Strategy Switching in Sequential Decision Tasks: Evidence from the Iowa Gambling Task

Authors: Hu Xinyun, Shen Yue, Junyi Dai, DAI Junyi

Date: 2023-05-20T00:00:00+00:00

Abstract

Numerous studies have utilized sequential decision-making tasks to investigate decision strategies across various domains. These studies, which assume individuals employ a single strategy throughout all task trials and compare the fit of corresponding computational cognitive models to empirical data, have revealed that multiple decision strategies may be involved in different decision-making tasks. However, a common limitation of such research is the neglect of potential strategy switching during task performance. Study 1 addressed this issue by developing computational cognitive models for the Iowa Gambling Task that permit switching between reinforcement learning and heuristic strategies, and by comparing these models with single-strategy alternatives, thereby providing clear evidence that individuals do change decision strategies in such sequential tasks. Study 2 further demonstrated that the probability of strategy switching increases with the number of trials. These findings indicate that to properly characterize decision strategies across various tasks, it is essential to consider the possibility of strategy switching during sequential decision-making, particularly in tasks with many trials. Future research should examine the various forms that strategy switching may take, as well as the task-related and individual factors that precipitate such switches, to further advance our understanding of the psychological mechanisms underlying sequential decision-making tasks.

Full Text

Strategy Switching in Sequential Decision Tasks: Evidence from the Iowa Gambling Task

Xinyun Hu, Yue Shen, Junyi Dai*
junyidai@zju.edu.cn
(Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310058, China)

Abstract:

Much research has been devoted to studying decision strategies in various tasks. Such research usually involved a sequence of decision trials under the same task structure to provide sufficient information for inferring the underlying decision strategies. By assuming each individual adopted a single decision strategy across all decision trials and comparing corresponding computational cognitive models in terms of their performances in fitting empirical data, such studies have revealed multiple possible decision strategies for many major decision tasks. One common drawback of such research, however, was overlooking the possibility that individuals switched their strategies along the sequence of decisions. This might lead to inappropriate conclusions regarding the decision strategies underlying specific decision tasks or misleading inferences of potential cognitive and affective differences between normal and different clinical populations based on parameter estimates from models assuming single strategies.

To address this critical issue, two studies were conducted to examine the possibility of strategy switching in the Iowa Gambling Task (IGT), an experience-based decision task with a sequence of trials aimed at mimicking real-world decisions under uncertainty. By developing a computational cognitive model that allowed for switches between reinforcement learning strategies and heuristic strategies and comparing its performance with those of single-strategy models, Study 1 showed that data from about half of the 617 healthy participants in 10 previous studies were better fitted by the strategy-switching model than three single-strategy models that performed well in previous research, that is, the WSLS, PVL2, and VPP models as exemplar models assuming heuristic, reinforcement learning, and mixed strategies, respectively. This result provided clear support for the possibility of strategy switching in the IGT.

Since strategy switching might occur with accumulating experience or fatigue and an increasing number of trials is likely to facilitate such changes, 321 participants were recruited in Study 2 to further examine whether a larger number of trials would contribute to more strategy switching in the IGT. Specifically, 160 participants performed a 100-trial IGT, whereas the other 161 participants performed a 200-trial IGT under otherwise the same task structure. It was found that data from a larger proportion of individual participants were best fitted by the strategy-switching model when the IGT involved 200 trials rather than standard 100 trials. This result provided further evidence for strategy switching in the task.

Overall, the current results suggest that strategy switching is likely to occur in a sequence of decisions under the same task structure. Consequently, in order to obtain proper understanding of the decision strategies for various decision tasks, it is necessary to consider seriously the possibility of strategy switching, especially for a long sequence of decisions. For a more refined understanding of psychological mechanisms underlying sequences of decisions, future research might further investigate various forms of strategy switching such as gradual instead of abrupt switches and task and individual factors that trigger such

switches.

Keywords: Decision task with a sequence of trials, The Iowa Gambling Task, Strategy switching, Computational cognitive modeling, Reinforcement learning and heuristic strategies

---

As the ancient saying goes, "The wise adapt to the times, the knowledgeable adjust to circumstances." When repeatedly facing decisions with identical task structures (i.e., completing sequential decision tasks), the strategies people employ are not static.[1] Extensive research has demonstrated that various decision tasks involve multiple distinct decision strategies. For instance, multi-attribute decision tasks involve a range of compensatory strategies (where advantages and disadvantages across attributes can offset each other) and non-compensatory strategies (where such trade-offs are not permitted; e.g., Payne et al., 1988; Rieskamp & Otto, 2006; Walsh & Gluck, 2016). In risky decision tasks, individuals may adopt strategies based on expected utility or similar evaluations (e.g., Kahneman & Tversky, 1979; Von Neumann & Morgenstern, 1944) or simpler heuristic strategies (e.g., Brandstätter et al., 2006). Furthermore, researchers have investigated how factors such as information environment, task demands, and individual differences influence strategy selection (e.g., Bergert & Nosofsky, 2007; Pachur & Galesic, 2013), finding that changes in task environment or requirements can lead to corresponding shifts in decision strategies (e.g., Bröder & Schiffer, 2006; Lee et al., 2014).

Beyond strategy switching driven by changes in task environment and demands, might people also switch strategies due to self-regulation, adaptation, or intrinsic exploratory motivation even when the task environment and requirements remain relatively stable? In most empirical studies on decision strategies, participants must complete a series of decision trials under identical task structures to provide researchers with sufficient information for inferring their decision strategies. Although previous research has examined the diversity of strategies individuals employ when facing specific decision tasks and the factors influencing strategy selection, few studies have investigated the possibility of strategy switching during a relatively stable sequential decision task. If such switching does occur, previous research on decision strategies may have drawn erroneous conclusions by overlooking this possibility. To better understand individuals' decision strategies across various tasks, a fundamental question must first be addressed: Does strategy switching indeed occur in sequential decision-making under relatively stable task environments and demands? This paper examines this important theoretical and practical issue using the Iowa Gambling Task (IGT) as a typical sequential decision task.

The Iowa Gambling Task (IGT) is an experience-based decision-making simulation originally developed to examine decision-making deficits in patients with ventromedial prefrontal cortex damage when facing uncertain real-world situations (Bechara et al., 1994). The task comprises four card decks (labeled A, B, C, and D), from which participants must repeatedly choose. After each deck selection, the top card is drawn and turned over, yielding a monetary reward based on the card information. However, selecting a deck sometimes also incurs losses. Participants are unaware of each deck's payoff structure and the total number of trials before starting, with their goal being to maximize total earnings through their choices. Consequently, participants must learn each deck's gain/loss patterns through repeated selections and adopt specific strategies to complete the task. The IGT has been widely used to identify decision-making deficits across various clinical populations, including individuals with brain injuries (Hochman et al., 2010), substance abuse disorders (Ahn et al., 2014; Bechara & Damasio, 2002; Bechara et al., 2001), neurological conditions (Stout et al., 2001), and mental health disorders (Li et al., 2019; Xu, 2012).

In addition to examining decision deficits in clinical populations, the IGT has been used to investigate decision strategies in both normal and clinical populations when facing uncertain situations. To this end, researchers have proposed a series of computational cognitive models corresponding to different strategies, which can be broadly categorized into reinforcement learning models and heuristic models. Reinforcement learning models assume the IGT involves three processes: a motivational outcome evaluation process, a cognitive expected valence updating process, and a probabilistic response selection process. Busemeyer and Stout (2002) proposed the first reinforcement learning model for the IGT—the Expectancy-Valence Learning (EVL) model. This model assumes individuals use an Expectancy Utility (EU) function to evaluate the utility of each outcome (Ahn et al., 2008), employs a Delta-Learning (DEL) rule to update each deck's expected valence (Rescorla & Wagner, 1972), and uses a Trial-Dependent Choice (TDC) rule to guide selection in the next trial (Luce, 1959). Building upon the EVL model, Ahn et al. (2008) further explored different mathematical formulations for each of the three processes involved in reinforcement learning models and proposed the Prospect-Valence Learning (PVL) model. This model assumes individuals use a Prospect Utility (PU) function (Kahneman & Tversky, 1979) to evaluate the net outcome of a choice (i.e., the sum of rewards and potential simultaneous losses), updates expected valence using the Decay-Reinforcement Learning (DRL) rule proposed by Erev and Roth (1998), and makes responses using a Trial-Independent Choice (TIC) rule (Yechiam & Ert, 2007). More recent research employing systematic model comparison methods has shown (Dai et al., 2015) that when evaluating outcomes, individuals are more likely to first separately evaluate simultaneously occurring rewards and losses using the prospect utility function and then integrate the evaluated results. The corresponding model is called the Prospect-Valence Learning 2 (PVL2) model.

Among heuristic models of the IGT, the most representative and best-performing in fitting empirical data is the Win-Stay-Lose-Shift (WSLS) model (Worthy et al., 2012). This model assumes that each choice depends only on the deck chosen in the previous trial and its outcome, independent of earlier choices and their results. Therefore, compared to reinforcement learning models that consider choices and corresponding outcomes from all previous trials, the WSLS model assumes a simpler psychological mechanism. Specifically, the model assumes that the probability of continuing to select the same deck depends on the outcome of the current deck selection. If the net outcome of the current selection is non-negative (i.e., a win), there is a higher probability of continuing to select the same deck; conversely (i.e., a loss), there is a higher probability of switching to a different deck in the next trial.

Despite rich research findings on decision strategies in the IGT, few studies have considered the possibility of individuals switching strategies during task completion. Busemeyer and Stout (2002) proposed a Strategy-Switching Heuristic Choice model. However, the so-called “strategy switching” in this model does not refer to a fundamental change in decision strategy but rather to the process of reallocating selection probabilities between disadvantageous decks (i.e., Decks A or B) and advantageous decks (i.e., Decks C or D) as individuals experience increasing losses from selecting disadvantageous decks. Additionally, some researchers have proposed computational cognitive models that combine reinforcement learning and heuristic strategies. For example, Worthy et al. (2013) proposed the Valence-Plus-Perseverance (VPP) model. This model suggests that in each IGT trial, people comprehensively consider each deck’s expected valence along with their previous trial’s choice and its outcome before deciding on the current trial’s selection. Although this model incorporates both reinforcement learning and heuristic components and demonstrates better performance in fitting empirical data compared to EVL, PVL, and WSLS models, it still assumes individuals use a single, albeit more complex, hybrid strategy for every trial in the IGT.

In summary, research on decision strategies in the IGT has not yet examined the possibility of strategy switching during task completion. If individuals do switch decision strategies during the task for various reasons, previous studies that only compared single-strategy models may have drawn incorrect conclusions about individuals’ strategy choices. Moreover, research that infers the psychological mechanisms underlying decision-making differences between populations based on parameter estimates from single-strategy models (e.g., Ahn et al., 2014; Yechiam et al., 2005) may produce biased estimates, leading to misinterpretation of population differences. The present study addresses whether strategy switching occurs in the IGT by developing models that allow strategy switching and comparing them with traditional single-strategy models, aiming to provide a more reliable basis for conclusions about decision strategies in the IGT and population differences, and to contribute to the broader theoretical and practical discussion of decision strategy switching.

2.1 Methods

2.1(1) IGT Introduction

As described above, the IGT includes four decks (labeled A, B, C, and D). In each trial, participants select a deck and receive a reward based on the information presented on the top card, potentially incurring simultaneous losses. Participants aim to maximize total earnings without knowing the total number of trials. For example, Bechara et al.'s (1994) original IGT study included 100 trials (unknown to participants) and employed the payoff scheme shown in Table 1 . Specifically, participants received \$100 for each selection of Deck A or B. However, for every 10 selections of Deck A, participants incurred five losses ranging from \$150 to \$350 in increasing amounts, with the specific positions of these five losses varying across each block of 10 selections. Similarly, for every 10 selections of Deck B, participants incurred one loss of \$1,250, with its position varying across each block of 10 selections. For Decks C or D, each selection yielded \$50. However, for every 10 selections of Deck C, participants incurred five losses totaling \$250, while for every 10 selections of Deck D, they incurred one loss of \$250, with loss positions varying across each block of 10 selections for both decks. Subsequent studies have used identical or similar task setups, with primary modifications involving the number of trials and whether real monetary payments were used. When using real payments (i.e., paying participants based on their final total earnings), the original payoff amounts from Bechara et al. are typically reduced by a factor of 100 to control experimental costs (e.g., Dai et al., 2015). Regardless of the specific payoff scheme, all IGT studies share three key characteristics: (1) Decks A and B yield high immediate rewards but larger total losses, making them disadvantageous (i.e., negative total returns) in the long run; (2) Decks C and D yield lower immediate rewards but smaller total losses, making them advantageous (i.e., positive total returns) in the long run; and (3) Decks A and C incur losses more frequently than Decks B and D.

2.1(2) Single-Strategy Models

To provide appropriate baseline models for investigating strategy switching in the IGT, this study considers three major categories of single-strategy models from existing literature: reinforcement learning models, heuristic models, and hybrid models, using the PVL2, WSLS, and VPP models as representatives of each category. These models have demonstrated good performance in previous research; therefore, if the new strategy-switching model shows better performance, it would support the existence of strategy switching in the IGT. The following sections introduce the specific mathematical formulations of these three computational cognitive models.

Reinforcement learning models for the IGT assume that people complete the task through three processes: outcome evaluation, expected (or prospective) valence updating, and probabilistic selection. According to the PVL2 model (Dai et al., 2015), after selecting a deck, individuals separately evaluate rewards and potential losses using prospect theory's value function and then combine the results. The corresponding utility function is called the Prospect Utility 2 (PU2) function, with the specific form:

$$(\ ) = [\quad (\ )]! - [| \quad (\ )|]!$$

where ( ) and ( ) represent the reward and potential simultaneous loss amounts obtained in trial , respectively, and ( ) represents the aggregated utility evaluation for trial . The parameter is a shape parameter measuring the sensitivity of subjective utility to objective value, with a range between 0 and 1, while represents the loss aversion parameter from prospect theory, with a range between 0 and 5.

After completing outcome evaluation, the PVL2 model assumes that individuals update each deck's prospective valence using the decay-reinforcement learning rule:

“( ) = ”(− 1) + ”( ) ( )

where “( ) represents the expected valence of deck after trial (= 1, 2, 3, 4, corresponding to Decks A, B, C, and D), is the memory decay parameter ranging from 0 to 1 (with larger indicating smaller memory decay effects on expected valence), and ”( ) is a dummy variable equal to 1 if deck was selected in trial and 0 otherwise. In other words, the expected valence of the selected deck is updated through both memory decay and current utility evaluation, while the expected valence of unselected decks undergoes only memory decay.

Finally, the PVL2 model assumes that individuals determine the probability of selecting each deck in the next trial based on their expected valences and make random selections accordingly (Sutton & Barto, 1998):

Pr[(+ 1) =] = #(%) (!(%) #(%) ("(%)

where the left side Pr[(+ 1) =] denotes the probability of selecting deck in trial + 1, and the denominator on the right side is a normalization factor ensuring that the predicted selection probabilities for all decks sum to 1. The parameter ( ) is the sensitivity parameter of the choice function; larger ( ) indicates that choices depend more heavily on decks' expected valences. According to the PVL2 model, ( ) remains constant across trials (i.e., ( ) is a constant function of), with the form:

( ) = = 3- − 1

where is a free parameter ranging from 0 to 5. Larger values produce larger values, representing a higher likelihood of selecting decks with higher expected valence. Overall, the PVL2 model comprises three processes—utility evaluation, expected valence updating, and probabilistic selection—with four free parameters: , , , and .

As a representative heuristic model, the WSLS model assumes a substantially simpler decision strategy than the PVL2 model. According to this model, individuals probabilistically determine their next choice based only on the deck selected in the previous trial and its net outcome (i.e., the sum of reward and loss). The model has two parameters: the first represents the probability of continuing to select the same deck when the previous deck's net outcome was greater than or equal to zero:

Pr(|) = PrF “( ) G %., = ” & (− 1) 0M

where %., = ” indicates that deck was selected in trial − 1, (− 1) 0 indicates that this selection yielded a non-negative net outcome, and ”( ) denotes continuing to select deck in trial . The model’s second parameter represents the probability of switching to a different deck when the previous deck’s net outcome was negative:

Pr(|) = 1 − PrF “( ) G %., = ” & (− 1) < 0M

where the symbols have the same meaning as in Equation 5. The model further assumes that when selecting a different deck from the previous trial, all possible decks have equal selection probabilities. Therefore, to ensure that total selection probabilities for all decks sum to 1, the probability of selecting any alternative deck in trial is ,./0(1%23 | 678) when the net outcome in trial − 1 was non-negative, and the probability of continuing to select the same deck in trial is 1 − Pr(|) when the net outcome in trial − 1 was negative, with the probability of selecting any alternative deck being /0(1:7;% | <=11).

In addition to reinforcement learning and heuristic models, the hybrid-strategy VPP model proposed by Worthy et al. (2013) has also demonstrated good performance. Worthy et al. argued that reinforcement learning models using decay-reinforcement rules confound the tendency to perseverate on the same deck with the tendency to select the deck with the highest expected valence. Therefore, they separated these two tendencies and proposed the VPP model. According to this model, individuals first use the PU function to evaluate the utility of an outcome and update each deck’s expected valence using the delta-learning rule:

( ) = P ( )! 当 ( ) 0 − | ( )|! 当 ( ) < 0

“( ) = ”(− 1) + ”( ) [( ) − ”(− 1)]

where ( ) represents the net gain of the current trial’s outcome, and other symbols have the same meaning as above.

On the other hand, individuals also determine their current tendency to perseverate on deck based on whether deck was selected in the previous trial and whether the net gain from selecting deck was non-negative:

7( ) = S * 7(− 1) + >=1 当 ( ) 0 * 7(− 1) + 8?@ 当 ( ) < 0

where “( ) represents the tendency to perseverate on deck in trial , is a decay parameter ranging from 0 to 1 (similar to parameter in Equation 2), and >=1 and 8?@ are two free parameters representing the change in perseveration tendency determined by the net outcome after a deck is selected, both ranging from –1 to 1. Finally, each deck’s overall value (”( ), i.e., comprehensive evaluation) is a weighted average of its expected valence and perseveration tendency:

“( ) = (! ”( ) + X1 − (!Y ”( )

where (! is a weight parameter ranging from 0 to 1.

Finally, similar to the PVL2 model, the VPP model assumes that participants determine the probability of selecting each deck in the next trial based on deck values and make random selections accordingly:

Pr[(+ 1) =] = #(%) A!(%) #(%) A”(%)

where the sensitivity parameter ( ) is calculated using a trial-independent rule (i.e., Equation 4). The VPP model includes eight free parameters: , , and for utility evaluation and updating; , >=1, and 8?@ for perseveration; (! for comprehensive evaluation; and for choice response.

2.1(3) Strategy-Switching Model

Since the IGT typically involves 100 or more trials, individuals may switch strategies during the task for various reasons. In this study, we assume two possible types of switching: (1) starting with a heuristic strategy that requires less information when little is available, then switching to a more complex and sophisticated reinforcement learning strategy after gaining more knowledge about the decks; or (2) starting with a reinforcement learning strategy, then switching to a heuristic strategy due to fatigue, boredom, or the need to reduce cognitive load as the task progresses. From a modeling perspective, given the dominant positions of the PVL2 model among reinforcement learning models and the WSLS model among heuristic models, this study uses these two models to represent potential reinforcement learning and heuristic strategies, respectively, to explore the possibility of strategy switching in the IGT.

Specifically, we developed a Switching-Strategy-Once (SSO) model that assumes individuals make one strategy switch between heuristic and reinforcement learning strategies during the IGT, with the specific computational cognitive mechanisms used during heuristic or reinforcement learning phases being identical to those assumed by the WSLS or PVL2 models, respectively. In addition to the parameters involved in the WSLS and PVL2 models, the SSO model includes two new parameters representing the trial at which the strategy switch occurs, denoted as (i.e., Switching Point), and the type of strategy switch, denoted as (i.e., Switching Type). = 1 represents switching from reinforcement learning to heuristic strategy during the IGT, while = 2 represents the opposite switching direction. Therefore, the model has eight parameters: , , , and for the reinforcement learning strategy; Pr(|) and Pr(|) for the heuristic strategy; the switching point parameter ; and the switching type parameter . Since strategy-switching models may become too similar to corresponding single-strategy models when the switching point occurs at the beginning or end of the task, making them difficult to distinguish, this study restricts to range from trial 21 to the 21st-to-last trial.

2.1(4) Data

To systematically compare the ability of strategy-switching and single-strategy models to fit empirical data, we selected a series of representative datasets from previous IGT studies as model-fitting targets (Steingroever et al., 2015). Specifically, these data came from 10 studies comprising 617 healthy participants across different age ranges, with IGT trial counts of 95, 100, and 150. All IGT tasks were completed on computers, with payoff schemes identical or similar to those used by Bechara et al. (1994) shown in Table 1. Basic information about the individual studies can be found in Table 1 of Steingroever et al.

2.1(5) Model Fitting and Comparison Methods

Each computational cognitive model examined in this study (i.e., WSLS, PVL2, VPP, and SSO) can predict the probability of selecting each deck in the next trial based on participants' previous choices and outcomes (i.e., one-step-ahead prediction; Ahn et al., 2008). Therefore, we first used Maximum-Likelihood Estimation (MLE) to fit each model to individual participants' choice data, finding for each model the parameter value combination that maximized the likelihood of the observed data, using the predicted probability of the observed data as a preliminary indicator of model fit. Specifically, the likelihood under particular model parameter values was defined as the probability of the model predicting the individual's observed choice sequence, and the Log-Likelihood (LL) was defined as:

=]] ln XPr X “(+ 1)YY × ”(+ 1)

where represents the total number of trials, Pr X “(+ 1)Y denotes the model-predicted probability of selecting deck in trial + 1 based on the participant's previous choices and their outcomes, and ”(+ 1) is a dummy variable equal to 1 if deck was selected in trial + 1 and 0 otherwise. This means that only the predicted probability of the actually selected deck in each trial contributes to the log-likelihood calculation. We then used the Particle Swarm Optimization (PSO) algorithm in MATLAB (Clerc, 2010) to find the maximum log-likelihood for each model and obtain the corresponding parameter estimates.

Generally, more complex models achieve better fit. Since the models examined have different numbers of parameters, their complexity levels differ. Therefore, we used two commonly employed criteria suitable for maximum-likelihood estimation that consider both fit and complexity: the Akaike Information Criterion with second-order bias correction (AICC; Akaike, 1974; Sugiura, 1978) and the Bayesian Information Criterion (BIC; Schwarz, 1978). We used these criteria scores to evaluate and select among models, with calculations as follows:

B = −2 C + 2 + 2 (+ 1) − − 1

= −2 C + ×ln( )

where represents the number of free parameters in the model, is the number of data points to be fitted (i.e., total trials minus 1), and C is the model's maximum log-likelihood. Smaller AICC (or BIC) values indicate better model performance (Broomell et al., 2011).^{2}

^{2}When the ratio of sample size to number of parameters is small (i.e., sample size/number of parameters < 40), using AICC can compensate for potential overfitting that may occur with AIC (Burnham & Anderson, 2004). Therefore, this paper uses AICC rather than AIC as a model evaluation criterion.

2.1(6) Model Recovery Test

Since AICC and BIC differ in the degree to which they penalize model complexity, and BIC generally imposes a higher penalty (i.e., $*$ ln( )) than AICC (i.e., 2 + D(Œ,) 8.*., using these two criteria may lead to different model selection results. Therefore, we conducted model recovery tests to determine which criterion is more appropriate for model selection with the observed data (Wagenmakers et al., 2004; Worthy et al., 2012). Specifically, this test involves two main steps: First, for each model, we used the optimal parameter values obtained from fitting the observed data to generate simulated participant data. During simulation, participants' actual choices and outcomes were not used; instead, data were randomly generated based on model-predicted selection probabilities and the IGT's inherent structure. Second, we fitted the simulated data from each model with different models and compared model performance using specific selection criteria. If a particular model selection criterion yields relatively good performance for each model only with data generated by itself, this indicates good model discriminability under that criterion. Conversely, if some models show relatively good performance even with data generated by other models under a particular criterion, this indicates poor model discriminability. In other words, such a criterion cannot accurately identify the true data-generating model and is therefore unsuitable for model selection based on observed data.

In this study, we fitted models to the observed data of 617 participants to obtain optimal parameter values for each participant under each model. Then, for each model, we generated three sets of simulated data using each participant's optimal parameter values, producing 1,821 (= 617 ×3) sets of simulated participant data. We then fitted these simulated data with the WSLS, PVL2, VPP, and SSO models using the same method as for the observed data. Finally, by analyzing model discriminability under different criteria (i.e., AICC and BIC), we could select the more appropriate model selection criterion for the observed data.

2.2 Results

2.2(1) Model Fitting and Comparison

Table 2 presents the results of fitting each model to the observed data of all 617 participants. When using AICC as the model selection criterion, the SSO model performed best both at the group level (mean) and at the individual level, with VPP, PVL2, and WSLS models performing progressively worse. When using BIC as the criterion, the PVL2 model showed the best performance at the group level (mean), followed by the SSO model. At the individual level, the WSLS and PVL2 models performed relatively well, showing the best performance for 30.79% and 33.87% of participants, respectively, while VPP and SSO models performed similarly. Regardless of whether AICC or BIC was used, the SSO model showed the best performance for a subset of participants (AICC: 43.27%, BIC: 18.96%).

2.2(2) Model Recovery Test

Since AICC and BIC penalize model complexity differently, with AICC tending to select models with more parameters than BIC, it is not surprising that the more complex VPP and SSO models performed better under AICC. To select the more appropriate model selection criterion, we conducted model recovery tests. Tables 3 and 4 present the model recovery test results. When using AICC as the model selection criterion, models showed good discriminability. For simulated data generated by each model, that model itself showed the best performance for the largest proportion of individual simulated datasets. When using BIC as the criterion, for simulated data generated by each model, the simplest WSLS model showed the best performance for the largest proportion of individual simulated datasets, indicating that BIC could not adequately discriminate between WSLS and other models. Therefore, AICC is more appropriate than BIC for model selection in this study.

2.3 Discussion

This study proposed a one-time strategy-switching model for the IGT and compared its data-fitting performance with that of representative single-strategy models—the PVL2 model (reinforcement learning strategy), WSLS model (heuristic strategy), and VPP model (hybrid strategy)—using data from 617 healthy participants from previous studies. The relative performance of models differed when using AICC versus BIC as selection criteria, but the strategy-switching model showed the best performance for a certain proportion of individual datasets. Model recovery test results indicated that AICC is more suitable for model selection in the current study than BIC because it is more likely to recover the correct data-generating model. When using AICC as the selection criterion, the SSO model outperformed the other three models at both group and individual levels, and the strategy-switching model showed the best performance for nearly half (43.27%) of participants' observed data. These results suggest that individuals are highly likely to switch decision strategies during the IGT.

As mentioned earlier, factors such as accumulating experience or fatigue may cause strategy switching in sequential decision tasks like the IGT. As the number of trials increases, it is reasonable to assume that such factors are more likely to take effect, making individuals more likely to switch decision strategies during the task. Therefore, as a supplement to the main study, we also compared model performance across IGT studies with different numbers of trials to further examine the possibility of strategy switching. Among the 617 participants examined in this study, 15 completed a 95-trial IGT, 504 completed a 100-trial IGT, and 98 completed a 150-trial IGT. Table 5 presents the corresponding results using AICC as the model selection criterion for IGT data with different trial counts. The advantage of the strategy-switching model over other models increased with trial count, both in terms of mean AICC and the proportion of participants for whom each model performed best. This was particularly evident in the proportion of individual participants for whom each model performed best, increasing from 13.33% to 53.06%.

To more deeply understand the specific circumstances of strategy switching, we further analyzed the distribution of the switching point (i.e.,) parameter under different trial counts. Specifically, we calculated the distribution information for the parameter corresponding to individual datasets best fitted by the SSO model for various IGT tasks. For the 95-trial IGT, the mean estimate was 48.5 (SD = 37.48); for the 100-trial IGT, the mean was 48.92 (SD = 19.47); and for the 150-trial IGT, the mean was 81.42 (SD = 36.03). Notably, as IGT trial count increased, the average position of strategy switching also moved later. For each switching type (i.e., switching from reinforcement learning to heuristic strategy, = 1, or from heuristic to reinforcement learning strategy, = 2), we further used one-sided Mann-Whitney tests to analyze differences in mean switching points between the 100-trial and 150-trial conditions (only 2 participants who completed the 95-trial IGT had data best fitted by the SSO model, so this condition was not analyzed). Results showed that for both switching types, the mean switching point was significantly later in the 150-trial IGT than in the 100-trial IGT (both ps < 0.001).^{3} This pattern may occur because as IGT trial count increases, a higher proportion of participants switch strategies during the task, and the switching points of newly added switchers tend to be later. This provides further evidence that individuals may switch strategies during the IGT and that the likelihood of strategy switching increases with trial count.

It should be noted that although the above analyses support the possibility of strategy switching in the IGT, these analyses examined data from different studies with varying task setup details, and the range and spacing of trial counts were not optimal, with uneven numbers of participants completing different trial counts. Therefore, these results can only be considered as providing limited evidence for strategy switching in the IGT. In Study 2 reported below, we manipulated trial count more systematically, collecting data from nearly equal numbers of participants under each trial count using identical task setups, to better test the key hypothesis that increased trial count promotes strategy switching.

^{3}We also used one-sided Mann-Whitney tests to examine differences in switching points between switching types, finding significant differences only in the 100-trial condition. In Study 2, mean switching points did not differ significantly between switching types for either the 100-trial or 200-trial IGT.

3.1 Methods

3.1(1) Participants

This study manipulated IGT trial count experimentally, with two conditions: 100 trials and 200 trials. A total of 321 college students were recruited (134 males, 187 females), with a mean age of 20.54 years (SD = 2.41). Among them, 160 completed the 100-trial IGT, while 161 completed the 200-trial IGT. Participants were required to be non-psychology majors who had not previously participated in IGT research. All participants provided informed consent before the experiment and participated voluntarily. After the experiment, participants received base compensation plus additional rewards contingent on their IGT performance, with higher performance yielding greater additional rewards.

3.1(2) Experimental Design and Procedure

This experiment used a single-factor between-subjects design to examine and compare the likelihood of strategy switching in the IGT under different trial counts. The experiment included two conditions: 100 trials (the standard setting in most IGT research) and 200 trials (which effectively increases the distance from the standard condition while controlling total experimental duration). Before the task, participants read standardized IGT instructions and were informed they had 2,000 units of experimental currency as initial wealth. During the task, participants saw four decks positioned at the top, bottom, left, and right of the screen and could select corresponding decks using the ‘up,’ ‘down,’ ‘left,’ and ‘right’ arrow keys. Participants were unaware of the total number of trials before completing the task. After each selection, the screen center displayed the current trial’s reward and loss, along with the updated total wealth (see Figure 1 [Figure 1: see original paper]). The top-bottom-left-right deck arrangement was used to reduce non-random influences on deck selection associated with traditional left-to-right arrangements, such as sequentially selecting Decks A, B, C, and D in initial trials and subsequently selecting spatially adjacent decks. Additionally, this study used the same payoff scheme shown in Table 1, with loss positions randomized across each block of 10 selections for each deck.

3.1(3) Data Analysis

This study used the same model fitting and comparison techniques as Study 1 to analyze and compare the performance of three single-strategy models and the one-time strategy-switching model in fitting individual IGT data, and conducted model recovery tests. Additionally, we used independent samples Z-tests for proportions to analyze the effect of trial count on the likelihood of strategy switching in the IGT.

3.2 Results

3.2(1) Model Analysis and Comparison

Since model recovery tests indicated that AICC was still more likely than BIC to yield correct model selection in this study (see below), only AICC-based results are reported here. Table 6 presents model comparison results for the 100-trial and 200-trial groups using AICC. The SSO model performed best in both conditions, both at the group level (mean) and at the individual level. Moreover, for both the 100-trial and 200-trial IGT, the SSO model showed the best performance for at least half of the participants. As in Study 1, VPP, PVL2, and WSLS models performed progressively worse. Independent samples Z-tests for proportions indicated that the likelihood of strategy switching (i.e., the proportion of participants for whom the SSO model performed best in fitting observed data) was higher in the 200-trial condition (65.22%) than in the 100-trial condition (50.00%, $z = 2.76$, one-tailed $p = 0.003$, 95% CI for the difference in proportions = [0.045, 0.259], Cohen's $h = 0.31$, corresponding to a small effect size).

As in Study 1, we also analyzed the estimated parameter values for participants whose data were best fitted by the SSO model under both trial count conditions. For the 100-trial IGT, the mean estimate was 47.03 ($SD = 20.39$); for the 200-trial IGT, the mean was 95.38 ($SD = 54.21$).^{4} One-sided Mann-Whitney tests indicated that for both switching types, the mean switching point was significantly later in the 200-trial condition than in the 100-trial condition (both $ps < 0.001$).

^{4}In both this study and Study 1, the mean estimate for participants best fitted by the SSO model was close to the midpoint of the allowed range. A possible reason is that for individuals who switch strategies, the likelihood of the switching point occurring at any position within the model's allowed range is roughly equal, with the overall distribution showing a unimodal shape.

3.2(2) Model Recovery Test

This study generated $3 \times 321 = 963$ sets of individual participant data using each model and fitted each simulated dataset with the four models. Table 7 presents the model recovery test results based on AICC for the 100-trial and 200-trial groups. In both conditions, each examined model showed the best performance for the largest proportion of simulated data generated by itself. Overall, the proportion of correctly recovered data-generating models was higher in the 200-trial condition (71.74%) than in the 100-trial condition (64.69%, $z = 4.70$, one-tailed $p < 0.001$, 95% CI for the difference in proportions = [0.041, 0.100], Cohen's $h = 0.15$, corresponding to a small effect size).

Table 8 presents the model recovery test results based on BIC. As in Study 1, when using BIC for model selection, the WSLS model showed the best performance in almost all cases, regardless of which model generated the individual simulated data, indicating that BIC could not adequately discriminate between WSLS and other models. Only in the 200-trial condition could the PVL2 and SSO models show the best performance for data generated by themselves. Overall, the proportion of correctly recovered data-generating models was higher in the 200-trial condition (59.06%) than in the 100-trial condition (49.17%, $z = 6.16$, one-tailed $p < 0.001$, 95% CI for the difference in proportions = [0.068, 0.130], Cohen's $h = 0.20$, corresponding to a small effect size).

3.3 Discussion

The purpose of this study was to examine whether increasing trial count leads to a higher likelihood of strategy switching in the IGT. Results showed that, as in Study 1, the strategy-switching model demonstrated the best performance for at least half of the participants in both the standard 100-trial and the extended 200-trial IGT. More importantly, compared to the 100-trial IGT, the strategy-switching model showed the best performance for a higher proportion of participants when the IGT included 200 trials. This suggests that people are more likely to switch strategies in the IGT with 200 trials. This result rules out the alternative explanation that the strategy-switching model's superior performance for some participants was merely due to random variation in model comparison results, providing further support for the possibility that individuals may switch strategies in sequential decision tasks like the IGT. Additionally, model recovery test results indicated that AICC remains a more appropriate criterion for correct model selection than BIC. Therefore, this study continues to use AICC as the basis for model selection and strategy inference. Finally, model recovery performance under both AICC and BIC was better in the 200-trial condition than in the 100-trial condition, consistent with the traditional view that larger data volumes help better discriminate between different models.

4 General Discussion

Sequential decision tasks are ubiquitous in daily life and extensively used in empirical research on decision strategies and influencing factors. For example, human resources personnel must frequently make choices among job candidates when hiring for various positions, and laboratory tasks like the IGT that require participants to make repeated decisions under identical task structures are also common. Previous research on decision strategies in sequential decision tasks typically assumes individuals use the same strategy across all trials, with multiple trials required solely to provide sufficient information for strategy inference. However, in such decision tasks, people not only learn about the specific characteristics of task stimuli but may also learn and adjust their decision strategies at a higher level. A fuller understanding of this latter type of learning will help us draw more accurate inferences about strategy selection and examine factors influencing strategy choice and switching, thereby better serving the goal of improving decision-making.

This study systematically examined the possibility of strategy switching in sequential decision tasks using the IGT. Results indicate that people not only switch strategies in the IGT but also that this likelihood increases with trial count. This suggests that when investigating individuals' decision strategies through various sequential decision tasks, it is essential to consider the possibility of strategy switching, particularly in tasks with many trials. Specifically, researchers can develop computational cognitive models that allow strategy switching, as reported in this paper, and compare them with single-strategy models to infer whether and when individuals switch strategies. This approach promises more accurate understanding of individuals' strategy use during different task phases and more reliable inferences from subsequent analyses based on model parameter estimates from different phases.

Although the studies reported here provide clear evidence that individuals may switch strategies in the IGT, the strategy switching considered represents only a subset of possible switching types. Specifically, we assumed that individuals make only one switch between reinforcement learning and heuristic strategies during the entire task and that this switch occurs abruptly. It is also possible that individuals make multiple strategy switches during the task or that strategy switching occurs gradually, transitioning from primarily using reinforcement learning to primarily using heuristic strategies across consecutive trials, or vice versa. From a modeling perspective, the former possibility requires introducing multiple switching points, while the latter would need to employ hybrid models like the VPP model and assume that the weighting coefficient for different strategies (i.e., (!) is a gradual function of trials. While implementing such models is more challenging analytically, it is not impossible. For example, in multi-attribute sequential decision tasks, Lee et al. (2019, 2021) used Bayesian methods to explore models allowing multiple strategy switches, finding that some participants' data were better explained by strategy-switching models, with a small number of participants' data supporting multiple switches. Future research could follow Lee et al.'s approach to examine the possibility of multiple strategy switches in the IGT or other important decision tasks (e.g., risky and intertemporal choice tasks), and could also develop gradual strategy-switching models and compare them with (single or multiple) abrupt-switch models to deepen understanding of the various possibilities of strategy switching. In addition to final choices, other task-related data may also reflect people's decision processes, such as response times and eye-tracking data. Fang et al. (2023) proposed a Machine Learning Strategy Identification (MLSI) method for multi-attribute decision tasks based on mouse-tracking data. This novel approach of using machine learning algorithms to extract decision features and identify decision strategies could be further extended to strategy-switching research in the future. It should be noted that regarding core decision strategies, models for multi-attribute decisions (e.g., Take-the-Best model) are generally deterministic, while models for the IGT are generally probabilistic, making the latter more complex analytically. Additionally, there are important differences between the IGT and multi-attribute decision tasks. For example, each decision in the IGT has explicit feedback, and choices and their outcomes affect subsequent decisions, thus involving more memory and experience accumulation factors. In contrast, each decision in multi-attribute tasks is relatively independent, generally lacks feedback, and thus does not require memory or experience. Furthermore, multi-attribute decision tasks are typically completed under certain information conditions, while the IGT involves more complex uncertain information. Therefore, evidence of strategy switching from multi-attribute decision tasks and the IGT represents converging evidence from different types of tasks, suggesting that strategy switching may exist across various sequential decision tasks with different characteristics.

After confirming the possibility of strategy switching in sequential decision tasks, a key question for further investigation is: What are the conditions that produce strategy switching, or what task factors, individual factors, or their interactions may trigger strategy switching? For example, when task difficulty or personal aspiration levels are high, individuals may switch to different strategies because existing strategies cannot achieve their goals. It can thus be inferred that increasing task difficulty (e.g., requiring wealth levels to increase in the IGT) or raising individuals' aspiration levels might trigger more strategy switching. Additionally, the existence of a dominant strategy may be another factor influencing strategy switching. After trying different strategies and discovering a dominant strategy, individuals' tendency to switch strategies may decrease. Conversely, if multiple strategies yield similar performance, the likelihood of strategy switching will depend on individuals' motivation to achieve better performance and their degree of motivation to explore different strategies. Investigating the triggers of strategy switching will further enhance our understanding of decision strategies and their switching.

Author Contribution Statement: Xinyun Hu: Data cleaning and analysis, manuscript drafting; Yue Shen: Study design, data collection, cleaning, and analysis; Junyi Dai: Research conceptualization, study design, final manuscript revision.

(Corresponding Author: Junyi Dai, E-mail: junyidai@zju.edu.cn)

Note: Figure translations are in progress. See original paper for figures.