Materials and Methods

2.1.1 | Experimental Procedure

The subjects in the experiment were recruited from the participant pool of the Prolific.co platform, and we conducted the experiment online. The interface of this web browser–based experiment was programmed using oTree (Chen et al. 2016). A total of 122 non-student subjects from the United Kingdom participated in the experiment (Male = 61, Female = 61; Mean Age = 31.7 years, Age range = 20–48 years). Each subject was paid £3 for participation, and they earned an additional amount from completing monetarily incentivized risky decisions. On average, each subject was paid £8.38 in total. All experimental procedures were approved by the institutional review board of the university with which the author is affiliated.

The experiment was composed of two sections. In the first section, each subject completed the multiple-choice format of the seven-item version of the CRT validated by Sirota and

Juanchich (2018). The administered CRT can be seen in the Supplement (Section S1.1). The second section was composed of a series of risky gambles. At the beginning, each subject was given 100 experiment points, which is a fictitious currency that has a 20 to 1 exchange rate with the Pound Sterling, that is 20 experiment points = £1. Each subject then completed a sequence of 200 decision trials in the form of risky gambles. For each trial, the subject needs to decide whether to accept or reject a gamble that has a 50% chance to gain and a 50% chance to lose experiment points. Specifically, the amounts to be gained or lost are taken from a set of {10, 20, 30, 40, 50, 60, 70, 80, 90, 100} experiment points. With this set of values, 100 unique gambles are generated. Each subject played each gamble twice, switching the sides of the display of the gain and loss values on the interface, resulting in 200 gambles. Moreover, the order of the 200 gambles was randomized to encourage the subjects to consider each gamble as an independent decision problem. An illustrative example of a decision trial is shown in Figure 1. This paradigm of risky decision making was adapted from Zhao et al. (2020). Subjects were informed that at the end of the experiment, one trial from the risky gamble decisions would be randomly selected to count. The subject would receive the outcome of the gamble if she chose to accept the gamble. Alternatively, the subject would finish the section with the endowed 100 experiment points if she chose to reject gamble. By this design, the subject would not lose more than the endowed experiment points. The experiment points are converted to Pound Sterling when the subjects receive the payment.

To make sure the instruction about the decision trials was clearly understood by the subjects, eight check-up questions were administered, such that additional instructions would be provided if incorrect answers are submitted. A subject could only proceed after answering all check-up questions correctly. Ten practice decision trials were also administered for the subjects to become familiar with the procedure. Furthermore, subjects were instructed to concentrate on the decision tasks without interruptions, until reaching the resting points, which are set after trials 50, 100 and 150. Detailed information about the instruction, check-up questions and practice trials are presented in the Supplement (Section S1.2–S1.4).

2.1.2 | The CRT Scores

There are different scoring methods pertaining to the CRT. The regular approach of coding the CRT score is to tally the correct answers; this produces the CRT Reflectiveness Score (CRT-RS). The CRT-RS is thought to measure the test taker's tendency in cognitive reflection, that is the tendency to inhibit intuition and engage in reflective reasoning. However, this scoring method has been criticized for measuring not only cognitive reflection but also numeracy because obtaining correct answers on the CRT requires numerical skills as well, making it difficult to interpret the CRT-RS. Therefore, researchers have adopted alternative scoring methods that are thought to better capture cognitive reflection (Erceg and Bubić 2017).

It has been proposed that the CRT Intuitiveness Score (CRT-IS), coded by tallying incorrect intuitive answers, is a better measurement of the non-inhibition of intuitive thinking (Dorigoni et al. 2022; Sinayev and Peters 2015). It has been theorized that obtaining correct answers on the CRT results from a two-step process. In the first step, the decision maker suppresses the intuitive response and notices that the intuitive answer is incorrect; then, in the second step, she engages in problem solving, the success of which depends on her numeracy. By this view, an intuitive answer represents the decision maker's failure in the first step, that is in the inhibition of intuitive thinking (Dorigoni et al. 2022; Sinayev and Peters 2015). As such, the

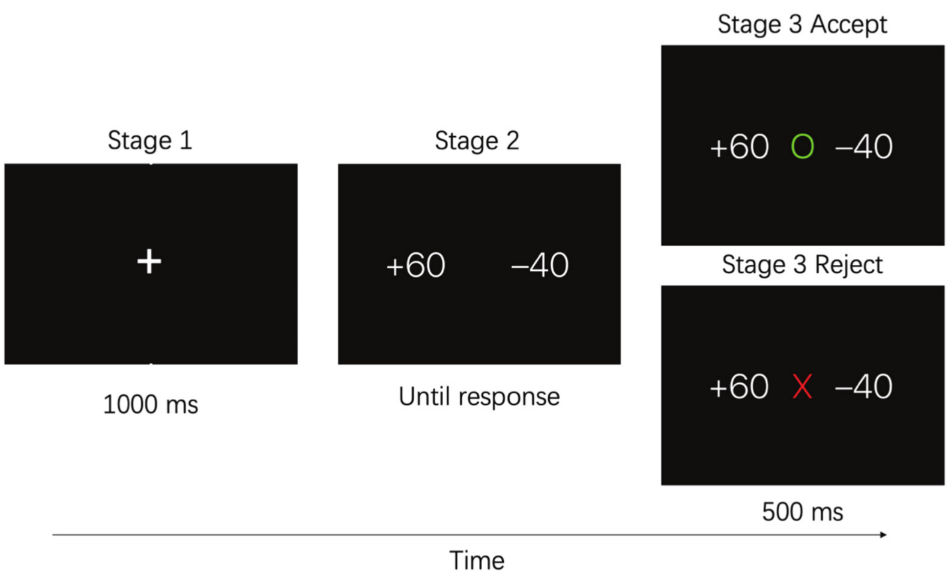

FIGURE 1 | An example of a decision trial. At Stage 1, a fixation cross is shown in the middle of the screen for the duration of 1000 ms. Thereafter, the cross disappears and the trial proceeds to Stage 2, where the gain and loss values are displayed on two sides of the screen. The value with a "+" sign indicates the potential gain, and the value with a "−" sign indicates the potential loss. Subjects can accept or reject the gamble by pressing the up or down keys on the keyboard. There is no time limit in Stage 2. After a decision is made, the trial proceeds to Stage 3, which shows either a green circle in the middle of the screen if the subject chooses to accept the gamble, or a red cross if the subject chooses to reject the gamble. Stage 3 lasts for 500 ms before the next trial starts.

CRT-IS is thought to be a specific measurement of a decision maker's tendency to rely on intuition, free from the influence of numeracy. In practice, however, studies often report strong negative correlations between the CRT-RS and CRT-IS, indicating that they largely capture the variance of the same trait (Erceg and Bubić 2017). Despite this, given the theoretical distinction between the two CRT measurements, we inspected both the CRT-RS and CRT-IS with respect to their associations with intuitive thinking. For ease of interpretation, we coded the CRT-IS as 7 (the number of questions in the CRT) minus the number of incorrect intuitive answers. As such, incremental values of the CRT-IS represent higher judgment ability at inhibiting error-prone intuition by the original conceptualization of the CRT.

2.1.3 | Expected Value Maximizing Response and Decision Domains

Our experiment involves 200 risky gambles that have even chances to gain or lose certain monetary values (expressed in experiment points). The expected value maximizing response is to accept the gamble when the gain value is higher and to reject the gamble when the loss value is higher. The risky decision trials were assigned into two domains according to the relative gains and losses. Those with higher gains belong to the *advantageous domain*, and those with higher losses belong to the *disadvantageous domain*. If the CRT is associated with expected value maximization, we can expect to observe high-ability individuals to accept gambles more often in the advantageous domain and reject gambles more often in the disadvantageous domain than their low-ability counterparts.

It has been shown that different sets of cognitive mechanisms are responsible for risky preference formation in gain prospects (losing is not possible) and loss prospects (gaining is not possible; Brooks et al. 2014; Clay et al. 2017). These findings suggest that intuition may affect decision making differently in the advantageous and disadvantageous domains, which are featured by different prospects of returns. Therefore, it is necessary to examine the role of intuition separately in the advantageous and disadvantageous domains. To note, decision trials from the advantageous and disadvantageous domains were mixed in the experiment. Therefore, separate analysis of intuition in the two decision domains implies that decision makers could intuitively process the relative size of gain and loss values at the onset of decision making. This is a reasonable assumption as it has been theorized that intuition is used for recognizing and interpreting cues (De Neys and Pennycook 2019). In other words, it is possible that decision makers' intuitive judgment was triggered by the relative size of gain and loss values.

2.1.4 | Drift Diffusion Modeling

Based on the decision outcomes, decision times and gain and loss values from the risky decision trials, DDMs were fitted to each subject. Figure 2 shows a schematic for the DDM used to model the risky decisions in the experiment. Each trajectory is a hypothetical decision trial that initiates evidence accumulation from a biased starting point and ends when one of the thresholds is reached. The red trajectory represents a decision trial where the gain is higher valued in the subjective utility and vice versa for the blue trajectory. The decision maker has a predisposition towards rejecting the gamble, given by the shifted starting point, such that relatively less (more) evidence needs to be accrued for the decision maker to reject (accept) the gamble. As a result, the blue trajectory quickly reaches the threshold for rejection, and it takes the red trajectory a much longer time to reach the threshold for acceptance. The fluctuations in the trajectories indicate the effect of randomness, or noise, on evidence accumulation. A decision maker could favor one decision according to her subjective utility but due to randomness of the trajectory end up making the opposite decision if the noise is strong.

Importantly, we examine both signed starting point values and absolute starting point values estimated from the DDM. The signed starting point value was estimated as the distance from zero to the shifted point, divided by the distance from zero to a threshold. A starting point shifted towards accepting (rejecting) the gamble has a positive (negative) value. The signed starting point value is interpreted as the effect of intuition for causing a subject to accept or reject risky gambles. In other words, the signed starting point value ranges from −1 to 1. From 0 to 1, acceptance is favored more than rejection by intuition. Conversely, in the range from 0 to −1, rejection is favored more than acceptance. Overall, higher values of the signed starting point represent the acceptance is favored relatively more than rejection by intuition. The absolute starting point value is calculated as the absolute value of a signed starting point. It measures the extent to which a starting point shifts away from the midpoint between the two thresholds, regardless the direction of the shift. The absolute value of the starting point is interpreted as the decision maker's degree of reliance on intuition, irrespective of whether the intuition favors acceptance or rejection. A higher value of absolute starting point indicates the decision maker relied more on intuition for making decision.

The DDM starting point value is simultaneously estimated with other model components, including the drift rate component that indexes the computation in evidence accumulation. The drift rate is commonly interpreted as to represent utility evaluation process dependent on task-specific conditions. Therefore, in the present study, the drift rate was specified as a utility function depending on the gain and loss values in each decision trials. Two forms of the DDM with different

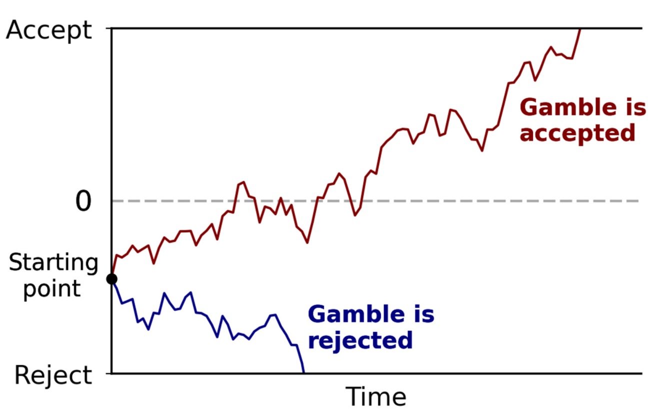

FIGURE 2 | The drift diffusion model. The horizontal and vertical axes represent time and evidence accumulation state, respectively.

specifications of the drift rate function were estimated. In DDM-1, the drift rate function is represented by the *utility weighting effects* associated with gain and loss values, as suggested by the prospect theory (Kahneman and Tversky 1979). In DDM-2, the drift rate function has an additional *fixed utility bias*, representing a valuation process independent from specific gain and loss values offered by the gambles in risky decisions (Zhao et al. 2020). As such, a decision maker's risk preference is mapped onto both the DDM starting point, representing the effect of intuition, and the drift rate, which reflects the utility weighting effect and (in DDM-2) fixed utility bias. In other words, individual differences in subjective utility are accounted for in the estimation of how intuition affects decision making. The computational details are given in the Supplement (Section S2). Both DDM-1 and DDM-2 have been used for modeling risky decisions in similar experimental paradigms (Sheng et al. 2020; Zhao et al. 2020). In the present analysis, we inspected the relationship between the CRT and intuition by using both sets of the starting point values estimated from the DDM-1 and DDM-2. Additionally, as DDM-1 is a nested model of DDM-2, the goodness-of-fit between the two models can be assessed using the likelihood ratio test. We also examine research questions using starting point values estimated from the better-fitting model between DDM-1 and DDM-2 (hereafter referred to as "best-fit model set"; Lee et al. 2019).

One question that needs to be addressed is whether a decision maker's intuitive position in a decision domain is related to their subjective preference in that domain. For example, one might suspect that, compared with a risk-neutral decision maker, someone who is strongly risk-averse or risk-seeking would be more likely to have a stronger intuition favoring a particular type of response. From this perspective, the absolute starting point (i.e., the distance between 0 and a given starting point) cannot reliably measure the degree of reliance on intuition, as it does not account for the extent to which subjective preference may influence the starting point position. However, separating a decision maker's intuitive position from their subjective preference is challenging, as preference is at least partly shaped by intuition. One way to examine the relationship between intuition and subjective preference is to analyze the relationship between the starting point and drift rate components (i.e., utility weighting effect and fixed utility bias). The drift rate components index computation in evidence accumulation and represent a decision maker's preference, assuming intuition is unbiased. If signed starting point values strongly correlate with drift rate components, and do not explain additional variance in choice probability beyond what is accounted for by drift rate components, this would suggest that the starting point position reflects subjective preference arising from evidence accumulation. Alternatively, if the correlation between signed starting point values and drift rate components is weak or only modest, and both independently explain variance in choice probability, this would suggest that subjective preference derived from evidence accumulation is not the primary factor determining the starting point position that represents intuition.

As a robustness test, we used an alternative measure to estimate reliance on intuition. We calculated the absolute deviation between each participant's signed starting point value and the value predicted from a linear regression on their choice proportion (the proportion of acceptance among all decisions). This deviation quantifies how much a participant's starting point position differs from what would be expected based on their observed choices, providing a behavior-relative measure of reliance on intuition. Unlike the absolute starting point value, which may be confounded by individual differences in subjective preference, this approach helps distinguish reliance on intuition from subjective preference by measuring the extent to which the participant's starting point position is more extreme than predicted by their choice behavior alone. If starting point position primarily reflects preference arising from evidence accumulation, then choice proportion should account for most of the variation in starting point, and the deviation should be small. In contrast, larger deviations suggest that reliance on intuition contributes to starting point position independently of preference arising from evidence accumulation.

Our experiment does not include manipulations intended to alter a person's intuitive judgment throughout the period of risky decision making. Hence, the starting point was estimated as constant values throughout the decision trials used for model fitting in both DDM-1 and DDM-2. Previous research has shown that the DDM starting point would move in response to changes in the expected favorability of available options (Desai and Krajbich 2022; Zhao et al. 2020). This finding is in line with the prediction that intuitive thinking is used for recognizing patterns and interpreting cues (De Neys and Pennycook 2019). In our experiment, the incentive structure is not altered throughout the decision trials. It is also theorized that intuitive thinking can develop through learning and experience accumulation (De Neys 2023). In our study, no feedback was provided during the decision trials and the payment from risky decision making was only revealed at the end of the experiment. While there is no apparent reason to assume that the intuitive position of decision makers would change over the course of risky decision making, there has been little evidence about how intuition may develop overtime. As a robustness test, we estimated two additional forms of the DDM, DDM-1-VSP and DDM-2-VSP (Variable Starting Point), where the starting point from DDM-1 and DDM-2 was estimated separately on the first and second half of the decision trials used for model fitting. We examine whether the starting point that represents intuition would alter overtime and how findings related to our research questions may change.

In addition to the starting point and drift rate, the DDM also includes a threshold component that indexes the balance in emphasizing the speed versus accuracy in decision making, and a non-decision time component that indexes the time used for information processing unrelated to the task conditions. The threshold height and non-decision time components were estimated as constant values throughout the decision trials used for model fitting. The estimation of DDM components was performed using PyDDM, a freely available software package in Python (Shinn et al. 2020).

2.1.5 | Statistical Testing

If the CRT can predict expected value maximization in the risky decision trials, high CRT measurements, including the

CRT-RS and CRT-IS, should associate with greater likelihood of accepting gambles in the advantageous domain, and lower likelihood of accepting gambles in the disadvantageous domain. To assess these relationships, each subject's decision trials were pooled together, and generalized linear mixed models (GLMMs; Baayen et al. 2008) were used. The analysis was implemented using the lme4 package in R (Bates et al. 2015). Three control variables were included in the regression model: the value difference between the gain and the loss in a trial (Difference), the mean value of the gain and loss as a proxy of the stake of the risky decision (Stake), and their interaction. The value difference between the gain and the loss is the most theoretically important predictor of the decision type in a trial, and its effect should be moderated by the stake size of the risky decision. For example, a pair of gain and loss with values of 10 and 20 and another pair with values of 90 and 100 are both measured to have a value difference of 10, but the effect of value difference can be expected to be much stronger in the (10, 20) pair than in the (90, 100) pair.

We aim to understand the predictive power of the CRT by investigating the two proposed research questions (RQ1A and RQ1B). For RQ1A, a correlational analysis was performed to assess how CRT measurements, including the CRT-RS and CRT-IS, relate to the degree of reliance on intuition proxied by the absolute starting point values estimated from the DDM. If the CRT's original conceptualization holds, one can expect to observe significant negative correlations between CRT measurements and absolute starting point values. For RQ2, we assessed whether the effect that CRT measurements have on the likelihood of accepting gambles is mediated by signed starting point values that represent the effect of intuition. A positive result would indicate that intuitive thinking is indeed the mechanism underlying the CRT's predictive power for superior decision performance. The mediation analysis was performed using the R package 'mediation' (Tingley et al. 2014). The quasi-Bayesian Monte Carlo method (Imai et al. 2010) was used to calculate confidence intervals for assessing the significance of *Average Causal Mediation Effect* (ACME) and *Average Direct Effect* (ADE). The ACME represents the indirect effect an independent variable has on the dependent variable through the mediator. The ADE represents the direct effect of the independent variable on the dependent variable. The ACME and ADE are seen as significant when the 95% confidence interval (95% CI) does not contain zero. Based on the estimates of the ACME and ADE, the percentage of the ACME in the total effect (ACME + ADE) was also calculated.

2.2 | Results

2.2.1 | CRT Measurements

Figure 3 shows the distributions of the CRT-RS and CRT-IS measured from the subjects in the experiment. Both scores take on integer values from 0 to 7 (for CRT-RS: $M+SD=2.762+2.085$; for CRT-IS: $M+SD=3.770+1.957$). Neither score is normally distributed (Kolmogorov–Smirnov tests, $p<0.001$ for both scores). A very high correlation is found between the CRT-RS and CRT-IS ($\rho=0.887$, $p<0.001$), suggesting that these two scales largely capture the variance of the same trait. Moreover, age is found to have a weak negative correlation with the CRT measurements (CRT-RS: $\rho=-0.184$, $p=0.043$; CRT-IS: $\rho=-0.173$, $p=0.056$), and male subjects tend to score higher on this measurement (CRT-RS: Mann–Whitney U test statistic $=2575$, $p<0.001$; CRT-IS: Mann–Whitney U test statistic $=2575$, $p=0.002$).

2.2.2 | Risky Decision-Making Patterns

A total of 24,400 risky decision trials were completed by 122 subjects (200 for each subject). Trials where the decision time was either too short (below 300 ms) or too long (above 10,000 ms) were excluded from the analysis as they likely represent misclicks and procrastination, respectively. As a result, 346 trials are excluded from further analysis (1.4% of the dataset).

Figure 4a shows the acceptance rate by the value difference between gains and losses (gain value—loss value). Overall, subjects tended to accept gambles in the advantageous domain (acceptance rate: 69.8%) and reject gambles in the disadvantageous domain (acceptance rate: 5.5%). In both domains, the acceptance rate increases as the gain value becomes higher relative to the loss value. This pattern is also confirmed by the GLMM analysis summarized in Table 1, where the value difference variable predicts the likelihood of accepting gambles in both domains. Figure 4b shows the mean decision time by the value difference. The distribution has the expected pattern that decision time increases with the difficulty of decision making, that is as gain and loss values become closer. Interestingly, trials with the largest differences between gain and loss values are also shown to have long decision times. This may be because trials with extreme value differences are infrequent, and subjects spent longer time on these less common trials due to unfamiliarity. The GLMM analysis summarized in Table 2 indicates that infrequent trials with extreme value differences do not disrupt the overall pattern between decision time and value difference. In both advantageous and disadvantageous domains, closer value differences between gains and losses consistently predict longer decision times (the value difference variable increases from 10 to 90 in the advantageous domain, and from −90 to −10 in the disadvantageous

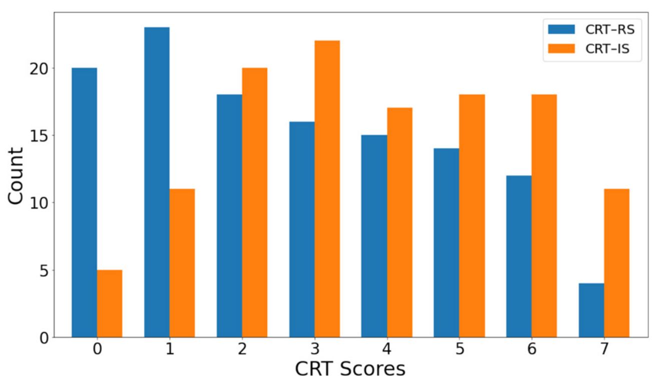

FIGURE 3 | Distributions of the CRT-RS and CRT-IS. The distribution of the CRT-RS is skewed towards lower scores, whereas the CRT-IS tends to have mid-range of scores.

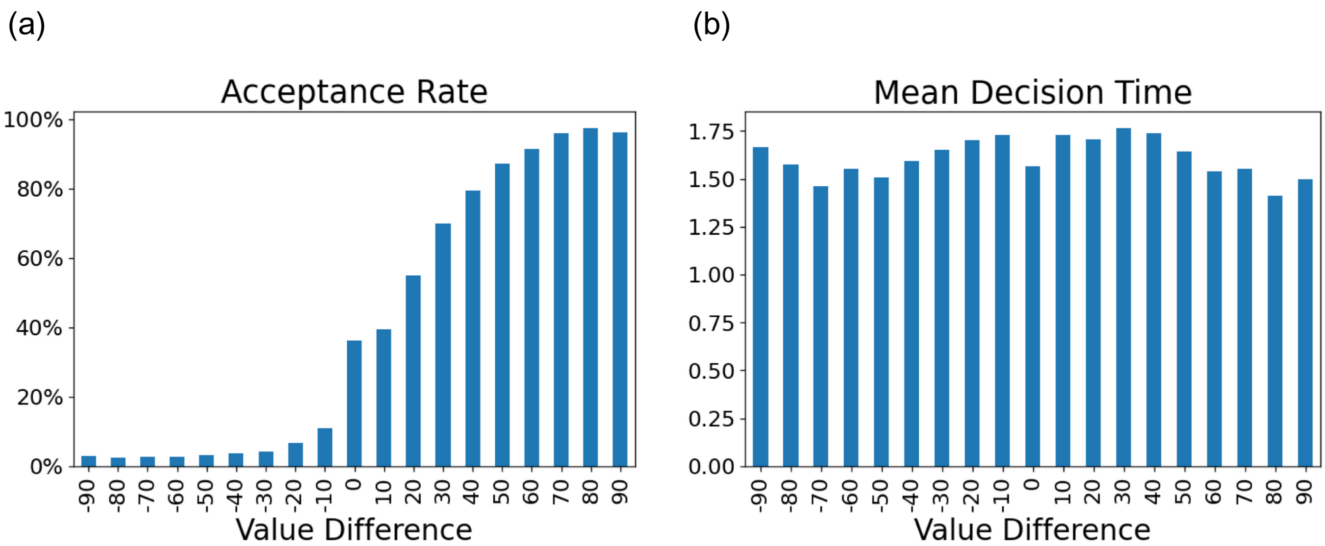

FIGURE 4 | Acceptance rate and mean decision time by value differences in decisions. It should be noted that smaller gaps in the value difference occurred more often than larger ones. The frequencies of absolute value differences (|gain−loss|) in risky decision trials are 90: 2%, 80: 4%, 70: 6%, 60: 8%, 50: 10%, 40: 12%, 30: 14%, 20: 16%, 10: 18%, and 0: 10%.

TABLE 1 | GLMM analysis of factors influencing the likelihood of accepting gambles.

	Model 1 Aggregate		**Model 2 Advantageous**		**Model 3 Disadvantageous**	
Fixed effects†	***β***	**SE**	***β***	**SE**	***β***	**SE**
Intercept	−1.300***	0.139	1.831***	0.182	−4.745***	0.248
Difference	3.184***	0.044	1.984***	0.049	0.810***	0.063
Stake	−0.305***	0.020	−0.839***	0.046	0.128	0.074
CRT-RS	−0.315*	0.139	−0.205	0.180	−0.626**	0.223
Difference x Stake	−0.606***	0.042	−0.267***	0.050	−0.122	0.079
Observations	24,054		10,819		10,838	
AIC	14574.3		7648.6		3088.9	

All variables except for binary variables are scaled and centered values. Logistic regression. Significance levels: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE 2 | GLMM analysis of factors influencing the decision time.

	Model 4 Advantageous		**Model 5 Disadvantageous**	
Fixed effects†	***β***	**SE**	***β***	**SE**
Intercept	1.708***	0.046	1.668***	0.048
Difference	−0.060***	0.007	0.040***	0.006
Stake	0.055***	0.009	−0.014	0.008
CRT-RS	−0.004	0.046	−0.063	0.048
Difference × stake	0.066***	0.011	0.004	0.009
Observations	10,819		10,838	
AIC	24,229.4		22,531.3	

All variables except for binary variables are scaled and centered values. Gamma regression. Significance levels: ***$p < 0.001$.

domain). Altogether, these behavioral patterns suggest that subjects considered the size of value difference between gains and losses in decision making, rather than simply making decisions based on whether the gain or loss value was greater.

The GLMM analysis of how the CRT-RS predicts the likelihood of accepting gambles and decision time in risky decisions is presented in Tables 1 and 2, respectively. The CRT-RS is not found to predict decision time in either advantageous or disadvantageous domain. This result does not support the proposition that high-ability individuals tend to engage in reflective thinking. With respect to the likelihood of accepting gambles, results suggest that the CRT-RS predicts lower likelihood of accepting a gamble at the aggregate level of the data (all trails are included). A closer examination shows that this pattern is driven by the behaviors in the disadvantageous domain where those who score high on the CRT-RS are more likely to submit expected value maximizing responses, that rejecting gambles. We do not find that the CRT-RS predicts the likelihood of accepting gambles in the advantageous domain. The same patterns related to the decision time and likelihood of accepting gambles are also found with the CRT-IS (see Tables S1 and S2 in the Supplement); this result is unsurprising given the high correlation between the CRT-RS and CRT-IS. Together, these findings suggest that high measurements on the CRT predict expected value maximization in the disadvantageous domain, and they do not predict decision types in the advantageous domain.

It is also worth noting that the stake size variable in the GLMM analysis is shown to be a nonsignificant in moderating the effect of value differences in the disadvantageous domain. This pattern, in conjunction with the varying predictive effects of CRT measurements across the two domains, suggests that the decision processes underlying risky preference formation are likely affected by different mechanisms in the advantageous and disadvantageous domains. It is thus necessary to investigate the role of intuition separately in the two domains.

2.2.3 | DDM Analysis

The DDM-1, DDM-2, DDM-1-VSP and DDM-2-VSP were fitted to decision trials in the advantageous and disadvantageous domains. We found near perfect correlations in all conditions between the observed statistics for each subject and the statistics simulated by the estimated DDMs, including the acceptance rate, mean decision time, and decision time at the 10th, 30th, 50th, 70th, and 90th percentiles. All correlation coefficients were above 0.89 with $p < 0.001$ (see Tables S3–S8 in the Supplement). To demonstrate such high correlations, Figure 5 shows the scatter plots for using the DDM-1 to simulate acceptance rate and mean decision time at the aggregate level (decisions from the advantageous and disadvantageous domains pooled together). Additional plots comparing acceptance rates, mean decision times, and decision time percentiles across model forms are provided in the Supplement (see Figures S1–S4); all indicate high correspondence between observed and simulated statistics. These results show that the estimated DDMs accurately capture subject-level variation in both decision outcome and decision time, which are thought to be affected by intuition.

We obtain signed and absolute starting point values from the estimated DDMs. Figure 6 shows the boxplots of signed and absolute starting point values by decision domains. The paired samples t test suggests that, by both the DDM-1 and DDM-2, subjects are estimated to have higher signed starting point values in the advantageous domain (by the DDM-1: $t = 10.341$, $p < 0.001$; by the DDM-2: $t = 3.267$, $p < 0.001$), and higher absolute starting point values in the disadvantageous domain (by the DDM-1: $t = 8.311$, $p < 0.001$; by the DDM-2: $t = 5.562$, $p < 0.001$). These results warranted separate estimation of the DDM starting point in the two domains and provide evidence that intuition affects decision making differently in each domain. Using the likelihood ratio test, the DDM-2 is found to be the better-fitting model than the DDM-1 for 46 out of 122 subjects in the advantageous domain, and 88 out of 122 subjects in the disadvantageous domain. We obtain a set of estimates of signed and absolute starting points from the better-fitting model between the DDM-1 and DDM-2 (the "best-fit model set").

We further examined whether intuition can indeed affect risky decisions. Table 3 presents the GLMM analysis examining the relationship between signed starting point values and the likelihood of accepting gambles. A subject's decision trials in the advantageous domain are associated with the signed starting point estimated from the same domain for the subject, and vice versa for decision trials in the disadvantageous domain. The GLMM analysis suggests that, by both the estimated values from the DDM-1 and the DDM-2, higher values of the signed starting point are associated with a greater likelihood of accepting gambles, providing evidence that intuition can affect risky decisions. Moreover, we examined the relationship between the starting point position and drift rate components that represent

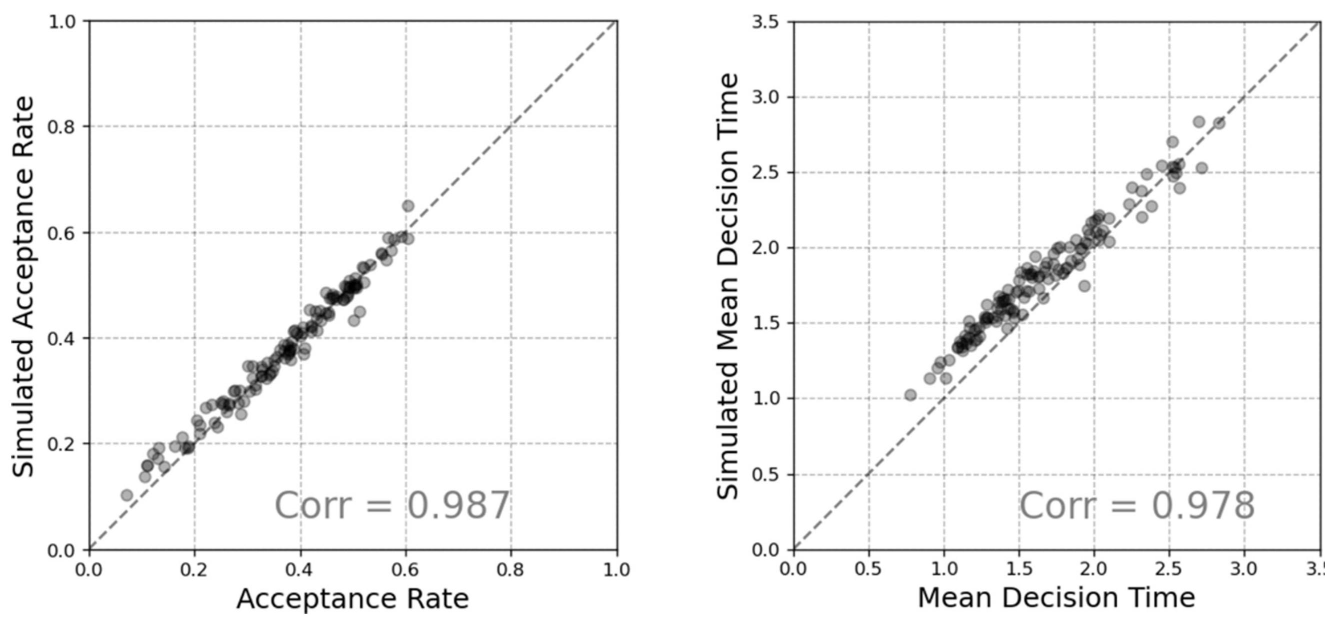

FIGURE 5 | Scatter plots of observed and simulated acceptance rates and decision times. In both plots, the data points are allocated close to the diagonal line, indicating high correlations between the observed and simulated data. 'Corr' corresponds to Pearson's correlation coefficients.

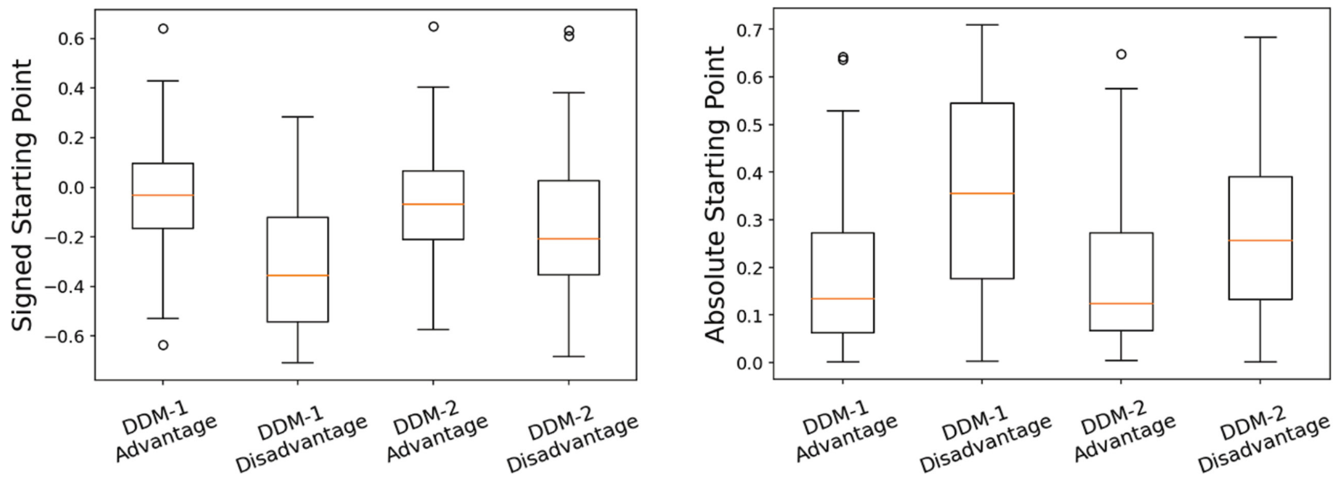

FIGURE 6 | Boxplots for the estimated values of starting points. The box edges represent the first quartile (Q1) and the third quartile (Q3) values of the data; the line in the middle represents the median (Q2). The whiskers extend from the box edges by 1.5 times the interquartile range (Q3–Q1), stopping at the farthest data point within this limit. Outliers are displayed as individual dots.

TABLE 3 | Regression analysis of factors influencing the likelihood of accepting gambles.

	Model 6 DDM-1		Model 7 DDM-2	
Fixed effects†	β	SE	β	SE
Intercept	−1.565***	0.090	−1.404***	0.117
Difference	2.722***	0.045	3.112***	0.045
Stake	−0.278***	0.025	−0.265***	0.024
Signed starting point	1.298***	0.056	0.473***	0.042
Difference × stake	−0.577***	0.045	−0.606***	0.044
Observations	21,657		21,657	
AIC	11,699.1		12,191.6	

All variables except for binary variables are scaled and centered values. Logistic regression. Significance levels: ***$p < 0.001$.

subjective preference arising from evidence accumulation. By the estimates from the best-fit model set, the signed starting point is found to have a weak positive correlation with the utility weighting effect in the advantageous domain (Spearman's correlation coefficient = 0.220, $p = 0.015$). In the disadvantageous domain, the correlation between the signed starting point and the utility weighting effect is nonsignificant (Spearman's correlation coefficient = 0.100, $p = 0.272$). As such, the starting point position does not appear to associate strongly with subjective preference arising from evidence accumulation. Further analysis using the GLMM finds that the utility weighting effect and fixed utility bias estimated with the DDM also associate with the likelihood of accepting gambles when included as independent variables alongside the signed starting point (see Tables S9 and S10, and Section S3.3 of the Supplement for detailed analysis; see also Table S26 for a correlational analysis consistent with the GLMM results). In other words, the starting point and drift rate components independently contribute to the observed choice pattern.

Altogether, these findings suggest that the starting point that represents intuition is not a mere reflection of subjective preference arising from evidence accumulation, and the starting point is not the only factor that could mediate the relationship between CRT measurements and expected value maximization.

In order to address RQ1A, we calculate Pearson's correlation coefficients between the two CRT measurements, the CRT-RS and CRT-IS, and the degree of reliance on intuition represented by absolute starting point values. The results are shown in Table 4. In the advantageous domain, CRT measurements are not found to have significant correlations with absolute starting point values. By contrast, CRT measurements positively correlate with absolute starting point values in the disadvantageous domain. This pattern of correlations is consistent across estimates from the DDM-1, DDM-2 and the best-fit model set. Boxplots that show the distribution of absolute starting point values by the CRT-RS can be seen in Figure S5. Because CRT measurements are categorical variables, we also perform nonparametric analysis by using Spearman's rank correlation; the results are qualitatively identical from those obtained from Pearson's correlation (Table S11). Therefore, the results suggest that in the disadvantageous domain where CRT measurements predict expected value maximization, high-ability individuals tend to rely more on intuition in decision making. Whereas in the advantageous domain in which CRT measurements are not found to predict decision performance, judgment ability in term of the CRT does not associate with the degree of reliance on intuition. To note, while correlational patterns differ across the two decision domains, results from neither domain suggest that high CRT measurements associate with a lower degree of reliance on intuition, as is postulated by the CRT's original conceptualization.

For RQ1B, the mediation analysis was carried out by assessing the significance of ACMEs and ADEs of CRT measurements after introducing the signed starting point as an additional variable in regression models that assess the likelihood of accepting gambles (Model 2, 3, S2, S3). The results are summarized in Table 5. In the advantageous domain, 95% CIs all span across 0, indicating that ACMEs and ADEs for both CRT measurements

TABLE 4 | Pearson's correlation coefficients between the CRT measurements and absolute starting point values.

Domains	CRT Scores	Absolute Starting Point Value		
		DDM-1	DDM-2	Best-fit
Advantageous	CRT-RS	0.077	0.023	0.036
	CRT-IS	−0.068	−0.119	−0.111
Disadvantageous	CRT-RS	0.246**	0.296**	0.282**
	CRT-IS	0.203*	0.206*	0.198*

Observations = 122. Significance levels: * $p < 0.05$; ** $p < 0.01$.

TABLE 5 | The mediating effects of the starting point.

Domain	CRT score	Starting point†	ACME		ADE		Percentage mediated
			Estimate	95% CI	Estimate	95% CI	
Advantageous	CRT-RS	DDM-1	−0.019	[−0.048, 0.008]	−0.003	[−0.030, 0.022]	75.8%
Advantageous	CRT-RS	DDM-2	−0.019	[−0.042, 0.004]	−0.004	[−0.036, 0.027]	69.8%
Advantageous	CRT-RS	Best-fit	−0.018	[−0.040, 0.003]	−0.004	[−0.035, 0.028]	67.9%
Advantageous	CRT-RS	DDM-1	−0.018	[−0.046, 0.007]	−0.008	[−0.037, 0.018]	64.8%
Advantageous	CRT-RS	DDM-2	−0.015	[−0.036, 0.005]	−0.012	[−0.044, 0.019]	50.8%
Advantageous	CRT-RS	Best-fit	−0.015	[−0.038, 0.005]	−0.012	[−0.043, 0.019]	50.0%
Disadvantageous	CRT-RS	DDM-1	−0.014	[−0.026, −0.004]	−0.004	[−0.014, 0.007]	75.9%
Disadvantageous	CRT-RS	DDM-2	−0.005	[−0.014, 0.004]	−0.015	[−0.028, −0.001]	23.5%
Disadvantageous	CRT-RS	Best-fit	−0.004	[−0.014, 0.005]	−0.015	[−0.029, −0.004]	22.4%
Disadvantageous	CRT-RS	DDM-1	−0.011	[−0.023, −0.001]	−0.015	[−0.016, 0.005]	66.7%
Disadvantageous	CRT-RS	DDM-2	−0.003	[−0.012, 0.006]	−0.015	[−0.028, −0.003]	16.3%
Disadvantageous	CRT-RS	Best-fit	−0.003	[−0.012, 0.007]	−0.015	[−0.028, −0.003]	15.9%

The type of the DDM from which the mediator signed starting point is estimated. Observations = 10,819 for analyses in the advantageous domain, and observations = 10,838 for analyses in the disadvantageous domain. Number of simulations = 1000. Seed = 12345.

are nonsignificant regardless of whether estimates from the DDM-1, DDM-2 or best-fit model set are used. This is expected because CRT measurements neither predict the decision type nor correlate with signed starting point values in this domain (the correlational analysis is presented in Tables S12, S13 and Figure S6). In the disadvantageous domain, the patterns are consistent between the CRT-RS and CRT-IS but differ by the set of starting point estimates used. With estimated values from the DDM-1, the ACMEs are significant for the CRT measurements and account for about 70% of the total effects of the CRT's predictive power, leaving the associated ADEs nonsignificant. However, the percentage of total effects mediated by the signed starting point drops to around 20% with estimates from the DDM-2 and the best-fit model set, and the ACMEs become nonsignificant. In this case, the predictive power of CRT measurements is not mediated by the signed starting point that represents the effect of intuitive thinking. The inconsistency in the results can be attributed to the difference in the assumptions made by the models. With the assumption that the risky decisions do not involve a valuation process independent from task-specific conditions (DDM-1), we found evidence that intuition mediates the predicative power of CRT measurements for expected value maximization. However, the model with this assumption does not provide the best fit to the data for all decision makers. When the evaluation process is allowed to have a fixed utility bias, the mediating effect of intuition becomes nonsignificant. Overall, our analysis suggests that intuition does not contribute to the CRT's predicative power for decision performance.

To assess the robustness of our findings in relation to RQ1A, we used the deviation between each participant's signed starting point value and the value predicted from their choice proportion as a behavior-relative measure of reliance on intuition. By this alternative measure, the findings are consistent with those in the main analysis, where the reliance on intuition is measured using the absolute starting point. Specifically, in the advantageous domain, the deviation does not significantly correlate with the CRT-RS (Pearson's correlation coefficient = −0.040, $p = 0.664$) based on the estimates from the best-fit model set, whereas a positive correlation is observed in the disadvantageous domain (Pearson's correlation coefficient = 0.188, $p = 0.038$). A detailed analysis of this robustness test is presented in S3.5 in Supplement, with supporting results shown in Figure S7 and Tables S14 and S15. In line with the main analysis, these results suggest that high-ability individuals relied more on intuition in the disadvantageous domain, but not in the advantageous domain.

For another robustness test, we examine estimates from the DDM-1-VSP and DDM-2-VSP, where the starting point position was estimated separately for the first and second half of decision trials. By both models, subjects in the disadvantageous domain were estimated to have higher signed starting point values from the first half of the decision trials than from the second half (by the DDM-1-VSP: $t = 7.381$, $p < 0.001$; by the DDM-2-VSP: $t = 7.708$, $p < 0.001$). Whereas signed starting point values do not differ significantly between the first and second half of the decision trials in the advantageous domain (by the DDM-1-VSP: $t = -0.773$, $p = 0.441$; by the DDM-2-VSP: $t = -0.853$, $p = 0.395$). With the absolute starting point, estimates from the second half of decision trials are greater than those from the first half in both the advantageous domain (by the DDM-1-VSP: $t = 3.750$, $p < 0.001$; by the DDM-2-VSP: $t = 3.044$, $p = 0.001$) and disadvantageous domain (by the DDM-1-VSP: $t = 5.264$, $p < 0.001$; by the DDM-2-VSP: $t = 3.807$, $p < 0.001$). The distribution of estimated starting point values can be viewed in Figure S8. These findings are in line with the proposition that intuition thinking can develop as experience accumulates. Importantly, results related to our research questions are consistent with those from the DDM-1 and DDM-2, that is high-ability individuals relied more strongly on intuition in the disadvantageous domain, and intuition contributed to the CRT's predictive power in this domain by estimates from the DDM-1-VSP, but not by estimates from the DDM-2-VSP and the best-fit model set. Correlational and mediation analyses addressing RQ1A and RQ1B are presented in Tables S16–S20 in Section S3.6 of the Supplement.

2.3 | Discussion

In our experiment, CRT measurements predict normatively correct risky decisions in the disadvantageous, but not advantageous domain. In other words, we do not find that CRT measurements consistently predict superior performances. Such inconsistency is also found in the systematic review by Lilleholt (2019), who hence argues that the association between CRT measurements, as well as other measurements of cognitive ability, and risk preference may be domain-specific. Importantly, in the disadvantageous domain where CRT measurements are indeed found to predict expected value maximizing responses, DDM analysis provides evidence against the CRT's original conceptualization regarding the role of intuition. To begin with, the degree of reliance on intuition is found to be stronger among high-ability individuals; this result stands in contradiction to the conceptualization that high-ability individuals tend to inhibit intuitive thinking. With respect to how intuition contributes to the CRT's predictive power for superior performance, the results differ from the assumptions made for the risky preference formation. With the assumption that risky preference formation is affected by intuition and utility weighting, as suggested by prospect theory, DDM analysis suggests that intuition helps high-ability individuals to achieve expected value maximization. However, this relationship diminishes with the additional assumption that risky preference formation involves a valuation process independent from task-specific conditions. Overall, results from the best-fit DDM set indicate that intuition does not contribute to the association between CRT measurements and decision performance. Regardless, these results contradict to the CRT's original conceptualization that intuition is error-prone and causes decision makers to deviate from expected value maximizing responses. Consequently, the findings from the present study suggest the need to reassess how the implications from previous studies would vary if high-ability individuals are instead assumed to be more reliant on intuition, particularly in decisions that involve risk assessment (e.g., Bucher-Koenen and Ziegelmeyer 2014; Thoma et al. 2015).

As it was mentioned, studies on the topic have employed various methods to evaluate intuition and have produced conflicting results, including results that are in favor of the CRT's original conceptualization (e.g., Campitelli and Gerrans 2014; Travers et al. 2016), and results that indicate intuition is irrelevant to the CRT (e.g., Otero et al. 2022; Sinayev and Peters 2015). One question that needs to be addressed is how the DDM analysis should be viewed when there are conflicting results from studies using other methods to evaluate intuition. To this, we argue that the validity of the DDM, as a tool for evaluating intuition, is demonstrated by its ability to accurately simulate key behavioral patterns based on a framework that aligns with how intuition is thought to affect decisions. Fast judgment that yields default response is a defining feature of intuition thinking; hence, the involvement of intuition is thought to affect both decision outcome and decision time patterns (Evans and Stanovich 2013; Kahneman 2011). Similarly, the DDM starting point affects both these patterns by generating a bias at the onset of decision making; a shifted starting point towards one option would result in higher likelihood and faster decision times for choosing that option (and lower likelihood and longer decision times for choosing the opposite option). The fact that estimates of DDM starting point, as well as other DDM components, are capable of simulating the decision outcome and decision time patterns in the real data (Section 3) provide evidence that the DDM starting point adequately represents how intuition affects decisions. This feature that is not shared by other methods for evaluating intuition highlights the value of using the DDM to inform the ongoing debate about how intuition contributes to the CRT's predictive power for decision performance.

While the internal validity of our findings is supported by the DDM's capacity to simulate key behavioral patterns, further research is needed to explore whether these findings can be generalized to other contexts where the CRT is found to have predictive power. As it has been discussed, the CRT is found to predict behaviors and performance in a variety of contexts, and it is possible CRT measurements may relate to intuition differently depending on the specific context. To assess the generalizability of the present findings, future research could apply the DDM to examine how CRT measurements relate to intuition in other domains, such as social and intertemporal decision making, where behaviors have also been linked to the CRT and analyzed using the DDM (Chen and Krajbich 2018; Corgnet et al. 2015; Frederick 2005; Zhao et al. 2019). Moreover, future research may examine how task complexity affects the role of intuition in decision making. The risky decision tasks in our experiment are relatively simple. In order to deduce the expected value maximizing response, subjects only needed to compare the sizes of gain and loss values offered in each decision trial. It has been hypothesized that overly simple decision tasks could undermine the utility of relying on intuitive ability for making inferences, and the association between cognitive ability, such as that measured by the CRT, and effective use of intuition would become more pronounced in more complex decision problems (Ghasemi et al. 2022). Nevertheless, the opposite pattern is also plausible. Previous research using self-reports has found that decision makers claim to rely more on reflective, rather than intuitive, thinking as decisions become more complex (Inbar et al. 2010). To address this issue, future work can include studies that manipulate the complexity of the decision tasks administered to participants. For instance, researchers could compare and contrast how CRT measurements associate with intuitive thinking in simple risky decisions, such as those in our experiment, and more complex ones, such as risky decisions whose assessments require combining probabilities and gain or loss values. Overall, these investigations will help advance our understanding of how performance on the CRT relates to intuitive processes across different contexts and contribute to evaluating the generalizability of the present findings.

3 | Study 2

Study 2 aims to understand the cross-domain stability of the preference for intuition; we investigated whether the preference for intuition is stable across risky decision and social decision domains. Intuitive thinking has been found to play an important role in both risky decisions and social decisions (Chen and Krajbich 2018; Montero-Porras et al. 2022; Rand et al. 2012; Reyna 2004; Sheng et al. 2020; Zhao et al. 2020, 2022). However, these two decision domains involve distinct motivations related to how the effects of the decisions are distributed among different actors. In risky decisions, a decision maker needs to evaluate options that may have multiple outcomes that will occur at certain probabilities, but all outcomes affect only the decision maker herself. By contrast, in social decisions, whilst the decision maker needs to evaluate options that may again involve multiple outcomes, these outcomes affect not only the decision maker himself but also others. The motivations related to these two domains are also called 'self-regarding motivation' in the risky domain, and 'other-regarding motivation' in the social domain and, broadly speaking, all decisions are formed by varying degrees of self- and other-regarding motivations (Grund et al. 2013). Therefore, the findings from the present study will have broad implications because, to an extent, the investigation in the present study can be seen as a way to assess the stability of the preference for intuitive thinking in both, self- and other-regarding decisions.

The present analysis uses the experimental data from Smith and Krajbich (2018), in which subjects complete a series of risky decisions in the form of lottery choices, and social decisions in the form of mini-dictator games. For both decision domains, we use the DDM to measure each subject's preference for intuition by her degree of reliance on intuition in decision making. It should be noted that the cross-domain stability of the preference for intuitive thinking can be interpreted in different ways. From the perspective of an individual decision maker, the preference can be seen as stable if the extent to which she relies on intuition for making decisions is comparable across different domains. Alternatively, from the perspective of individual differences, the preference for intuition across domains can be seen as stable if the pattern of individual differences in one domain is indicative of the patterns in other domains. This is plausible because, even if the domain-specific preferences for intuition of an individual decision maker are not at a comparable level, they may all strongly relate to the decision maker's domain-general preference. In this case, the degree of reliance on intuition in risky decision making should correlate with that in social decision making. Based on these different interpretations about the stability of preference for intuition, the following two specific research questions regarding the decision making of subjects in the experiment are proposed to guide the analysis:

> Research Question 2A (RQ2A): Do subjects have similar degrees of reliance on intuition in the risky and social decisions?
>
> Research Question 2B (RQ2B): Is there a correlational relationship between the degree of reliance on intuition in the risky and social decisions?

Answers to these two questions will test whether and how the preference for intuition is stable across decision domains.

3.1 | Materials & Methods

3.1.1 | The Experiment

The risky and social decisions that we analyze are collected from the experiment conducted by Smith and Krajbich (2018), whose data is publicly available in the Open Science Framework.[2] In the experiment, 36 subjects completed 200 risky decisions in the form of monetary gambles, and 200 social decisions in the form of mini-dictator games. A representative decision trial for each of these two types of decisions is shown in Figure 7.

For each decision trial in the risky decisions case, a subject was asked to choose between accepting two gambles, where each gamble had an even probability (50/50) of winning a larger or a smaller monetary reward. Subjects were presented with a graphical display similar to that shown in Figure 7a (without orange labels and boxes). The amounts of monetary rewards are indicated by the height of the white bars displayed on the screen. A bar that fills the whole box of 300 pixels high indicates the highest reward of $10. As can be seen in Figure 7a, four white bars are displayed, the top-left and bottom-left bars consist of one gamble, and the top-right and bottom-right bars consist of another gamble. Importantly, the gambles are designed to represent a riskier option and a safer option in each decision trial, such that the riskier option has higher maximum reward but also higher variability between the rewards, and the safer option has less variable rewards but the maximum reward is lower than that in the riskier option. The subjects thus need to evaluate the risk and potential rewards and choose between accepting the riskier or the safer gamble.

For each trial in the social decisions trial, the subjects were presented with a display similar to the one shown in Figure 7b, and asked to choose between two options with different monetary rewards for the subject herself and another subject. These types

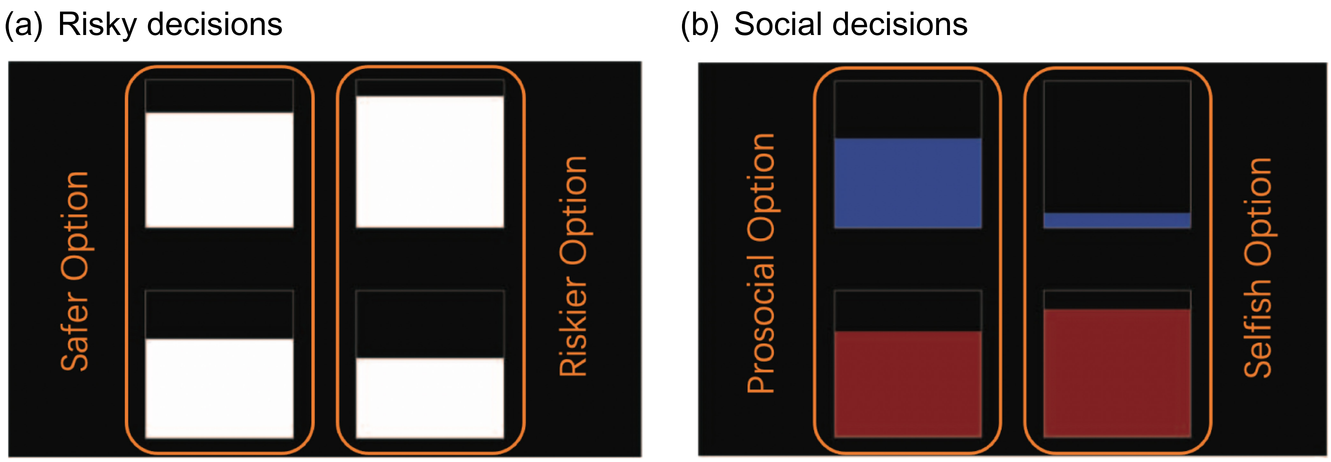

FIGURE 7 | Representative decision trials. The side of the display of different types of options (riskier and safer options in risky decisions; prosocial and selfish options in social decisions) on the screen is randomized across decision trials. These figures are adapted from Smith and Krajbich (2018).

of social decisions are often referred to as mini-dictator games, in which the subjects have the power to allocate money between themselves (dictator) and another subject (receiver). The monetary reward for the subject herself is indicated by the heights of the red bars, and the reward to be received by the other subject is indicated by the heights of the blue bars. The highest reward that corresponds to a bar filling the whole box is $5. As in the risky decision trials, the red and blue bars shown on the left side of the screen represent one option, and the two bars shown on the right side represent another option. In each decision, the two options can be thought of as a pair consisting of a selfish option and a prosocial option. In comparison with the prosocial option, the reward for oneself is greater in the selfish option but the reward for the other is lower. The subject needs to evaluate the trade-off between the outcome to herself and the outcome to the other and choose between accepting the selfish or the prosocial option.

In both decision domains, subjects indicated their decisions by pressing keys on the keyboard. The option on the left can be selected by pressing the "F" key, and the option on the right can be selected by pressing the "J" key. The decision outcome and decision time information from each decision trial were recorded. In addition, subjects' left-eye fixation locations on the computer screen were also recorded by an eye-tracker. The recording was performed at 1000 Hz.

3.1.2 | Drift Diffusion Modeling

Analogous to the modeling approach in Study 1, the DDM was separately fitted to each subject's risky decision trials and social decision trials. The DDM was used to assess how a subject chose between the riskier option (upper threshold) and the safter option (lower threshold) in risky decision making, as well as between the selfish option (upper threshold) and the prosocial option (lower threshold) in social decision making. The absolute starting point value is used to represent the decision maker's degree of reliance on intuition.

Two forms of the DDM were estimated for the main analysis. The first form is the standard DDM (sDDM) estimated based on the decision outcomes, decision times and reward values in each decision trial. The DDM starting point, threshold and non-decision time components were estimated as constant values throughout the decision trials. The drift rate component was estimated as utility functions that vary across decision trials depending on reward values, following the practice in Smith and Krajbich (2018). For the risky decisions, the utility function follows the assumption of constant relative risk aversion. For the social decisions, the utility function follows the Charness-Rabin model (Charness and Rabin 2002). As such, decision makers' subjective utilities are captured by the drift rate component, and individual differences in choices do not necessarily reflect variations in the starting point that represents intuition. Secondly, a special form of the DDM called attentional drift diffusion model (aDDM) was estimated. This model requires visual fixation data as an additional input in model estimation, and has an additional attentional discounting component, which captures the value-amplifying effect of visual attention on the evaluation of the available options (Krajbich et al. 2010). The attentional discounting component in the aDDM is estimated as a constant value throughout the decision trials. As in Study 1, we performed two robustness tests to assess the findings from the main analysis. Firstly, we calculated the absolute deviation between each participant's signed starting point value and the value predicted from a linear regression on their choice proportions. Choice proportion is defined as the proportion of decisions in which the riskier option was chosen in risky decision making, and the proportion of decisions in which the selfish option was chosen in social decision making. This deviation reflects the extent to which a participant's starting point position is more extreme than would be expected based on their observed choices, providing a behavior-relative measure of reliance on intuition. Secondly, we estimated two additional forms of the DDM, sDDM-VSP and aDDM-VSP, as a robustness test. In these models, the starting point was estimated separately for the first and second half of the decision trials used for model fitting in sDDM and aDDM. The results will help to understand whether the preference for intuition across decision domains may converge or diverge over time. The details about the modeling are given in the Supplement (Section S4). The estimation of DDM components was performed using PyDDM.

3.1.3 | Statistical Testing

Given the relatively small sample of subjects in the experiment ($N = 36$), the assessments of the research questions relied on non-parametric statistical testing methods. Specifically, RQ2A was assessed by comparing the absolute starting point values estimated from the risky and social decisions using the Wilcoxon signed-rank test. For RQ2B, Spearman's rank correlation was used to assess the correlational relationship between the estimated absolute starting point values from the two decision domains. Furthermore, for each research question, the results from using the estimated values obtained from both the sDDM and aDDM were inspected to determine if a robust pattern could be identified in both scenarios. To note, comparing the goodness-of-fit between the two models is challenging because the samples used for the sDDM and aDDM analyses were different. This difference occurred because recordings of eye fixation locations failed in some decision trials. However, we argue that the aDDM is a more accurate representation of decision making than the sDDM because the former takes into account visual attention that is known to affect the valuation of available options. Consequently, the results from the aDDM will be given more weight in addressing the research questions.

3.2 | Results

Each of the 36 subjects completed 200 risky decisions and 200 social decisions, amounting to a total of 7200 decisions in each decision domain. Decision trials where the decision time was either too short (below 300 ms) or too long (above 10 s) were excluded. As a result, in the fitting of the sDDM, 142 and 48 trials are excluded from the risky and social decision domains, respectively. For the aDDM, a larger amount of decision trials is removed in the model fitting due to additional exclusion of decision trials with failed recordings of eye fixation locations, resulting in 988 and 897 trials being excluded from the risky and social decision domains, respectively.

The estimated DDMs are shown they can accurately simulate decision outcome and decision time patterns in both the risky and social decision domains. Figure 8 shows the scatter plots of the observed statistics against the simulated data by the aDDM, including choice probability and mean decision time. Results for simulated decision time percentiles are shown in Figures S10, S11 and S14a of the Supplement. In all scatter plots, the data points are close to the diagonal line, indicating high predictive accuracy of the simulations. Pearson's correlations suggest that the simulated mean decision times and rates of choosing certain types of options are highly correlated with the observed statistics; all correlation coefficients are above 0.97 with $p < 0.001$. A similar pattern is also found with the simulations generated by the sDDM (Figures S9, S12, S13, and S14b). These results validate the DDM as tool for evaluating intuition.

With respect to the signed starting point that represents the effect that intuition has on decisions, the distributions of its estimated values from the aDDM are shown in Figure 9 (the distributions of the attentional discounting component estimated from the aDDM are shown in Figure S16 of the Supplement). For both the risky and social decision domains, there exist considerable variations in the signed starting point values (in the risky decision domain: $M + SD = 0.049 + 0.189$; in the social decision domain: $M + SD = 0.218 + 0.165$). Twenty-one subjects from the risky decision domain and 32 subjects from the social decision domain are estimated to have positive values of the signed starting point. At the group level, the estimated values were not found to significantly differ from 0 in the risky decision domain (W-statistic $= 228$, Z-statistic $= -1.650$, $p = 0.099$, effect size $= -0.275$), whereas the estimated values are greater than zero in the social decision domain (W-statistic $= 25$, Z-statistic $= -4.839$, $p < 0.001$, effect size $= -0.806$). Taken together, these results suggest that intuition does not cause decision makers to favor choosing riskier or safer options in the risky decision domain. By contrast, in the social decisions, intuition causes decision makers to favor choosing selfish options. A similar pattern is also found with the estimated starting point values from the sDDM (Figure S15). To note, in both decision domains, the signed starting point is observed to moderately correlate with drift rate components that index risk and social preferences in evidence accumulation (Spearman's correlation coefficients range between −0.3 to −0.5; Table S21). Moreover, using these estimated drift rate components, we calculated subjective values of each option in risky decisions, following the constant relative risk aversion model, and social decisions, following the Charness-Rabin model. The GLMM analysis shows that both the signed starting point and subjective values associate with choices in the two decision domains (Tables S22 and S23; see also Table S27 for a correlational analysis with results consistent with those from the GLMM). Therefore, while the starting point position appears to associate with subjective preference arising from evidence accumulation, the two independently explain variance in the likelihood of choices. In other words, the starting point position that represents intuition is not primarily determined by subjective preferences arising from evidence accumulation. Based on estimated values of the signed starting point, the absolute starting point values that represent the degree of reliance on intuition are calculated.

The boxplots in Figure 10 show the comparisons between the ranges of absolute starting point values in the risky and social decision domains. As indicated by the plots, the absolute starting point values in the social decision domain appear to have higher mean and variance than those in the risky decision domain. Using the Wilcoxon signed-rank test, evidence is found that, at the group level, the absolute starting point values in the social decision domain are greater than the values in the risky decision domain by the estimated results from the aDDM (W-statistic $= 492$, Z-statistic $= 2.498$, $p = 0.006$, effect size $= 0.416$). The same pattern is also found with the estimated results from the sDDM (W-statistic $= 449$, Z-statistic $= 1.822$, $p = 0.034$, effect size $= 0.304$). These results suggest that subjects relied more on intuition in the social decision domain than in the risky decision domain, which addresses RQ2A.

Figure 11 shows the scatter plots for the absolute starting point values from the two decision domains plotted against each other by the estimated values from both the aDDM and sDDM. While the best fitting lines in both figures show a slight downward trend, the Spearman's rank correlation between the absolute

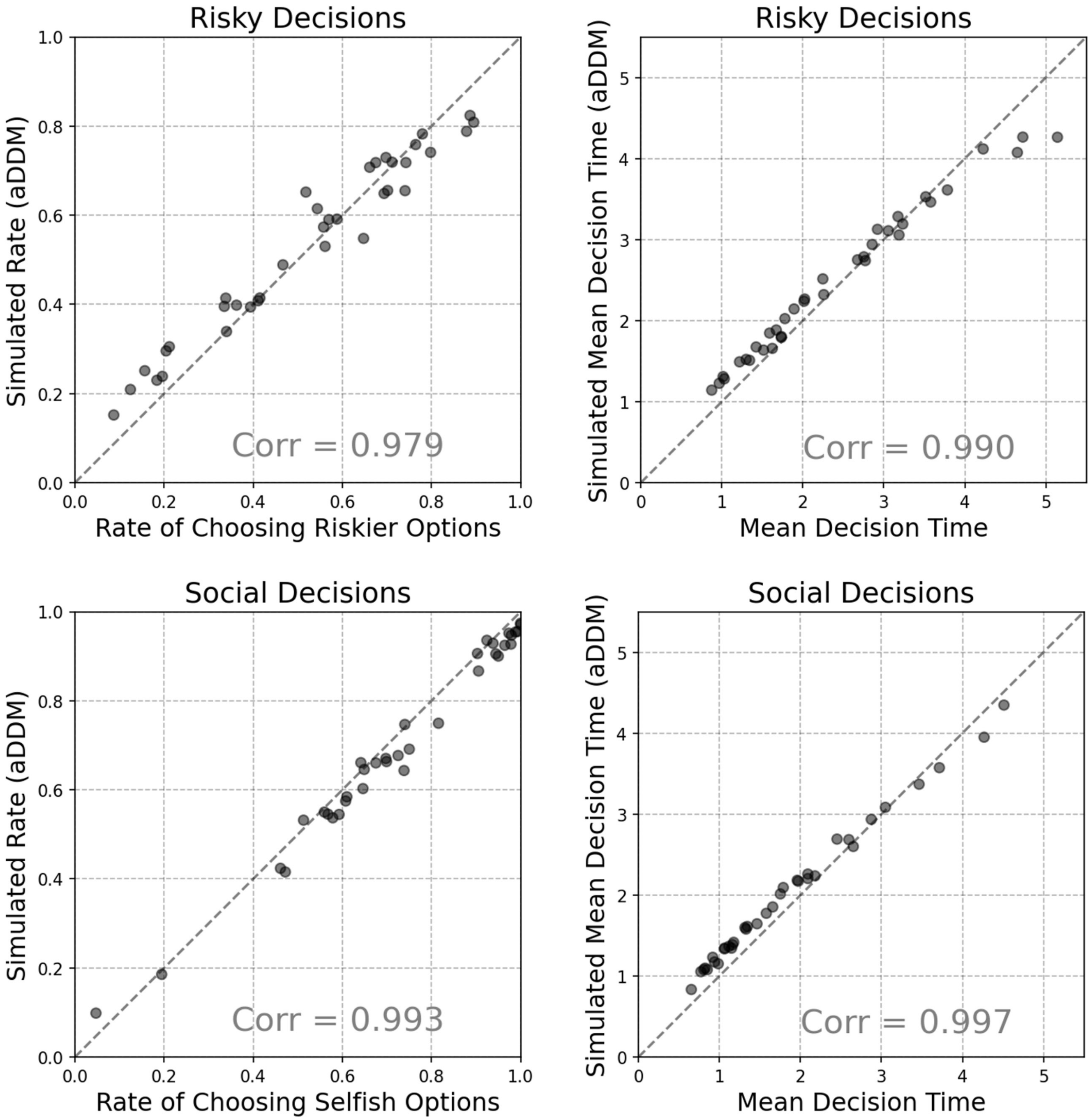

FIGURE 8 | Scatter plots for the observed and simulated statistics. Observed and simulated rates of choosing a certain type of options and mean decision times. Corr corresponds to Pearson's correlation coefficients.

starting point values from the two decision domains is not significant in either scenario. For the estimated values from the aDDM, the correlation coefficient is -0.247 ($p=0.147$), and for the estimated values from the sDDM, the correlation coefficient is -0.067 ($p=0.697$). In other words, we do not find a significant correlational relationship between the degrees of reliance on intuition in the risky and social decisions, which addresses RQ2B.

For a robustness test, we use the behavior-relative measure of reliance on intuition to assess the research questions. In this approach, reliance on intuition is quantified by the deviation between each participant's signed starting point value and the value predicted from their choice proportion. A detailed analysis of this robustness test is presented in S5.4 in Supplement (see Tables S24–S25 and Figures S17–S19). By both measures derived from the sDDM and aDDM, reliance on intuition in the risky and social decisions do not correlate significantly with each other. These results align with those from the main analysis regarding the RQ2B. However, with respect to RQ2A, we find no significant difference in the degree of reliance on intuition between the two decision domains. This contrasts with the main analysis, where reliance on intuition was found to be stronger in social decisions than in risky decisions. As it was discussed, the starting point position estimated from both risky and social decisions moderately

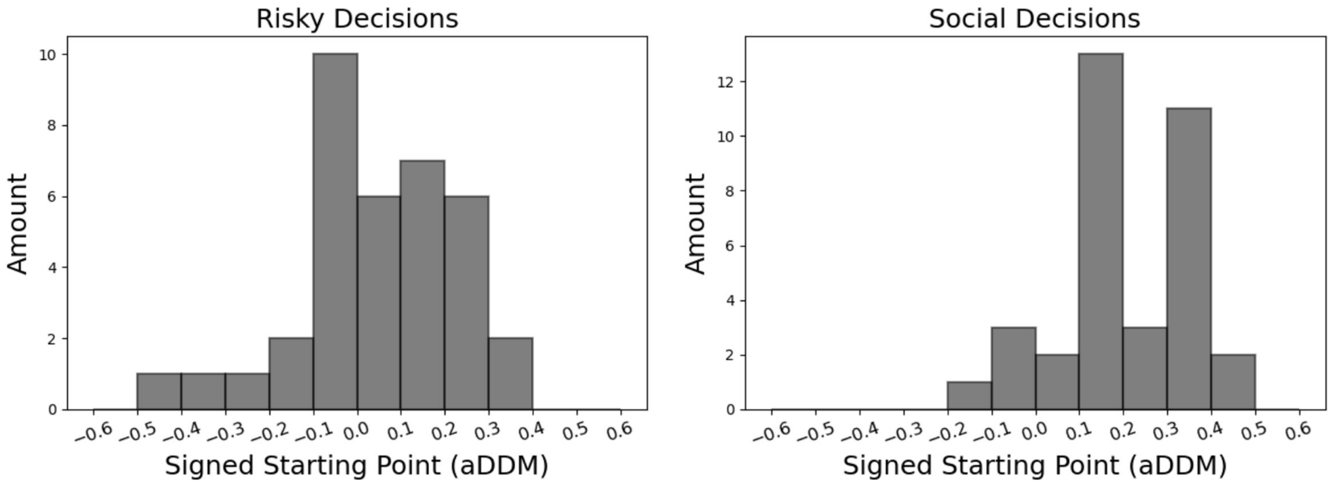

FIGURE 9 | Distribution of signed starting point values. Distributions of the estimated starting point values in the risky and social decision domains.

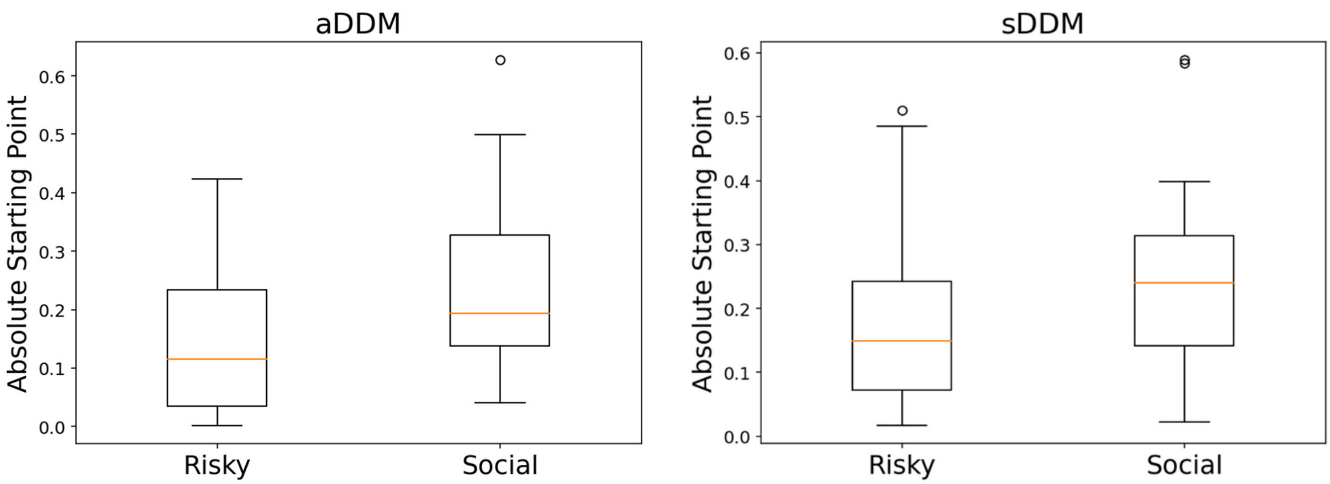

FIGURE 10 | Boxplots for the absolute starting point values. By both the estimated values from the aDDM and sDDM, the distributions of absolute starting point values in the social decision domains occupy higher value ranges than their counterparts in the risky decision domain.

correlate with subjective preference arising from evidence accumulation. Moreover, decision makers tended to favor a particular type of options more strongly in social than in risky decisions (by samples used for aDDM analysis, proportion of decisions being choosing riskier options: $M + SD = 0.523 + 0.237$; proportion of decisions being choosing selfish options: $M + SD = 0.734 + 0.230$). Therefore, the cross-domain difference in reliance on intuition observed in the main analysis may, in part, reflect underlying subjective preferences rather than a stable, domain-specific tendency to rely on intuition.

We further assess the robustness of our findings by examining research questions using starting point estimates obtained from the sDDM-VSP and aDDM-VSP. Plots of the value distributions are shown in Figure S20, and statistical analyses for this robustness test are presented in Section S5.4 of the Supplement. The results are consistent between the two forms of the DDM. We find that the absolute starting point value is estimated to be greater in the second half of decision trials in both risky and social decision making, indicating that subjects became more reliant on intuition over the course of decision-making tasks. In the first half of decision trials, the absolute starting point values are greater in the social decision domain than in the risky decision domain. However, the absolute starting point values from the two decision domains become comparable in the second half of decision trails. These patterns suggest that, while subjects tend to rely more on intuition in the social decision domain than in the risky decision domain, accumulated experience increases their reliance on intuition over the course of decision making and moderates the influence of decision domains on their preference for intuition. With respect to the pattern of individual differences, by estimates from both the first and second half of the decision trials, no significant correlations are found between the absolute starting point values across the two decision domains. The results of the correlational analyses align with those derived from the sDDM and aDDM, highlighting the instability of the preference for intuition.

3.3 | Discussion

The unstable preference for intuition found by the DDM analysis in the present study broadly accords with findings from

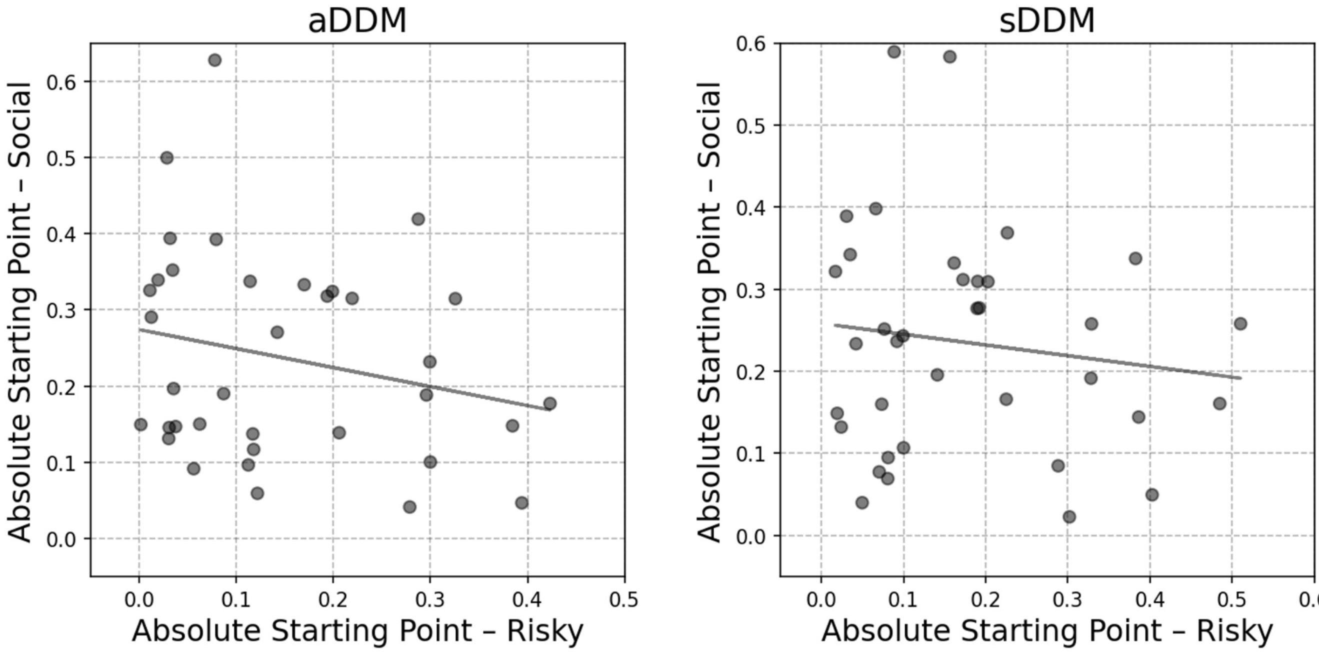

FIGURE 11 | Scatter plots for the absolute starting point values. The horizontal and vertical axes represent the absolute starting point values estimated from the risky and social decision domains, respectively. The solid lines in the plots are the best fitting lines for the data.

previous studies that investigated the topic using self-report questionnaires. Firstly, we found that decision makers tend to rely more on intuition in the social decision domain than in the risky decision domain. This result aligns with the finding that self-reported reliance on intuition vary across decision domains (Berman et al. 2018; Gallo et al. 2017; Inbar et al. 2010; Oktar and Lombrozo 2022; Pachur and Spaar 2015). The difference in the degree of reliance on intuition between risky and social decision making may be attributed to the roles of authenticity in the two decision domains. As identified in prior research by Oktar and Lombrozo (2022), decision makers report to rely more on intuition in decisions where authenticity is thought to be important and decisions need to be reflective of true self. It is plausible that decision makers placed a greater emphasis on authenticity in social decisions than in risky decisions because the former involves other-regarding motivation. Such a speculation is not without reason as Berman et al. (2018) find that, compared with financial investments, decision makers are less responsive to objective outcome metrics in donation decisions and frequently report to choose less effective donation options that represent more subjectively preferred causes. Similarly, in social decisions, decision makers may feel the need to be authentic and decide based on intuitive beliefs about the importance of others' well-being, rather than to decide based on careful examination of reward values for oneself and others. However, it should be noted that the stronger reliance on intuition observed in social compared with risky decisions was not replicated in the robustness test, which used a behavior-relative measure to account for the influence of subjective preference. This suggests that the observed difference may, in part, reflect underlying subjective preferences regarding economic incentives, rather than a stable, domain-specific tendency to rely on intuition. Furthermore, it is important to point out that in both decision domains, decision makers became more reliant on intuition over the course of decision making, with the degree of reliance converging in the second half of the decision trials. This finding supports the idea that experience accumulation fosters increased reliance on fast, intuitive judgment, and the level of experience in a decision domain can predict the preference of intuition in the domain (De Neys and Pennycook 2019; Pachur and Spaar 2015). Overall, our findings suggest that reliance on intuition can vary across contexts and is shaped by underlying subjective preferences and the level of experience. In our study, decision makers tended to favor a particular type of response more strongly in social than in risky decisions, which may have contributed to greater reliance on intuition in social decisions. Additionally, increased familiarity over time fostered more intuitive processing in both decision domains. These findings indicate that reliance on intuition is not rigidly domain-specific but emerges from dynamic interactions between context, preferences, and experience.

Secondly, we found that the degree of reliance on intuition in the risky decision domain does not correlate with that in the social decision domain. This non-correlation holds both for the measured reliance on intuition in the first and second half of decision trials, as well as when using the behavior-relative measure of reliance on intuition. In other words, individual differences in the preference for intuition in one decision domain do not predict the preference in another domain. Previously, Pachur and Spaar (2015) show that correlations between self-reported preferences for intuition across different decision domain are modest to weak. Our findings indicate that correlations between the preference for intuition across domains may be even weaker than it was previously found. They also suggest that measuring decision makers' domain-general preferences for intuition is not a reliable method for representing preferences within specific domains. Taken together, findings from the present study confirm previous research that finds unstable preference for intuition using results from self-reports. Additionally, our findings highlight that the tendency to use intuition in decision making is shaped by characteristics of decision domains, subjective preferences, and accumulated experience, rather than being a stable, individual trait. These results challenge the notion of a consistent, domain-general preference for intuition and emphasize the importance of considering both task-specific preference and experiential factors when studying intuitive judgment.

One limitation of the present study is the relatively small sample size (decisions were collected from 36 subjects), which may limit the generalizability of the findings. Future research should seek to replicate these results in larger samples to confirm the robustness of the patterns observed here. Another important direction for future work concerns the validity of using self-report to measure the preference for intuition. As it has been discussed in this section, results from the DDM analysis are broadly consistent with those obtained from self-reports in previous research. This consistency indicates that decision makers may be more sentient about the use of intuition than that suggested by extant evidence. While decision makers cannot accurately judge their ability at using intuition to assist decision making (Leach and Weick 2018), they may be capable of assessing their preference for intuition in hypothetical contexts. In future work, it will be important to examine whether self-reported use of intuition aligns with DDM measurements. The results will help to understand whether the preference for intuition can be reliably assessed by self-reports. The findings from this investigation will have implications for the vast body of literature that investigates intuition using self-reports (Akinci and Sadler-Smith 2012; Sadler-Smith and Sparrow 2008). More broadly, future studies should consider comparing DDM-based measures with other process-tracing methods, such as mouse tracking (Travers et al. 2016), to better establish convergent validity across different approaches to measuring intuition in decision making.

4 | General Discussion

In summary, the findings from the two studies in this article contribute to the understanding of the role of intuition in decision making and how intuition relates to cognitive ability. Intuition is usually conceived as a thinking process responsible for decisions that are made without conscious use of rules and knowledge for inference (Evans 1989; Evans and Stanovich 2013). As such, the conventional view suggests that many judgment biases can be attributed to misapplication of intuition in situations that require reflective (or deliberate) reasoning (Grayot 2020; Kahneman 2002, 2011). In line with this position, the tendency to inhibit intuition is often assumed to be a manifestation of high cognitive ability that is linked to superior decision performance, providing the theoretical foundation for the CRT (Frederick 2005; Kahneman and Frederick 2002). Study 1 in this article produced evidence against this conventional view by showing that individuals measured with high cognitive abilities by the CRT rely more on intuition in risky decision making, and under a certain assumption, intuition guides high-ability individuals to achieve superior decision performance. Furthermore, these findings support the revisionist view represented by "logical intuition theory," which posits that decision makers are capable of assessing the logical principles in reasoning by fast, intuitive thinking processes, and a decision maker's intuitive ability, or the ability to form logical intuition, contributes to her overall cognitive ability for decision making (De Neys and Pennycook 2019; Erceg et al. 2022; Raoelison et al. 2020; Voudouri et al. 2024). Nevertheless, given the cross-domain instability of preference for intuition found in Study 2, it is plausible that the relationship between cognitive ability and intuition varies across decision domains. It should be noted that cognitive ability includes various aspects, such as numeracy, verbal ability, working memory capacity, etc. (Otero et al. 2022). While intuitive ability constitutes one facet of cognitive ability, it is likely to be only relevant in decision domains where intuition is the favored mode of thinking. In other words, the association between cognitive ability and intuition may depend on the context of decision making. Future investigations are necessary to validate this proposition; the results will be able to advance the knowledge regarding the relationships among intuition, cognitive ability, and behaviors.

A notable pattern that emerged from the two studies is the variability of intuition in decision making. Firstly, as it has been discussed, findings from Study 2 suggest that the preference for intuition is unstable across decision domains. Secondly, in both studies and across all model estimations, we found that decision makers became more reliant on intuition over the course of decision making. This finding contributes to the emerging body of evidence suggesting that accumulated experience increases reliance on intuitive judgment (Boissin et al. 2021; Raoelison et al. 2021). Specifically, it has long been assumed that decision makers would rely more on the fast, intuitive judgment once key principles have been practiced to automaticity (e.g., Kahneman 2011). Nevertheless, it is recognized that empirical evidence in support of this "automatization assumption" is scant as the development of intuition cannot be observed (De Neys 2023). The results from the present study provide evidence for the automatization assumption through the DDM analysis. Furthermore, in contrast to the automatization assumption, which suggests that intuitive judgment develops through feedback and learning (Boissin et al. 2021), increasing reliance on intuition was found in our experiments where no opportunities for feedback and learning were provided over the course of decision-making trials. In other words, our findings suggest that learning and feedback may not be necessary for the development of intuition. Overall, findings from the two studies presented in this article suggest that the tendency to rely on intuition is not a stable trait, but instead varies across decision domains and with experience. Additionally, prior research has shown that the DDM starting point, which is used to represent intuition in our studies, shifts in response to changes in the relative favorability of available options (Desai and Krajbich 2022; Zhao et al. 2020). Taken together, these findings underscore that the role of intuition in decision making is highly context dependent, varying across tasks and over time. Therefore, it is important to measure intuition in ways that are grounded in the specific contexts in which decisions are made. The DDM provides a valuable means to measure intuition directly from decision behavior, providing a naturally context-dependent approach to understanding how intuitive processes operate in specific decision scenarios.

In both studies, the role of intuition in decision making was evaluated by the starting point of the DDM process, which indexes a biased disposition prior to evidence accumulation. The estimated DDMs were shown to provide accurate simulation of decision outcome and decision time patterns in the real data, both patterns are thought to be influenced by intuition. As such, the validity of the DDM analysis on intuition is demonstrated by the fact that the DDM can accurately simulate key decision-making patterns based on a framework in line with how intuition affects decisions, that is by generating a fast default response at the onset of decision making. The successful application of the DDM in this article also adds to the evidence for the view that the latent role of intuition in decision making can be understood through sequential sampling models such as the DDM (e.g., Chen and Krajbich 2018; Diederich and Trueblood 2018). We note that the DDM is a descriptive model for binary choices, which therefore prevents its application to decision-making problems that involve more than two options. In such circumstances, other forms of sequential sampling models may be utilized for investigation. In particular, the racing model has been developed to model evidence accumulation among an unlimited number of options (Thomas et al. 2019; Tillman et al. 2020; Usher et al. 2002). In future research, the racing model may be applied to investigate intuition in decision-making problems with numerous options. For instance, choice overload describes the behavioral anomaly that decision performance and satisfaction of decision makers decline as the set of available options for the decision problem becomes larger (Chernev et al. 2015). For such a behavioral pattern, the binary choice framework offered by the DDM is not applicable due to the involvement of numerous options, whereas the racing model that is not constrained by the number of options can be potentially useful for investigating how intuition contributes to the observed behavioral anomaly, providing insights that would otherwise be difficult to obtain. To sum up, the DDM and other sequential sampling models offer new opportunities to study topics where progression in knowledge is hindered by the difficulty of evaluating the latent role of intuition.

5 | Conclusion

The role of intuition in decision making has attracted strong scholarly interest. Nevertheless, tools that can conclusively evaluate intuition is lacking; investigation on intuition often produce conflicting results and suffer from reliability issues. Recently, the DDM is adopted as a new tool to study intuition in decision making. The DDM is capable of simulating key behavioral patterns in the real data with high accuracy by a framework in line with how intuition affects decision, hence providing support to the validity to this approach. Study 1 in this article uses the DDM to examine how intuition relates to the CRT. Our results provide evidence against the conventional view that intuitive judgment is prone to error and that individuals with high CRT scores tend to inhibit intuition. In Study 2, the DDM analysis shows that the preference for intuition is unstable across decision domains. Previous evidence for this conclusion has exclusively come from self-reports, the validity of which to assess intuition is questionable. Our findings confirm the instability of the preference for intuition and suggest that decision makers may be more sentient about the use of intuition than that indicated by extant evidence. Overall, the two studies included in this article help to advance knowledge about the operation of intuition in decision making. Our investigation also demonstrates the usefulness and validity of the DDM analysis on intuition.

Acknowledgments

The experiment in Study 1 reported in this article was funded by Loughborough Business School at Loughborough University. We are grateful for the financial support of the school. Open access publishing facilitated by Aalto-yliopisto, as part of the Wiley - FinELib agreement.

Data Availability Statement

The data files and analysis scripts supporting the findings of this research are available on the Open Science Framework at https://osf.io/ck64y/.

Endnotes

[1] It should be emphasized that other studies have shown CRT measurements can predict diverse outcomes when numeracy is controlled (for a review, see Pennycook and Ross 2016).

[2] https://osf.io/g7cv6/

References

Ackert, L. F., R. Deaves, J. Miele, and Q. Nguyen. 2020. "Are Time Preference and Risk Preference Associated With Cognitive Intelligence and Emotional Intelligence?" *Journal of Behavioral Finance* 21, no. 2: 136–156.

Akinci, C., and E. Sadler-Smith. 2012. "Intuition in Management Research: A Historical Review." *International Journal of Management Reviews* 14, no. 1: 104–122.

Anderson, L. R., and J. M. Mellor. 2008. "Predicting Health Behaviors With an Experimental Measure of Risk Preference." *Journal of Health Economics* 27, no. 5: 1260–1274.

Baayen, R. H., D. J. Davidson, and D. M. Bates. 2008. "Mixed-Effects Modeling With Crossed Random Effects for Subjects and Items." *Journal of Memory and Language* 59, no. 4: 390–412.

Baldacchino, L., D. Ucbasaran, and L. Cabantous. 2023. "Linking Experience to Intuition and Cognitive Versatility in New Venture Ideation: A Dual-Process Perspective." *Journal of Management Studies* 60, no. 5: 1105–1146.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67, no. 1: 1–48.

Berman, J. Z., A. Barasch, E. E. Levine, and D. A. Small. 2018. "Impediments to Effective Altruism: The Role of Subjective Preferences in Charitable Giving." *Psychological Science* 29, no. 5: 834–844.

Betsch, C. 2008. "Chronic Preferences for Intuition and Deliberation in Decision Making: Lessons Learned About Intuition From an Individual Differences Approach." In *Intuition in Judgment and Decision Making*, 231–248. New York: Psychology Press.

Bogacz, R., E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen. 2006. "The Physics of Optimal Decision Making: A Formal Analysis of Models of Performance in two-Alternative Forced-Choice Tasks." *Psychological Review* 113, no. 4: 700–765.

Boissin, E., S. Caparos, M. Raoelison, and W. De Neys. 2021. "From Bias to Sound Intuiting: Boosting Correct Intuitive Reasoning." *Cognition* 211: 104645.

Boyle, P. A., L. Yu, A. S. Buchman, and D. A. Bennett. 2012. "Risk Aversion Is Associated With Decision Making Among Community-Based Older Persons." *Frontiers in Psychology* 3: 205.

Brañas-Garza, P., P. Kujal, and B. Lenkei. 2019. "Cognitive Reflection Test: Whom, How, When." *Journal of Behavioral and Experimental Economics* 82: 101455.

Brooks, P., S. Peters, and H. Zank. 2014. "Risk Behavior for Gain, Loss, and Mixed Prospects." *Theory and Decision* 77, no. 2: 153–182.

Bucher-Koenen, T., and M. Ziegelmeyer. 2014. "Once Burned, Twice shy? Financial Literacy and Wealth Losses During the Financial Crisis." *Review of Finance* 18, no. 6: 2215–2246.

Campitelli, G., and P. Gerrans. 2014. "Does the Cognitive Reflection Test Measure Cognitive Reflection? A Mathematical Modeling Approach." *Memory & Cognition* 42, no. 3: 434–447.

Charness, G., and M. Rabin. 2002. "Understanding Social Preferences With Simple Tests." *Quarterly Journal of Economics* 117, no. 3: 817–869.

Chen, D. L., M. Schonger, and C. Wickens. 2016. "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance* 9: 88–97.

Chen, F., and I. Krajbich. 2018. "Biased Sequential Sampling Underlies the Effects of Time Pressure and Delay in Social Decision Making." *Nature Communications* 9, no. 1: 1–10.

Chernev, A., U. Böckenholt, and J. Goodman. 2015. "Choice Overload: A Conceptual Review and meta-Analysis." *Journal of Consumer Psychology* 25, no. 2: 333–358.

Clay, S. N., J. A. Clithero, A. M. Harris, and C. L. Reed. 2017. "Loss Aversion Reflects Information Accumulation, Not Bias: A Drift-Diffusion Model Study." *Frontiers in Psychology* 8: 1708.

Clithero, J. A. 2018. "Response Times in Economics: Looking Through the Lens of Sequential Sampling Models." *Journal of Economic Psychology* 69: 61–86.

Cokely, E. T., M. Galesic, E. Schulz, S. Ghazal, and R. Garcia-Retamero. 2012. "Measuring Risk Literacy: The Berlin Numeracy Test." *Judgment and Decision Making* 7, no. 1: 25–47.

Cokely, E. T., and C. M. Kelley. 2009. "Cognitive Abilities and Superior Decision Making Under Risk: A Protocol Analysis and Process Model Evaluation." *Judgment and Decision Making* 4, no. 1: 20–33.

Corgnet, B., A. M. Espín, and R. Hernán-González. 2015. "The Cognitive Basis of Social Behavior: Cognitive Reflection Overrides Antisocial But Not Always Prosocial Motives." *Frontiers in Behavioral Neuroscience* 9: 287.

Dane, E., and M. G. Pratt. 2007. "Exploring Intuition and Its Role in Managerial Decision Making." *Academy of Management Review* 32, no. 1: 33–54.

De Neys, W. 2023. "Advancing Theorizing About Fast-and-Slow Thinking." *Behavioral and Brain Sciences* 46: e111.

De Neys, W., and G. Pennycook. 2019. "Logic, Fast and Slow: Advances in Dual-Process Theorizing." *Current Directions in Psychological Science* 28, no. 5: 503–509.

Desai, N., and I. Krajbich. 2022. "Decomposing Preferences Into Predispositions and Evaluations." *Journal of Experimental Psychology: General* 151, no. 8: 1883–1903.

Diederich, A., and J. S. Trueblood. 2018. "A Dynamic Dual Process Model of Risky Decision Making." *Psychological Review* 125, no. 2: 270–292.

Dorigoni, A., J. Rajsic, and N. Bonini. 2022. "Does Cognitive Reflection Predict Attentional Control in Visual Tasks?" *Acta Psychologica* 226: 103562.

Enax, L., I. Krajbich, and B. Weber. 2016. "Salient Nutrition Labels Increase the Integration of Health Attributes in Food Decision-Making." *Judgment and Decision Making* 11, no. 5: 460–471.

Epstein, S., R. Pacini, V. Denes-Raj, and H. Heier. 1996. "Individual Differences in Intuitive–Experiential and Analytical–Rational Thinking Styles." *Journal of Personality and Social Psychology* 71, no. 2: 390–405.

Erceg, N., and A. Bubić. 2017. "One Test, Five Scoring Procedures: Different Ways of Approaching the Cognitive Reflection Test." *Journal of Cognitive Psychology* 29, no. 3: 381–392.

Erceg, N., Z. Galic, A. Bubić, and D. Jelic. 2022. "Who Detects and Why: How Individual Differences in Cognitive Characteristics Underpin Different Types of Responses on Reasoning Tasks?" *Thinking & Reasoning* 29, no. 4: 594–642.

Evans, J. S. B. 1989. *Bias in Human Reasoning: Causes and Consequences*. Lawrence Erlbaum Associates, Inc.

Evans, J. S. B., and K. E. Stanovich. 2013. "Dual-Process Theories of Higher Cognition: Advancing the Debate." *Perspectives on Psychological Science* 8, no. 3: 223–241.

Fellnhofer, K. 2022. "Measuring and Enhancing Intuition." *Nature Reviews Psychology* 1, no. 9: 492.

Fellnhofer, K., and Y. Deng. 2024. "Investor Intuition Promotes Gender Equality in Access to Reward-Based Crowdfunding." *Entrepreneurship Theory and Practice* 48, no. 2: 675–718.

Figner, B., and E. U. Weber. 2011. "Who Takes Risks When and Why? Determinants of Risk Taking." *Current Directions in Psychological Science* 20, no. 4: 211–216.

Forstmann, B. U., R. Ratcliff, and E. J. Wagenmakers. 2016. "Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions." *Annual Review of Psychology* 67: 641–666.

Frederick, S. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19, no. 4: 25–42.

Frydman, C., and G. Nave. 2017. "Extrapolative Beliefs in Perceptual and Economic Decisions: Evidence of a Common Mechanism." *Management Science* 63, no. 7: 2340–2352.

Gallo, I., S. Sood, T. C. Mann, and T. Gilovich. 2017. "The Heart and the Head: On Choosing Experiences Intuitively and Possessions Deliberatively." *Journal of Behavioral Decision Making* 30, no. 3: 754–768.

Ghasemi, O., S. Handley, S. Howarth, I. R. Newman, and V. A. Thompson. 2022. "Logical Intuition Is Not Really About Logic." *Journal of Experimental Psychology: General* 151, no. 9: 2009–2028.

Grayot, J. D. 2020. "Dual Process Theories in Behavioral Economics and Neuroeconomics: A Critical Review." *Review of Philosophy and Psychology* 11, no. 1: 105–136.

Grund, T., C. Waloszek, and D. Helbing. 2013. "How Natural Selection Can Create Both Self-and Other-Regarding Preferences and Networked Minds." *Scientific Reports* 3, no. 1: 1–5.

Hogarth, R. M. 2010. "Intuition: A Challenge for Psychological Research on Decision Making." *Psychological Inquiry* 21, no. 4: 338–353.

Imai, K., L. Keele, and D. Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15, no. 4: 309–334.

Inbar, Y., J. Cone, and T. Gilovich. 2010. "People's Intuitions About Intuitive Insight and Intuitive Choice." *Journal of Personality and Social Psychology* 99, no. 2: 232–247.

Isler, O., and O. Yilmaz. 2023. "How to Activate Intuitive and Reflective Thinking in Behavior Research? A Comprehensive Examination of Experimental Techniques." *Behavior Research Methods* 55, no. 7: 3679–3698.

Jiang, R., X. Li, P. Xu, and Y. Lei. 2020. "Do Teachers Need to Inhibit Heuristic Bias in Mathematics Problem-Solving? Evidence From a Negative-Priming Study." *Current Psychology* 41, no. 10: 6954–6965.

Kahneman, D. 2002. "Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice." *Nobel Prize Lecture* 8, no. 1: 351–401.

Kahneman, D. 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., and S. Frederick. 2002. "Representativeness Revisited: Attribute Substitution in Intuitive Judgment." In *Heuristics and Biases*, edited by T. Gilovich, D. Griffin, and D. Kahneman, 49–81. Cambridge: Cambridge University Press.

Kahneman, D., and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47, no. 2: 263–292.

Krajbich, I., C. Armel, and A. Rangel. 2010. "Visual Fixations and the Computation and Comparison of Value in Simple Choice." *Nature Neuroscience* 13, no. 10: 1292–1298.

Leach, S., and M. Weick. 2018. "Can People Judge the Veracity of Their Intuitions?" *Social Psychological and Personality Science* 9, no. 1: 40–49.

Lee, M. D., A. H. Criss, B. Devezer, et al. 2019. "Robust Modeling in Cognitive Science." *Computational Brain & Behavior* 2: 141–153.

Lilleholt, L. 2019. "Cognitive Ability and Risk Aversion: A Systematic Review and Meta Analysis." *Judgment and Decision Making* 14, no. 3: 234–279.

Lufityanto, G., C. Donkin, and J. Pearson. 2016. "Measuring Intuition: Nonconscious Emotional Information Boosts Decision Accuracy and Confidence." *Psychological Science* 27, no. 5: 622–634.

Montero-Porras, E., T. Lenaerts, R. Gallotti, and J. Grujic. 2022. "Fast Deliberation Is Related to Unconditional Behaviour in Iterated Prisoners' Dilemma Experiments." *Scientific Reports* 12, no. 1: 1–10.

Moritz, B., E. Siemsen, and M. Kremer. 2014. "Judgmental Forecasting: Cognitive Reflection and Decision Speed." *Production and Operations Management* 23, no. 7: 1146–1160.

Moritz, B. B., A. V. Hill, and K. L. Donohue. 2013. "Individual Differences in the Newsvendor Problem: Behavior and Cognitive Reflection." *Journal of Operations Management* 31, no. 1–2: 72–85.

Mulder, M. J., E. J. Wagenmakers, R. Ratcliff, W. Boekel, and B. U. Forstmann. 2012. "Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff." *Journal of Neuroscience* 32, no. 7: 2335–2343.

Nisbett, R. E., and T. D. Wilson. 1977. "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84, no. 3: 231–259.

Oechssler, J., A. Roider, and P. W. Schmitz. 2009. "Cognitive Abilities and Behavioral Biases." *Journal of Economic Behavior & Organization* 72, no. 1: 147–152.

Oktar, K., and T. Lombrozo. 2022. "Deciding to Be Authentic: Intuition Is Favored Over Deliberation When Authenticity Matters." *Cognition* 223: 105021.

Otero, I., J. F. Salgado, and S. Moscoso. 2021. "Criterion Validity of Cognitive Reflection for Predicting job Performance and Training Proficiency: A meta-Analysis." *Frontiers in Psychology* 12: 668592.

Otero, I., J. F. Salgado, and S. Moscoso. 2022. "Cognitive Reflection, Cognitive Intelligence, and Cognitive Abilities: A Meta-Analysis." *Intelligence* 90: 101614.

Pachur, T., and M. Spaar. 2015. "Domain-Specific Preferences for Intuition and Deliberation in Decision Making." *Journal of Applied Research in Memory and Cognition* 4, no. 3: 303–311.

Patel, N., S. G. Baker, and L. D. Scherer. 2019. "Evaluating the Cognitive Reflection Test as a Measure of Intuition/Reflection, Numeracy, and Insight Problem Solving, and the Implications for Understanding Real-World Judgments and Beliefs." *Journal of Experimental Psychology: General* 148, no. 12: 2129–2153.

Pennycook, G., and R. M. Ross. 2016. "Commentary: Cognitive Reflection vs. Calculation in Decision Making." *Frontiers in Psychology* 7: 9.

Peters, E. 2012. "Beyond Comprehension: The Role of Numeracy in Judgments and Decisions." *Current Directions in Psychological Science* 21, no. 1: 31–35.

Phillips, W. J., J. M. Fletcher, A. D. Marks, and D. W. Hine. 2016. "Thinking Styles and Decision Making: A Meta-Analysis." *Psychological Bulletin* 142, no. 3: 260–290.

Rand, D. G., J. D. Greene, and M. A. Nowak. 2012. "Spontaneous Giving and Calculated Greed." *Nature* 489, no. 7416: 427–430.

Raoelison, M., E. Boissin, G. Borst, and W. De Neys. 2021. "From Slow to Fast Logic: The Development of Logical Intuitions." *Thinking & Reasoning* 27, no. 4: 599–622.

Raoelison, M., V. A. Thompson, and W. De Neys. 2020. "The Smart Intuitor: Cognitive Capacity Predicts Intuitive Rather Than Deliberate Thinking." *Cognition* 204: 104381.

Ratcliff, R., P. L. Smith, S. D. Brown, and G. McKoon. 2016. "Diffusion Decision Model: Current Issues and History." *Trends in Cognitive Sciences* 20, no. 4: 260–281.

Reyna, V. F. 2004. "How People Make Decisions That Involve Risk: A Dual-Processes Approach." *Current Directions in Psychological Science* 13, no. 2: 60–66.

Sadler-Smith, E., and P. R. Sparrow. 2008. "Intuition in Organizational Decision-Making." In *The Oxford Handbook of Organizational Decision-Making*, edited by G. P. Hodgkinson and W. H. Starbuck, 305–324. Oxford: Oxford University Press.

Sajid, M., and M. C. Li. 2019. "The Role of Cognitive Reflection in Decision Making: Evidence From Pakistani Managers." *Judgment and Decision Making* 14, no. 5: 591–604.

Scopelliti, I., C. K. Morewedge, E. McCormick, H. L. Min, S. Lebrecht, and K. S. Kassam. 2015. "Bias Blind Spot: Structure, Measurement, and Consequences." *Management Science* 61, no. 10: 2468–2486.

Sheng, F., A. Ramakrishnan, D. Seok, et al. 2020. "Decomposing Loss Aversion From Gaze Allocation and Pupil Dilation." *Proceedings of the National Academy of Sciences* 117, no. 21: 11356–11363.

Shinn, M., N. H. Lam, and J. D. Murray. 2020. "A Flexible Framework for Simulating and Fitting Generalized Drift-Diffusion Models." *eLife* 9: e56938.

Sinayev, A., and E. Peters. 2015. "Cognitive Reflection vs. Calculation in Decision Making." *Frontiers in Psychology* 6: 532.

Sirota, M., and M. Juanchich. 2018. "Effect of Response Format on Cognitive Reflection: Validating a Two-and Four-Option Multiple Choice Question Version of the Cognitive Reflection Test." *Behavior Research Methods* 50, no. 6: 2511–2522.

Smith, S. M., and I. Krajbich. 2018. "Attention and Choice Across Domains." *Journal of Experimental Psychology: General* 147, no. 12: 1810–1826.

Šrol, J., and W. De Neys. 2021. "Predicting Individual Differences in Conflict Detection and Bias Susceptibility During Reasoning." *Thinking & Reasoning* 27, no. 1: 38–68.

Thoma, V., E. White, A. Panigrahi, V. Strowger, and I. Anderson. 2015. "Good Thinking or gut Feeling? Cognitive Reflection and Intuition in Traders, Bankers and Financial Non-Experts." *PLoS ONE* 10, no. 4: e0123202.

Thomas, A. W., F. Molter, I. Krajbich, H. R. Heekeren, and P. N. Mohr. 2019. "Gaze Bias Differences Capture Individual Choice Behaviour." *Nature Human Behaviour* 3, no. 6: 625–635.

Tillman, G., T. Van Zandt, and G. D. Logan. 2020. "Sequential Sampling Models Without Random Between-Trial Variability: The Racing Diffusion Model of Speeded Decision Making." *Psychonomic Bulletin & Review* 27, no. 5: 911–936.

Tingley, D., T. Yamamoto, K. Hirose, L. Keele, and K. Imai. 2014. "Mediation: R Package for Causal Mediation Analysis." *Journal of Statistical Software* 59, no. 5: 1–38.

Toplak, M. E., R. F. West, and K. E. Stanovich. 2011. "The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-Biases Tasks." *Memory & Cognition* 39, no. 7: 1275–1289.

Travers, E., J. J. Rolison, and A. Feeney. 2016. "The Time Course of Conflict on the Cognitive Reflection Test." *Cognition* 150: 109–118.

Usher, M., Z. Olami, and J. L. McClelland. 2002. "Hick's law in a Stochastic Race Model With Speed–Accuracy Tradeoff." *Journal of Mathematical Psychology* 46, no. 6: 704–715.

Van Dooren, W., D. De Bock, F. Depaepe, D. Janssens, and L. Verschaffel. 2003. "The Illusion of Linearity: Expanding the Evidence Towards Probabilistic Reasoning." *Educational Studies in Mathematics* 53, no. 2: 113–138.

Voudouri, A., M. Białek, and W. De Neys. 2024. "Fast & Slow Decisions Under Risk: Intuition Rather Than Deliberation Drives Advantageous Choices." *Cognition* 250: 105837.

Welsh, M., N. Burns and P. Delfabbro (2013). "*The Cognitive Reflection Test: How Much More Than Numerical Ability?*". Paper Presented at the Proceedings of the 35th Annual Meeting of the Cognitive Science Society, Berlin, Germany. pp. 1587–1592.

Zhao, W. J., A. Coady, and S. Bhatia. 2022. "Computational Mechanisms for Context-Based Behavioral Interventions: A Large-Scale Analysis." *Proceedings of the National Academy of Sciences* 119, no. 15: e2114914119.

Zhao, W. J., A. Diederich, J. S. Trueblood, and S. Bhatia. 2019. "Automatic Biases in Intertemporal Choice." *Psychonomic Bulletin & Review* 26, no. 2: 661–668.

Zhao, W. J., L. Walasek, and S. Bhatia. 2020. "Psychological Mechanisms of Loss Aversion: A Drift-Diffusion Decomposition." *Cognitive Psychology* 123: 101331.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.