Received: 11 October 2019 | Revised: 1 December 2020 | Accepted: 1 December 2020

DOI: 10.1002/bdm.2226

RESEARCH ARTICLE

WILEY

Individual differences in exploring versus exploiting and links to delay discounting

Ethan A. Meyers | Derek J. Koehler

Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada

Correspondence
Ethan Meyers, Department of Psychology, University of Waterloo, 200 University Avenue West. Waterloo, Ontario N2L 3G1, Canada.
Email: emeyers@uwaterloo.ca

Funding information
This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada

Abstract

Sometimes, we must choose between obtaining an immediate reward or foregoing it in favor of searching for a better reward elsewhere. Such decisions have been characterized as involving exploration-exploitation trade-offs. Here, we studied the reliability and basis of individual differences in tasks involving choices between exploration and exploitation. In Studies 1, 2, and 4, we found little evidence for a stable individual difference in tendency to explore (vs. exploit). Additionally, we tested delay discounting as a potential predictor of individual differences in exploration. In Studies 3 and 4, we found that delay discounting was inconsistently predictive of exploration behavior. Our results support the claim that people adapt their exploration behavior to the environment in which they find themselves. This adaptation overrides any general preference to explore environments more or less than other people. Our results also suggest that predictors of exploration may be exclusively restricted to the particular environment in which they were observed. Implications for past and future research of exploration-exploitation decision making are discussed.

KEYWORDS
delay discounting, exploitation, exploration, individual differences, intertemporal choice, reliability

1 | INTRODUCTION

A common decision humans face is whether to explore new opportunities or to continue to exploit their current circumstances. The exploration-exploitation trade-off (Hills, Todd, Lazer, Redish, & Couzin, 2015; Mehlhorn et al., 2015) can be seen as one between gaining new knowledge that might afford future benefits, on the one hand, and extracting benefits from current knowledge, on the other. Likewise, this decision can be framed as whether to "stay" or to "leave" (Cohen, McClure, & Yu, 2007). Staying, in these contexts, usually provides the opportunity to benefit from previous experience or knowledge of the environment. In contrast, leaving offers the chance to seek out a greater benefit somewhere else.

How people (and nonhuman animals) make the decision to explore or exploit has been examined from widely different scientific lenses (neuroscience and RL: Cohen et al., 2007; decision making: Hills & Hertwig, 2010; organizational learning: March, 1991), using a variety of terms and concepts. This paper focuses on an agent faced with only two options: to explore or to exploit. Exploring provides information about potential benefits; exploiting reaps those benefits. Two tasks involving exploration-exploitation trade-offs are used in the current studies and are briefly described here as illustrations. In the observe-or-bet task (Rakow, Newell, & Zougkou, 2010; Tversky & Edwards, 1966), the participant must choose whether to "bet" by predicting which of two outcomes (which of two lights a machine will turn on) will occur, receiving a reward if correct, or instead to "observe" which outcome actually occurs to obtain information on their probabilities. Here, the goal is to use observing to learn the bias of the machine and use that information to predict its future actions. In the Secretary Problem (Ferguson, 1989; Gardner, 1960; Seale & Rapoport, 1997), participants play the role of an employer sequentially interviewing candidates in an effort to find the best person to fill a new job opening. Although only being told how the applicant ranks relative to all of those who have been interviewed before, the participant must decide whether to hire the applicant currently being interviewed or instead to pass on that applicant (who cannot be returned to) as a means of seeking other, better qualified applicants. Here, the goal is to use the information gained from not hiring a candidate to better inform each future applicant consideration. In both tasks, then, participants face a conflict between obtaining a reward now (predicting which light the machine will turn on; hiring the applicant) and gathering information on the distribution of rewards that can be used later (observing which light the machine turns on; passing on an applicant).

People differ in how they resolve exploration-exploitation trade-offs; that is, when faced with an identical environment, some people choose to explore more than others (Gigerenzer & Gaissmaier, 2011; Lee & Newell, 2011; Mehlhorn et al., 2015). Previous research has identified some individual characteristics that predict exploration behavior. People with greater working memory capacity, for instance, tend to explore more (Hills & Pachur, 2012). Older people tend to explore less (Mata, Wilke, & Czienskowski, 2013), whereas people with depression may explore more (von Helversen, Wilke, Johnson, Schmid, & Klapp, 2011).

This paper explores the reliability and basis of individual differences in exploration (vs. exploitation) behavior. We use the two tasks described above, and later, two additional tasks to measure individual differences in the tendency to explore versus exploit. The scarcity of correlation between the trade-off measures provides some insight into the unreliability of broad individual differences in exploration behavior. We then attempt to connect such individual differences with a potentially related construct on which people differ: delay discounting (also known as time preference).

Time preference, specifically in the form of delay discounting, provides an index of how much an individual values future rewards (or punishments, though here we focus exclusively on reward) relative to rewards that can be obtained immediately (Frederick, Loewenstein, & O'Donoghue, 2002; Odum, 2011). A common observation is that immediate rewards are generally preferred over delayed rewards (Kirby, Marlowe, Festinger, Garvey, & LaMonaca, 1999) but people differ in the extent to which they discount time.

We speculate that individual differences in delay discounting may be associated with individual differences in the tendency to explore versus exploit. Exploitation can be viewed as opting for a reward based on knowledge or information that was previously obtained. Exploration can be viewed as foregoing a reward in favor of information collection that may pay off in the future in the form of greater rewards (Mehlhorn et al., 2015). On this account, we would expect those who discount the future more (i.e., are more impatient) would exploit more and explore less than those who are lower in delay discounting (are less impatient).

It is also possible that one's preference to receive sooner but smaller sums is predictive of exploration behavior in only one of the tasks detailed above, or neither of them. These tasks can be further understood by the unique strategies they each require to achieve the best possible outcome. This might affect the association between time preference and exploration in each task. When observing and betting on the machine, an agent must decide how much to explore before predicting the option they believe is correct. Agents must be careful to avoid both underexploring and overexploring. Underexploring leads to more opportunities to earn points but having less information to earn such points. Overexploring leads to the opposite: fewer opportunities to earn points but having more information to earn such points. As indicated by their names, the optimal amount of exploring lies in between the two. A positive linear relation between exploration and a delayed reward preference would suggest that those who prefer the largest (but latest) rewards would tend to overexplore. Yet those looking to receive the highest possible payout should avoid overexploring. This contradiction might lead one to not expect any sort of correlation between exploration and time preference in this task.

When undergoing a sequential hiring practice to find a new employee, one must decide how many applicants it is worth interviewing before hiring the next top-ranked one. To accomplish such a goal, people are likely to apply a fast and frugal heuristic (Gigerenzer & Goldstein, 1996; Stein, Seale, & Rapoport, 2003), namely, to "wait and see." Using this heuristic, people are likely to recognize the need for rejecting some number of applicants initially interviewed before then deciding to hire the next best applicant they encounter. In this case, we can conceptualize such a strategy as delaying immediate (and potentially very good) candidates in order to hopefully later-on hire the best possible one. People should differ in the extent to which they are able to "wait," that is, continue rejecting applicants, even if they are highly ranked. In this circumstance, delay discounting might be uniquely predictive of exploration behavior. Specifically, those who prefer larger but later payoffs might explore (interview more candidates) to a greater extent than those who prefer immediate gratification.

To assess the reliability of exploration behavior across various exploration-exploitation trade-off scenarios, we conducted an initial pilot study and then four subsequent studies. In Studies 1, 2, and 4 as well as the pilot, we tested whether exploration behavior in one trade-off task was associated with exploration behavior in various others. In Studies 3 and 4, we tested whether delay discounting predicts a general tendency to explore (vs. exploit) across trade-off tasks.

2 | PILOT STUDY

2.1 | Method

The data and materials for each study are publicly available online on the Open Science Framework (https://osf.io/cmhnz/)

2.1.1 | Participants

A total of 200 participants were recruited. In all studies, participants resided in the United States and were obtained using Amazon's

Mechanical Turk. The study took about 15 min to complete and workers were paid U.S. $2.00 for their time.

2.2 | Materials

2.2.1 | Exploration-exploitation trade-off tasks

Secretary problem

In the Secretary Problem (adapted from Seale & Rapoport, 1997), participants were instructed that they would be interviewing a fixed number of applicants for a job. They were informed that the pool of applicants contained 40 prospective candidates and that each one had their own individual ranking from 1 to 40 (1 being the best and 40 being the worst). It was made clear that no two applicants were of the same rank, so no two were equal. Participants were told that their goal was to hire as highly ranked an applicant as possible. To accomplish such, they could sequentially interview as many applicants as they wished. The order of applicants was prerandomized and presented in the same order for every participant. The interview process went as follows:

Subjects were only provided with one piece of information during the "interview": the current applicant's relative rank as compared to the other candidates that had been interviewed to that point. Each potential employees' true rank was always hidden and was used to generate the relative ranking that would be displayed to the participant. Due to this, the ranking of the first interviewed candidate was always number one (Figure 1 displays what a participant would have seen when interviewing Candidates 1 and 5).

In addition to being given a relative ranking, participants were provided two options and instructed to select one. They could choose to hire the applicant, ending the task and advancing them to the next portion of the experiment, or they could reject the applicant. Rejection allowed the subject to interview the next candidate. It was made clear that rejection meant that the passed-on interviewee could not be recalled or rehired later.

During each trial of the task, the participant had to decide between continuing the search for the highest ranked applicant or hiring the current interviewee. In this case, we deemed our dependent variable—exploring—to be the number of rejections a participant made. The more rejections made by the participant, the more we deemed them to have explored. The prerandomized order had placed the highest ranked applicant at the 24th position.

Observe-or-bet task

The observe-or-bet task, adapted from Navarro and Newell (2014) after a task by Tversky and Edwards (1966), asked participants to try to predict the actions of an imaginary "Blox" machine. They were told that a Blox machine has two lights, one red and one blue. Every trial (there were 40), the machine would turn one of the two lights on. They were also informed that the machine had a bias favoring one of the two lights but were not told which was favored. In our experiments, the bias was that 66% of the time the red light would turn on (26/40 trials). Each trial the participant would choose from one of three options: (1) observe, (2) predict red, and (3) predict blue. Choosing to predict either of the lights would grant them an opportunity to earn a point (their total remained hidden) providing their prediction was correct. However, they would receive no feedback about which light turned on. After selecting either of the prediction options, the computer would restate their prediction before moving them on to the next trial. For example, if "predict red" was selected, the participant would see "You predicted the Red light will turn on." Because making a prediction would gather zero information about the Blox machine, observing was the only way to learn its habits. By forgoing the chance to earn a point, a participant could choose to observe which light turned on that trial. After selecting the observe option, the computer would inform the participant the color of the light that turned on and then advance them to the next trial. For this task, we considered the total number of observations made as our measure of exploration.

Delay discounting measure

The delay discounting measure (from Kirby et al., 1999) was used to examine each participant's time preference. The measure features 27 questions pitting an immediate and delayed reward against each other (e.g., "Would you prefer $27 today, or $50 in 21 days?"). Participants were instructed to select the amount that they would prefer to receive, although they would not actually collect the money. Each question varied in both dollar amount and time difference between the two rewards. Time preference was measured by the total number of times a participant selected a delayed reward.

Procedure. The study's design was within-subjects. Participants completed both exploration-exploitation tasks in a counterbalanced order. Then, subjects completed the delay discounting measure. Finally, demographic questions were answered, and the study concluded.

2.3 | Results

Nine participants were identified as outliers on at least one of the tasks completed (a score more than ±3 SDs away from the mean) and were subsequently excluded from analysis. We first assessed the reliability of exploration behavior across tasks. Notably, exploring in one

FIGURE 1 The information displayed to participants when interviewing Candidates 1 and 5

Candidate 1	Candidate 5
"This applicant's relative rank is: 1 out of 1 applicant seen so far."	"This applicant's relative rank is: 3 out of 5 applicants seen so far."

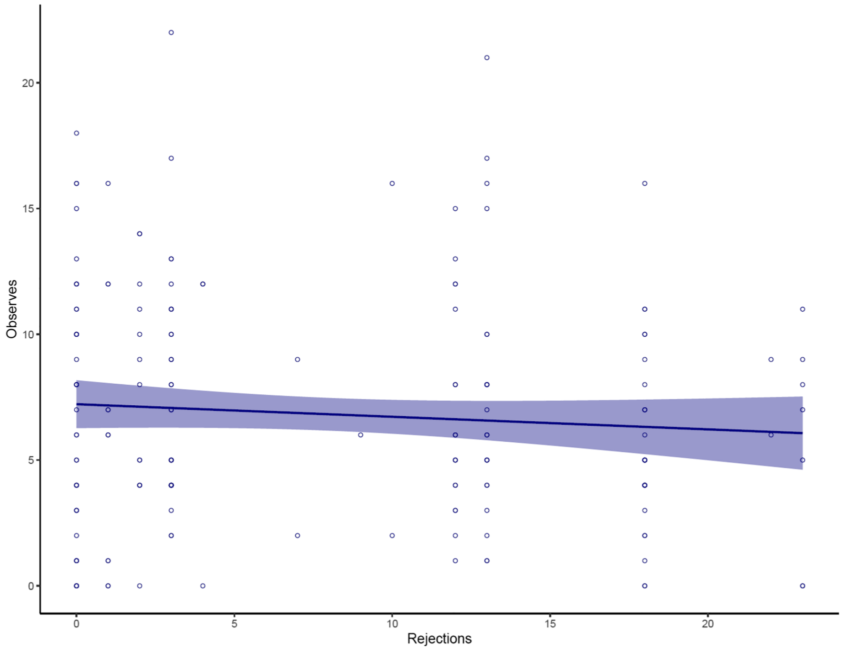

FIGURE 2 Scatterplot of Pilot Study data. The light shaded region surrounding the line of best fit represents a 95% confidence interval. Observes is the number of observations made in the observe-or-bet task. Rejections is the average number of candidates rejected in the secretary problem [Colour figure can be viewed at wileyonlinelibrary.com]

task was not associated with exploring in the other, $r(187) = -.08$, $p = .257$ (see Figure 2). That is, the number of candidates a participant rejected ($M = 8.01$, $SD = 7.50$) was not related to how many times they observed the actions of the Blox machine ($M = 6.82$, $SD = 4.55$). Next, we examined the link between delay discounting and exploration behavior. A preference to receive larger but later rewards was not related to the number of rejected candidates, $r(187) = .04$, $p = .547$, or the number of observations made, $r(187) = -.11$, $p = .093$. Participants elected to receive the delayed reward in less than half of the trials ($M = 11.88$, $SD = 6.15$), $t(190) = -3.64$, $p < .001$.

2.4 | Discussion

We found initial support for the claim that one's tendency to explore (vs. exploit) a given environment is not consistent across established trade-off tasks. We also found no evidence to support the idea that time preference is related to a tendency to explore.

A major concern for this study was data quality. As no comprehension questions were implemented, task understanding could not be assured. Moreover, we did not try to identify and differentiate respondents who were earnestly completing the study versus those who were attempting to complete it as fast as possible for the relatively large payoff ($2 USD). More than 20% of the sample hired the very first applicant in the secretary problem. We reasoned this was likely due to the two factors mentioned above: either the participants did not understand that the very first applicant they interviewed would be rank 1 or they simply wanted to skip through our study. As a result, we are hesitant to draw any firm conclusions based on the Pilot study alone.

We decided to more comprehensively assess task understanding in future work. In order to avoid increasing the length of the study (due to concerns of participant attrition), we parsed our two research questions and addressed them one at a time. So we conducted the next study to assess the reliability among exploration behavior with an explicit focus on ensuring task comprehension.

3 | STUDY 1

3.1 | Method

3.1.1 | Participants

A total of 200 participants were tested. The study took about 15 min to complete and workers were paid U.S. $2.00 for their time.

3.1.2 | Materials and procedure

This study featured only two measures: The secretary problem and the observe-or-bet task. In order to increase the rate of task comprehension in the sample, the instructions for each task were expanded, providing greater detail about how the tasks worked and what the responsibilities of each subject would be. In addition, a comprehension question was added to the end of the instructions for each task (see Figure 3). The multiple-choice questions each had multiple correct answers the participant was required to select in order to pass. Following the advice of Oppenheimer, Meyvis, and Davidenko (2009), for each question, failing to provide the correct answers would lead to the instruction set being repeated. Repeated failures (four times for the secretary problem and five times for the observe-or-bet task[1])

[1]These amounts were based off the number of possible answers there were to choose from.

FIGURE 3 The comprehension questions implemented for Study 2. The correct options are bolded

Secretary Problem	Observe – or – Bet
To check your understanding of the task you will be completing, please answer the following question (if your answer is incorrect the instructions will be repeated).	To check your understanding of the task you will be completing, please answer the following question (if your answer is incorrect the instructions will be repeated).
Suppose you have interviewed 6 out of the 40 applicants so far, and the relative ranking of the 6th applicant is 1 out of _6_ applicants seen so far. Which of the following is true (check all that apply)?	Which of the following is true (check all that apply)?
a) Applicant 6 is definitely the top-ranked applicant of all 40 applicants.	a) When I choose to 'Predict' I can earn a point if my prediction is correct and will be told whether or not my prediction was correct (whether the light was red or blue)
b) Applicant 6 may or may not be the top-ranked applicant of all 40 applicants.	**b)** When I choose to 'Predict' I can earn a point if my prediction is correct and will NOT be told whether or not my prediction was correct (whether the light was red or blue)
c) Applicant 6 is definitely not the bottom-ranked applicant out of all 40 applicants.	c) When I choose to 'Observe' I will be told whether the red or blue light turned on, and will also be able to earn a point for making a correct prediction
	d) When I choose to 'Observe' I will be told whether the red or blue light turned on, but will NOT be able to earn a point for making a correct prediction

would lead to the participant having their data removed prior to analysis. Importantly, because the correct answer required multiple choices to be selected and that participants were not informed which of their choices were or were not correct upon failure, a participant could not simply try all the unique combinations possible to pass through.

The secretary problem also had a second hiring trial added to it. The second trial was created by taking the prerandomized order of the original task and reversing it (the applicant with the true rank of 1 would be the 17th candidate interviewed rather than the 24th). All participants completed both trials (in counterbalanced order) that were framed as hiring for two separate positions.

A within-subjects design was used for this experiment as participants completed both tasks in counterbalanced order.

3.2 | Results

Removing all participants who failed at least one of the comprehension checks (n = 51) left a sample of 149. A further three participants were excluded from data analysis as outliers. Consistent with the Pilot study, we found no relation between exploring in one task and exploring in the other, $r(145) = .15$, $p = .071$ (see Figure 4). That is, the average number of candidates a participant rejected (M = 13.52, SD = 9.15) was not related to how many times that participant observed the Blox machine (M = 7.76, SD = 2.39).

We next assessed the reliability of the Secretary Problem. Rejections in one trial were strongly related to the number of rejections in the other trial, $r(145) = .71$, $p < .001$. So although a general tendency to explore within this task appears to be present, the previously discussed result suggests this is unrelated to the tendency to explore in the observe-or-bet task.

Finally, we assessed the internal consistency of the observe-or-bet task to ensure that the lack of association between exploration behavior was not due to random responding within this task. A split-half analysis revealed that responses on even-numbered trials were highly similar to responses on odd-numbered trials, $r_{sb} = .83$.[2] That is, throughout the task, participants were quite consistent with their choices. This suggests that the lack of association between exploration behavior was not due to abnormal responding patterns on the observe-or-bet task.

3.3 | Discussion

Consistent with the Pilot study, we found that exploration behavior was unreliable across tasks. Specifically, even after ensuring that participants both sufficiently understood the tasks were responding in a nonerratic manner, exploration behavior in one task was unrelated to exploration behavior in the other. This is unexpected as both tasks are considered to measure the exploration-exploitation trade-off (e.g., Mehlhorn et al., 2015) and have been widely used in this area of research.

[2]This is consistent with the coefficient of the Pilot study data, $r_{sb} = .73$.

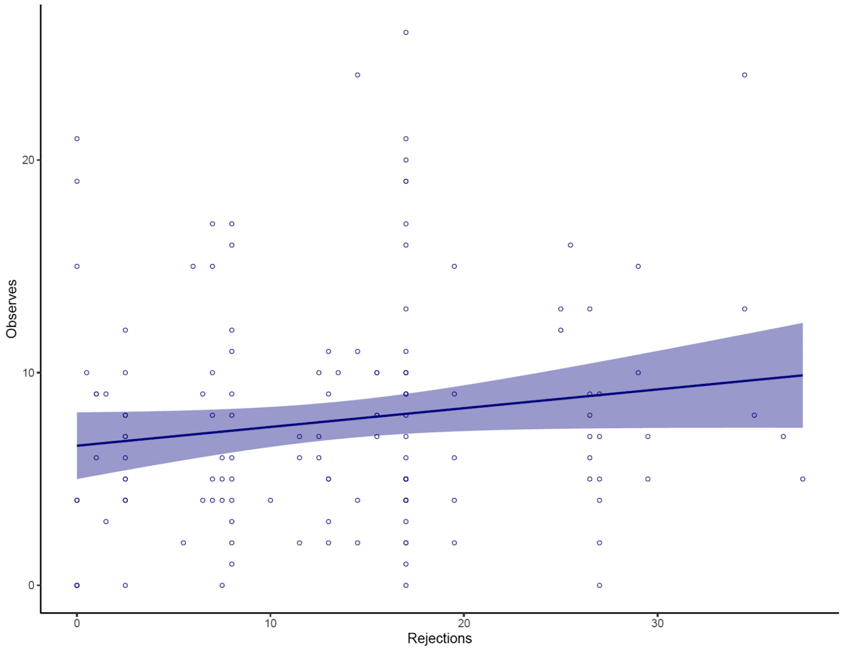

FIGURE 4 Scatterplot of Study 1 data. The light shaded region surrounding the line of best fit represents a 95% confidence interval. *Observes* is the number of observations made in the observe-or-bet task. *Rejections* is the average number of candidates rejected in the secretary problem [Colour figure can be viewed at wileyonlinelibrary.com]

Two major concerns of Study 1 exist. The first is that it is entirely possible that people have a general tendency to explore a given environment in which they differ in but that at least one of the two implemented measures is not actually assessing what it has been proclaimed to. Thus, the lack of reliability of exploration behavior might be an artifact of task selection. The second concern is that participants have not been provided any incentive to meaningfully engage in the tasks. So one possible interpretation of our current results is that *unincentivized* exploration behavior (or behavior that is not truly representative of a person's real tendencies) is not reliable across tasks.

The next study aimed to address the concerns described above. In addition to ensuring each task would be completed twice, we further added two more tasks that assess the exploration-exploitation trade-off. We also provided the opportunity for participants to earn several large bonuses to incentivize engagement.

4 | STUDY 2

4.1 | Method

4.1.1 | Participants

A total of 409 participants were tested. The study took about 10 min to complete, and workers were paid U.S. $2 for their time. As an incentive, participants were instructed that the Top 2 performers on each task would win a bonus $5 in addition to their remuneration for participating.

4.1.2 | Materials

The observe-or-bet game and the secretary problem each underwent a small modification for this study. The observe-or-bet game had a second round added to it. In this study, one round would last for 20 trials, and the other round would last for 30 trials. The secretary problem also had one of its rounds extended, featuring a 60-applicant pool, whereas the other round remained as a 40-applicant pool. In addition, two further trade-off tasks were added. These are described below.

4.1.3 | The sampling paradigm

In the sampling paradigm adapted from Hertwig and Erev (2009), participants are presented with two boxes representing lotteries on a screen. Each lottery has a distribution of possible payouts and an associated expected value that is unknown to the participant. Subjects are told they are to ultimately select the box that they would prefer to draw from where they going to receive the payment from whatever value they pulled. To learn the payoff distributions of both lotteries, they can sample an unlimited number of times from each box. Participants sample a lottery by clicking on its respective box and receiving a number taken from the payoff distribution. When ready, they can select the box that they would make their "true" draw from. Exploration is defined as the number of times a subject elected to sample from either box.

4.1.4 | Search task

In the search task adapted from Sang, Todd, Goldstone, and Hills (2018), participants try to accrue as many points as possible over a set number of turns (20 and 30 in our study). Participants "draw" without replacement from a deck of 99 cards each featuring a unique value (1 to 99) on its face. Each turn after the first (because the first turn has a mandatory draw) participants can choose to either gain the points of a card they have already flipped over (exploit) or to draw another card (explore). When a card is drawn, it remains on the player's screen for the rest of the game. That is, a card's points can be gained more than once. Players are always aware of their current point total. Exploration is defined as the number of cards drawn.

Procedure

To avoid running a lengthy online experiment, each participant was randomly assigned to complete only two of the four possible exploration tasks. They were also informed that the top two performers of each task they were to complete would be granted an additional $5 on top of their remuneration. Each task was played twice, with a slight variation between each round (e.g., the search task had one round of 20 turns and one round of 30 turns; the observe-or-bet task had one round of 20 turns and one round of 30 turns; the sampling paradigm had two rounds of two boxes each with a unique payoff distribution; the secretary problem one round had 40 candidates; and the other had 60). The order in which the tasks and their respective rounds were presented was randomized. After completing two tasks, each participant provided some basic demographic information and was asked to what extent they believed they would get paid the bonus if they were to win it. Following this, the study concluded.

4.2 | Results

A further 43 participants were removed prior to analysis due to either being identified as an outlier (more than 3 SDs above or below the mean) or not completing at least one of the assigned tasks. This left a final sample of 366. As responses to the sampling task were non-normally distributed, a square root transformation was applied. Descriptive statistics for this study can be found in Table 1.

Table 2 contains all correlations of interest as well as accompanying condition sizes. Importantly, in only two instances was exploration behavior in one task significantly associated with exploration behavior in another. The greater the number of candidates a participant rejected on average, the more samples they drew from the box-lotteries on average, $r(57) = .29$, $p = .025$. In other words, exploring in the secretary problem was correlated with exploring in the sampling paradigm. Surprisingly, the greater the number of observations a participant made on average during the observe-or-bet task, the smaller the number of box-lottery samples they drew on average during the sampling paradigm, $r(54) = -.33$, $p = .014$. In this case, although exploration behavior across these tasks is related, the correlation is negative, suggesting that a tendency to explore more in one environment is related to exploring less in the other. No other correlations between exploration behavior in any two tasks were significant.

The lack of reliability of two of the measures implemented in this study (the observe-or-bet task and the search task) raises questions about the validity of the results. To crudely test this concern, we took each round of the tasks independently and assessed the correlation between the tasks again. Although of similar magnitude to the previous statistic, the associations between each round of the observe-or-bet task and the sampling paradigm failed to retain statistical significance after correcting for Type 1 error rate inflation. Likewise, no significant correlation between these trials and exploration on the remaining tasks was found. In other words, the number of cards drawn and the number of times the Blox machine was observed appear unrelated to any other form of exploration behavior assessed in this study.

TABLE 1 Descriptive statistics for study 2

	n	M	SD
Observations	188	6.90	3.48
Cards drawn	193	15.88	4.12
Candidates rejected	194	13.90	11.34
Lottery samples	196	3.70	4.01

Note. Cards drawn is the average number of cards drawn in the search task; this variable was statistically transformed for the purposes of analysis; these are the untransformed values. *Observations* is the average number of times participants selected to observe in the observe-or-bet task. *Lottery samples* is the average number of times the box lotteries were sampled before choosing the final draw. *Candidates rejected* is the average amount of candidates rejected.

4.3 | Discussion

In response to the limitations of the previous study, we included additional trade-off tasks and incentivized engagement to further our assessment of the reliability of the general tendency to explore a given environment. When comparing exploration behavior across four trade-off tasks, we found that exploration in one task was seldom related to exploration in another. In only two cases was the correlation between tasks significant, yet both were of a modest size, and only one was in the expected direction. Moreover, providing the opportunity to earn large rewards based on performance did not appear to meaningfully change behavior or produce different results (i.e., lack of any significant correlations) relative to the previous studies. Thus, although all the tasks we have implemented so far are thought to assess one's tendency to tradeoff between exploring and exploiting, our results exhibit a stark lack of behavioral reliability across these tasks.

We consider three possible explanations for the lack of reliability across explore-exploit tasks. The first is that a general tendency to explore a given environment in which people can differ might simply not exist. The logic of this explanation would suggest that exploration in one task is unlikely to be related to exploration in another. This is because one individual is not more likely to explore multiple environments more so than any other person. Although our results can be interpreted to mostly support this claim, there are several well-documented examples whereby people differ in their tendency to explore from one another and this persists across environments. For example, “maximizers” tend to consistently overexplore environments in contrast to their typically underexploring “satisficer” counterparts (Bruine de Bruin, Parker, & Fischhoff, 2007; Schwartz et al., 2002).

TABLE 2 Correlation coefficients (Pearson's *r*) among exploration behavior across tasks

	1	*n*	2	*n*	3	*n*	4	*n*
1. Observations	(.20)*							
2. Cards drawn	−.20*	*61*	(.04)					
3. Rejections	−.02*	*71*	−.11	*61*	(.46)*			
4. Samples	−.33*	*56*	−.23	*52*	.29*	*59*	(.78)	

Note. Cards drawn is the average number of cards drawn in the search task with a square-root transformation applied. *Observations* is the average number of times participants selected to observe in the observe-or-bet task. *Samples* is the average number of times the box lotteries were sampled before choosing the final draw. *Rejections* is the average amount of candidates rejected. Values in parentheses represent the correlation between rounds of each task. These values for *rejections* and *samples* are significant at the $p < .01$ level; neither of the other two reach statistical significance.
*$p < .05$.

Likewise, those with a greater working memory capacity tend to explore more in general perhaps because they are able store more information gained from exploration than those with lower working memory capacity (Brydges, Heathcote, & Braithwaite, 2008; Hills & Pachur, 2012). So this explanation appears generally unsupported.

The second possible explanation is that some of, if not all, the tasks we have implemented in the present work do not accurately assess the exploration-exploitation trade-off. If this were to be true, our results would not be surprising as it would mean that the behavior being measured across tasks is not consistently "exploration" but instead is undefined. However, given that we only used tasks that have been previously implemented in trade-off research (e.g., secretary problem: Seale & Rapoport, 1997; Seale & Rapoport, 2000; observe-or-bet: Navarro, Newell, & Schulze, 2016; Rakow et al., 2010), there exists a set precedent that these tasks all measure exploration-exploitation decision making. Yet it remains possible that the way we have implemented the tasks could have altered the construct validity of any number of them. Although we challenge this idea, the point receives further consideration in the General Discussion.

The third possible explanation is that individual differences in the tendency to explore a given environment may exist but are eclipsed by the goal directed behavior of the agent that is specific to each task. In other words, people adapt their behavior to the environment in which they find themselves, and this takes precedent over any general tendency they may have to want to explore the environment more than another person. One way to conceptualize this is that in each task, there exists a general strategy to obtain success that most people will attempt to mimic (to varying degrees of success). For example, in the secretary problem, the optimal strategy is to reject a large portion of the initial applicants regardless of their rank before hiring the next highest-ranked applicant interviewed. This is quite distinct from optimal strategy in the observe-or-bet game wherein participants must decide whether or not they have acquired sufficient knowledge of which light the machine favors. Thus, the individual strategies one implements to succeed in the task (initial mass rejection vs. careful gathering and weighing of information) are likely to eclipse any broad individual tendency to explore (vs. exploit). Moreover, this claim is consistent with recent independent work by von Helversen, Mata, Samanez-Larkin, and Wilke (2018), who found that there was no single general factor underlying exploration behavior across three unique trade-off tasks. We suggest that exploration behavior may be domain specific instead of domain general and advance our line of research to assess this possibility.

Within the data collected already, we wondered if we could already assess this to some extent. The secretary problem is a one-shot type of trade-off where a person is unable to switch back and forth between exploring and exploiting. Once they make the decision to hire (exploit), the task is over, and only the initial amount of exploration completed is considered. In contrast, the observe-or-bet task freely permits the agent to swap back and forth. For example, a participant may choose to observe for the first three trials (explore), guess which light will turn on for the next several (exploit), and then decide they want more information and begin to observe again. Later switches back to exploring might be due to concerns that the participant has forgotten, misidentified the better option, or to relieve boredom (Newell, Koehler, James, Rakow, & van Ravenzwaaij, 2013). Alternatively, they may believe the better option has changed over trials. Thus, decisions to explore made later in the game (postexploiting) may be the result of reasons entirely separate from why someone may reject a candidate in the secretary problem. By using total observations, we may be treating two different types of exploration behavior as if they were equivalent. Initial observations, those made before the first decision to start exploiting has been made, may be more comparable to exploration in the secretary problem. Indeed, if exploration is reliable across environments, then we would expect an individual who rejected more candidates to also have initially observed more. So we correlated the initial observations made in the observe-or-bet task with the number of candidate rejections in the secretary problem. We found no support for this relation when examining the data of Study 2, as initial observations were not correlated with candidate rejections, $r(69) = .04$, $p = .745$. However, we found that the more a participant observed initially, the more likely they were to reject a greater number of candidates in both Study 1, $r(147) = .19$, $p = .022$, and the Pilot study, $r(198) = .20$, $p = .004$. So there exists some modest support for the contention that a broad individual difference in the tendency to explore (vs. exploit) is eclipsed by particular exploration strategies specific to a given environment.

Another way to test whether exploration behavior manifests uniquely across environments would be to look at individual differences that one might expect to map onto exploration behavior as it is broadly defined. Delay discounting represents a variable that one could reasonably expect to predict behavior across exploration-exploitation trade-off paradigms. If the trade-off is viewed as deciding between acquiring an immediate reward now (exploiting) versus delaying gratification and gathering more information to be used to obtain greater rewards later (exploring), then delay discounting might map onto exploration broadly. However, if exploration behavior is at least somewhat dependent on the environment, then delay discounting might not be predictive of exploration at all. The other potential outcome from exploration behavior being environmentally dependent is that delay discounting might be uniquely associated with exploration behavior in one paradigm and not in the other. In Study 3, we test these possibilities.

5 | STUDY 3

In order to test the link between delay discounting and exploration behavior, we drew from the experimental designs of Study 1 and the initial Pilot study. We combined the improved instruction sets for the secretary problem and the observe-or-bet task with the previously implemented delay discounting measure.

5.1 | Method

5.1.1 | Participants

A total of 403 participants were tested. The study took about 15 min to complete, and workers were paid U.S. $2.00 for their time.

5.1.2 | Materials and procedure

In this study, participants completed only one of the exploration tasks (observe-or-bet or the secretary problem) assigned at random and then completed the delay discounting measure. The exploration tasks featured the comprehension checks and more detailed instructions implemented in Study 1. Finally, some basic demographic questions were answered, and the study concluded.

5.2 | Results

After removing those who failed their respective comprehension check, the sample contained 364 subjects (39 failed). Of these, 186 completed the observe-or-bet task and 178 completed the secretary problem. A further three participants were excluded for outlier scores in the observe-or-bet task prior to analysis.

We assessed the relation between delay discounting and exploration behavior for each task. Rejecting candidates in the secretary problem (M = 11.54, SD = 9.42) was positively associated with a preference for delayed rewards, $r(176) = .44$, $p < .001$ (see Figure 5). That is, the more an individual elected to explore in the secretary problem, the more likely they were to prefer larger but later sums. In contrast, observations of the Blox machine (M = 7.91, SD = 6.15) were not associated with a preference for delayed rewards, $r(182) = .09$, $p = .242$. That is, decisions to explore in the observe-or-bet task were unrelated to one's time preference. Participants elected to receive the delayed reward in less than half of the trials (M = 10.39, SD = 6.70), $t(361) = -8.49$, $p < .001$.

Next, we assessed the internal consistency of the observe-or-bet task to ensure that the lack of association was not due to random responding within this task. A split-half analysis revealed that responses on even-numbered trials were highly similar to responses on odd-numbered trials, $r_{sb} = .94$. That is, across trials, participants made choices that were highly consistent. Likewise, we ensured the reliability of responses across iterations of the secretary problem. Rejections in one round of the game were highly correlated with rejections in the other round, $r(176) = .74$, $p < .001$. These findings are consistent with the previous studies, suggesting that these results were not the consequence of poor consistency among choices within our measures.

5.3 | Discussion

We found that a preference to forgo an immediate payoff in favor of a delayed but larger payoff predicted exploration (vs. exploitation) behavior, but in only one of the two tasks implemented. Specifically, delay discounting was associated with exploration in the secretary problem, but not with exploration in the observe-or-bet game. This finding is consistent with a domain-specific model of exploration behavior that suggests people adapt their exploration behavior to fit the environment in which they find themselves. If individual differences in exploration are measurable, one could reasonably expect them to be predicted by delay discounting (or not at all). However, if it is the case that delay discounting is uniquely predictive of exploration in one task but not another, it would lend support to the domain-specific claim. This is what we found.

Whereas the demonstrated (lack of) association between observe-or-bet and delay discounting was consistent with the Pilot study, the secretary problem association was not. We found a sizeable correlation in this study between a preference for delayed rewards and candidate rejections but found no evidence for this relation in the pilot work. Despite the data quality of the pilot work being questionable as previously discussed, such a conflicting result deserves additional testing. Our next study sought to further explore the reliability of individual differences in exploration behavior and the link to delay discounting. Specifically, we wanted to replicate the modest, but significant relation between exploration

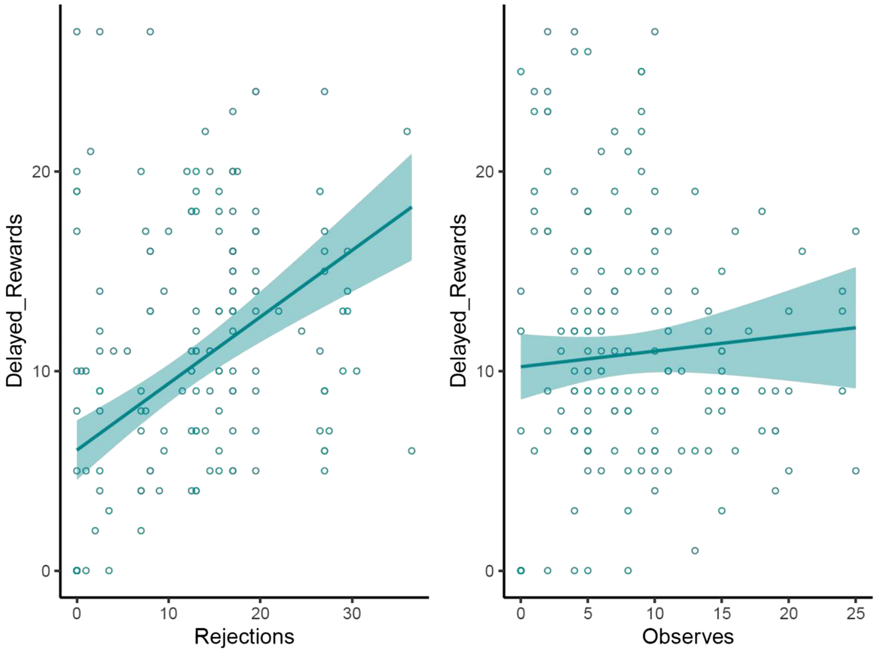

FIGURE 5 Associations between exploration and number of delayed rewards selected in Study 3. The light shaded region surrounding the line of best fit represents a 95% confidence interval. *Observes* is the number of observations made in the observe-or-bet task. *Rejections* is the average number of candidates rejected in the secretary problem [Colour figure can be viewed at wileyonlinelibrary.com]

behavior in the secretary problem and the sampling paradigm found in Study 2. As the reliability of the sampling paradigm was previously very low, we also looked to improve this through increasing the number of iterations of the task (the number of trials for the secretary problem was also increased). We also sought to assess the predictive link between delay discounting and the secretary problem, and by nature of the study's design, we would also be able to assess the same association for the sampling paradigm.

6 | STUDY 4

6.1 | Method

6.1.1 | Preregistration

This study was preregistered and the preregistration is available on the OSF here (https://osf.io/cmhnz/).

6.1.2 | Participants

We initially recruited 201 participants to provide responses for both exploration tasks but separately excluded participants from each task (32 for the secretary problem and 11 for the sampling paradigm) if they responded incorrectly to its comprehension questions. The study took about 20 min to complete and workers were paid U.S. $2.00 for their time. As an incentive, participants were instructed that the Top 2 performers for each task would win a bonus $5 in addition to their remuneration for participating.

6.1.3 | Materials and procedure

In this study, participants completed the secretary problem and the sampling paradigm in counterbalanced order and then completed the delay discounting measure. Each exploration task had its number of iterations increased to four (so each participant would make four hiring decisions and choose between a pair of lotteries four times). The instruction set for each exploration task featured a comprehension question that upon failing would restart the instructions. Participants who failed the comprehension question three times were automatically advanced, and their data would be excluded prior to analysis. The study concluded with basic demographic questions.

6.2 | Results

The data for the sampling paradigm were nonnormally distributed so a square-root transformation was applied. We first assessed the relation between exploration behavior for each task. Candidate rejections (M = 12.33, SD = 5.45) were significantly correlated with lottery samples (M = 10.48, SD = 8.93), $r(160)$ = .24, p = .002 (see Figure 6). That is, consistent with Study 2, we found a modest relationship between exploration in the secretary problem and the sampling paradigm. We then assessed the link to delay discounting. Secretary problem rejections were not associated with a preference to receive delayed rewards, $r(167)$ = .08, p = .302. Although consistent with the data from the Pilot study, this is inconsistent with the result found in Study 3. Lottery samples were associated with delay discounting, $r(188)$ = .21, p = .003, suggesting that the more samples an individual drew, the more likely they were to prefer receiving larger but later sums. In approximately half of the trials, participants elected to receive the delayed reward (M = 13.72, SD = 6.29), $t(161)$ = 0.45, p = .654.

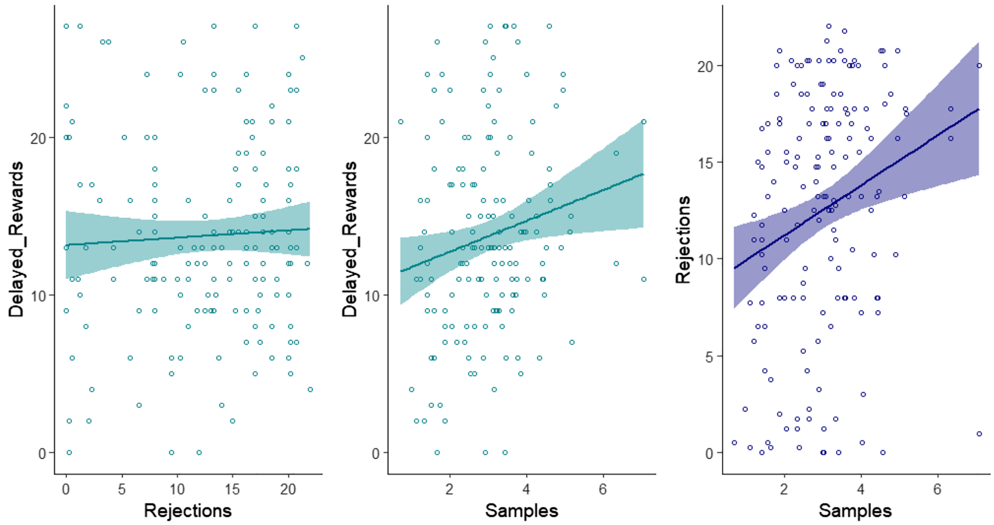

FIGURE 6 Associations between exploration behavior and number of delayed rewards in Study 4. The light shaded region surrounding the line of best fit represents a 95% confidence interval. *DelayedRewards* is the number of delayed rewards chosen out of 27 options. *Samples* is the average number of samples drawn from the box lotteries with a square-root transformation applied. Rejections is the average number of candidates rejected in the secretary problem [Colour figure can be viewed at wileyonlinelibrary.com]

The internal consistency of each exploration task was then assessed. Measurements across the four iterations of the sampling paradigm (α = .93) and the secretary problem (α = .88) possessed high reliability.

6.3 | Discussion

To further assess the reliability of exploration behavior and links to delay discounting, we conducted a well-powered, incentivized study. We found a modest but significant relation between exploration behavior across trade-off tasks. This pattern of data was highly consistent with the data of Study 2. Although one interpretation of this finding is that a general tendency for some people to explore more than others is measurable, we believe the weak strength of the relation lends greater support to the claim that exploration behavior is adaptive and highly variant across tasks. In further support of this claim, we also found that exploration behavior in the sampling paradigm was predictive of delay discounting, but inconsistent with Study 3, exploration behavior in the secretary problem was not.

7 | GENERAL DISCUSSION

The present investigation sought to shed new light on the reliability of assessing individual differences in the tendency to explore versus exploit. However, we found that exploration behavior in one trade-off task is not consistently related to exploration behavior in another. Although we did find one consistent relation between exploration behavior in the sampling paradigm and in the secretary problem, the strength of the association was modest and far weaker than what one would expect were a tendency to explore be highly reliable across trade-off environments. We suggest that this is because people adapt their exploration behavior to fit the environment (however imperfectly) in which they find themselves and that this takes precedent over any general tendency they may have to explore the environment more than another person. We found further evidence for this idea in Studies 3 and 4 by assessing the link between exploration behavior and delay discounting. In these studies, delay discounting was predictive of exploration behavior in one task (sampling paradigm), was not predictive in another (observe-or-bet), and was inconsistently predictive of a third (secretary problem). If exploration behavior was a domain-general phenomenon, one would expect that delay discounting would be reasonably consistent in its predictiveness of exploration behavior. However, our results suggest that in some cases it is, in some cases it is not, and in some cases it is not clear. So we suggest these results lend support to the claim that exploration behavior is a domain-specific phenomenon due to its apparent environmental malleability.

Although all the tasks we implemented in the present work are conceptualized as requiring a trading-off between obtaining new information and using the information one has accumulated (Mehlhorn et al., 2015; Sang et al., 2018), decisions in these types of situations appear only loosely related across environments. This is highly consistent with the recent work of von Helversen et al. (2018), which initially demonstrated the difficulty in detecting general exploration tendencies across unique exploration-exploitation trade-off tasks. One reason for why we may have not found reliable individual differences in exploration is the variance in "transition type" across the tasks implemented. In some (secretary problem, sampling task), the transition from exploration to exploitation is clear and instantaneous. In these cases, exploration can occur only before the first decision to exploit has been made. So all exploration decisions presumably serve the same single purpose (gather enough information with which I will make all my decisions). In contrast, in other tasks (observe-or-bet and search task), the transition is not as clear, as one is able to switch between the two modes freely. In these cases, exploration decisions can occur under several unique motivations and later exploring (exploring after one has exploited previously) may be conducted to serve an entirely different purpose than initial exploration. In the context of the observe-or-bet task, one may be first observing to gain more information to inform their prediction; later (after some predictions have been made), one may be exploring in an attempt to reduce boredom or to ensure their choice is correct (Newell et al., 2013). Notably, one could reasonably expect variance in transitions to affect assessments of reliability across tasks. However, it would make little sense if this were to entirely explain the lack of reliability we found. In two of the tasks that feature an identical transition (sampling paradigm and secretary problem), we did find a significant relationship in exploration behavior, but that relationship was small. Despite holding transition type constant, we were unable to find any substantive reliability in exploration behavior. Likewise, although the data are not as reliable in this instance, we also failed to find a relation between observe-or-bet exploration and search task exploration. So although transition type could play some small role in obscuring measurement of individual differences in exploration, it cannot solely account for our results.

Our claim is further supported by our examination into the relation between delay discounting and exploration behavior. If one conceptualizes exploration-exploitation trade-offs as choosing between foregoing an immediate payoff to obtain information that can be used to potentially gain a delayed but greater payoff (exploration), then our results are surprising. As delay discounting is assessing one's preference for sooner but smaller versus later but larger rewards, it should likely predict one's tendency to explore a given environment. However, if our claim is true that individuals adapt their exploration behavior in an attempt to fit the demands of the environment, then delay discounting should inconsistently predict exploration behavior across trade-off tasks. Our results support this interpretation. Delay discounting was predictive of exploration in the sampling paradigm and not in the observe-or-bet task. Moreover, we found evidence that delay discounting was predictive of exploring in the secretary problem in one case (Study 3), and not predictive of it in two cases (Pilot Study and Study 4).

Our findings have important implications for work on individual differences in this area of research. To date, most individual differences work on trade-off decision making has implemented only one task to assess exploration (e.g., Hausmann & Läge, 2008; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; Newell & Lee, 2011). As we show here with the link to delay discounting, individual differences may not generalize across all trade-off environments, even if it would make conceptual sense to. Researchers must be careful to avoid concluding that a factor that predicts one's tendency to explore versus exploit in one environment is true for all or even most trade-off environments. Instead, researchers should try to specify in what types of environments (or as we suggest, which types of strategies by way of environment) their individual difference predictor holds across. This recommendation is consistent with recent work demonstrating that within a single trade-off task, delay discounting correlates with one type of exploration but not another (Sadeghiyeh et al., 2020). Moreover, we suggest that it might be wise to express caution toward work forwarding a predictor of exploration behavior that has not been consistently demonstrated across multiple environments.

Our work draws parallel with recent advances in the study of risk. Comprehensive examinations of behavioral methods intended to elicit risk preferences have shown weak correlations (if any at all) between these measures and that people's preference for risk is dependent on the measure implemented (Frey, Pedroni, Mata, Rieskamp, & Hertwig, 2017; Pedroni et al., 2017). Consistent with the argument we are presenting here, it has been suggested that at least part of the reason why behavioral risk preference measures appear unreliable is because the same decision strategy is not applied evenly across all tasks (Pedroni et al., 2017). Likewise, this might further explain why some correlates of risk preference (e.g., fluid intelligence and household income) show only domain-specific associations dependent on the measure used to assess risk (Frey, Richter, Schupp, Hertwig, & Mata, 2020). We suggest that it might be wise to draw from this work to inform future investigation into improving the reliability of exploration-exploitation trade-off assessment. For instance, perhaps surprisingly, self-report assessments of risk preference are reliable (Frey et al., 2017; Steiner, Seitz, & Frey, 2019). Although it may not be the most intuitive way to assess exploration tendencies, the development of self-report measures might prove a fruitful avenue for which to pursue. Moreover, this possibility might also allow for a stronger test of self-report correlates of exploration behavior (e.g., delay discounting) as the increased error from the divergent response processes required for behavioral measures and self-report measures will be absent (see Dang, King, & Inzlicht, 2020).

7.1 | Limitations

A limitation consistent throughout our experiments was the smaller number of trials we implemented for each task as compared to their original versions. For example, the observe-or-bet game is typically played with hundreds of trials (e.g., Navarro & Newell, 2014), and the secretary problem is often played more than two or four times (e.g., Seale & Rapoport, 1997). Due to this, we accept the possibility that we do not have enough trials to measure individual differences in long-term exploration versus exploitation behavior (e.g., see Navarro et al., 2016, for evidence that humans take at least some time to adjust to near-optimal strategy in the observe-or-bet task). However, we would suggest that the most parsimonious explanation is that we are indeed assessing exploration (vs. exploitation) decision making, as it is unclear what else we might be measuring if not humans exploring a new environment and deciding between gaining more information and using what they have learned. Yet it remains possible that the trade-off behavior we observed is fundamentally different from the behavior our participants would have exhibited were the tasks at their full length. We suggest then that future work could explore the potential lack of reliability among these tasks with a greater number of trials.

Another limitation is that fiscal incentives were not offered in all the presented work. In Studies 2 and 4, participants were fiscally incentivized, whereas in Studies 1 and 3, they were not. So it is possible that for those unincentivized samples, we may not have been adequately assessing "true" behavior and preferences. However, this only matters to the extent that it can explain our pattern of results. We would argue that a lack of fiscal incentivization cannot. We make this argument on two grounds. First, the pattern of data and our conclusions are consistent across studies. Although we did not consistently incentivize participants, each study provides further evidence in support of our main claim. So, across both incentivized and unincentivized samples, we were unable to find reliable individual differences in exploration.[3] Second, we know from prior research that incentivization does not always meaningfully change behavior. In the case of delay discounting, participants respond similarly whether they are incentivized (by being told they would receive one of their choices at random) or not (Johnson & Bickel, 2002; Madden, Begotka, Raiff, & Kastern, 2003). So, although our conclusion cannot be definitive, we do not believe that a lack of incentives in two of the four main studies reported above can reasonably explain our pattern of results.

Ultimately, our research contributes to the broad question of how humans make the decision between "staying" or "leaving." We help shed light on a lack of measurable consistency of exploration behavior across tasks that are all reported to require trading-off between exploration and exploitation. Further, we claim that the environmental demands in each task are likely what is masking an ability to detect an individual's general tendency to explore (vs. exploit) more than other people. Our findings have direct implications for all individual differences research in exploration behavior in humans (and potentially nonhuman animals). Specifically, that individual differences should be shown to be predictive across a variety of trade-off environments as task-specific demands on strategy matters. Although we believe future work will continue to provide a more detailed answer to the question of "should I stay or should I go," we suggest the current definitive answer is "it depends."

ORCID

Ethan A. Meyers https://orcid.org/0000-0001-6171-6780
Derek J. Koehler https://orcid.org/0000-0002-3596-7882

REFERENCES

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*, 938–956. https://doi.org/10.1037/0022-3514.92.5.938

Brydges, N. M., Heathcote, R. J. P., & Braithwaite, V. A. (2008). Habitat stability and predation pressure influence learning and memory in populations of three-spined sticklebacks. *Animal Behaviour*, *75*(3), 935–942. https://doi.org/10.1016/j.anbehav.2007.08.005

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *362*(1481), 933–942. https://doi.org/10.1037/0033295X.111.4.939

Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical Science*, *4*(3), 282–289. https://doi.org/10.1214/ss/1177012493

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401. https://doi.org/10.1257/jel.40.2.351

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, *3*(10), e1701381. https://doi.org/10.1126/sciadv.1701381

Frey, R., Richter, D., Schupp, J., Hertwig, R., & Mata, R. (2020). Identifying robust correlates of risk preference: A systematic approach using specification curve analysis. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000287

Gardner, M. (1960). Scientific American Feb., Mar.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*(1), 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669. https://doi.org/10.1037/0033-295X.103.4.650

Hausmann, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment and Decision making*, *3*(3), 229–243.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523. https://doi.org/10.1016/j.tics.2009.09.004

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science*, *21*(12), 1787–1792. https://doi.org/10.1177/0956797610387443

Hills, T. T., & Pachur, T. (2012). Dynamic search and working memory in social recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 218–228. https://doi.org/10.1037/a0025161

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*(1), 46–54. https://doi.org/10.1016/j.tics.2014.10.004

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior*, *77*(2), 129–146. https://doi.org/10.1901/jeab.2002.77-129

Kirby, K. C., Marlowe, D. B., Festinger, D. S., Garvey, K. A., & la Monaca, V. (1999). Community reinforcement training for family and significant others of drug abusers: A unilateral intervention to increase treatment entry of drug users. *Drug and Alcohol Dependence*, *56*(1), 85–96. https://doi.org/10.1016/S0376-8716(99)00022-8

Lee, M., & Newell, B. R. (2011). Using hierarchical Bayesian methods to examine the tools of decision-making. *Judgment and Decision making*, *6*(8), 832–842.

Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology*, *11*(2), 139–145. https://doi.org/10.1037/1064-1297.11.2.139

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87. https://doi.org/10.1287/orsc.2.1.71

Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: Is there a reduction in exploration with aging? *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00053

Mehlhorn, K., Newell, B. R., Lee, M., Morgan, K., Braithwaite, V. A., Hausmann, D., … Gonzalez, C. (2015). Unpacking the exploration-

[3]One could argue that our incentives were not large enough to invoke motivated responding on behalf of our participants. However, we would suggest that our incentives—multiple prizes worth 250% of what the participant was receiving to complete the study that would require no additional work beyond engaging in the task—were most certainly large enough especially in the context of the recruitment platform (Amazon's Mechanical Turk).

exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, *2*(3), 191–215. https://doi.org/10.1037/dec0000033
Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(7), 2335–2343. https://doi.org/10.1523/JNEUROSCI.4156-11.2012
Navarro, D. J., & Newell, B. R., (2014). "Information versus reward in a changing world." Program of the 36th Annual Meeting of the Cognitive Science Society, 1054–1059.
Navarro, D. J., Newell, B. R., & Schulze, C. (2016). Learning and choosing in an uncertain world: An investigation of the explore-exploit dilemma in static and dynamic environments. *Cognitive Psychology*, *85*, 43–77. https://doi.org/10.1016/j.cogpsych.2016.01.001
Newell, B. R., Koehler, D. J., James, G., Rakow, T., & van Ravenzwaaij, D. (2013). Probability matching in risky choice: The interplay of feedback and strategy availability. *Memory & Cognition*, *41*(3), 329–338. https://doi.org/10.3758/s13421-012-0268-3
Newell, B. R., & Lee, M. D. (2011). The right tool for the job? Comparing an evidence accumulation and a naive strategy selection model of decision making. *Journal of Behavioral Decision Making*, *24*(5), 456–481. https://doi.org/10.1002/bdm.703
Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior*, *96*(3), 427–439. http://doi.org/10.1901/jeab.2011.96-423
Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009
Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*, *1*(11), 803–809. https://doi.org/10.1038/s41562-017-0219-x
Rakow, T., Newell, B. R., & Zougkou, K. (2010). The role of working memory in information acquisition and decision making: Lessons from the binary prediction task. *The Quarterly Journal of Experimental Psychology*, *63*(7), 1335–1360. https://doi.org/10.1080/17470210903357945
Sadeghiyeh, H., Wang, S., Alberhasky, M. R., Kyllo, H. M., Shenhav, A., & Wilson, R. C. (2020). Temporal discounting correlates with directed exploration but not with random exploration. *Scientific Reports*, *10*(1), 4020. https://doi.org/10.1038/s41598-020-60576-4
Sang, K., Todd, P. M., Goldstone, R. L., & Hills, T. T. (2018). Explore/exploit tradeoff strategies in a resource accumulation search task [revised/resubmitted to Cognitive Science].
Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology*, *83*(5), 1178–1197. https://doi.org/10.1037/0022-3514.83.5.1178
Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the "secretary problem". *Organizational Behavior and Human Decision Processes*, *69*(3), 221–236. https://doi.org/10.1006/obhd.1997.2683
Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of Behavioral Decision Making*, *13*(4), 391–411. https://doi.org/10.1002/1099-0771(200010/12)13:4%3C391::AID-BDM359%3E3.0.CO;2-I
Stein, W. E., Seale, D. A., & Rapoport, A. (2003). Analysis of heuristic solutions to the best choice problem. *European Journal of Operational Research*, *151*(1), 140–152. https://doi.org/10.1016/S0377-2217(02)00601-X
Steiner, M., Seitz, F., & Frey, R. (2019). Through the window of my mind: Mapping the cognitive processes underlying self-reported risk preference. *PsyArXiv Preprint*. https://doi.org/10.31234/osf.io/sa834
Tversky, A., & Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology*, *71*(5), 680–683. https://doi.org/10.1037/h0023123
von Helversen, B., Mata, R., Samanez-Larkin, G. R., & Wilke, A. (2018). Foraging, exploration, or search? On the (lack of) convergent validity between three behavioral paradigms. *Evolutionary Behavioral Sciences*, *12*(3), 152–162. https://doi.org/10.1037/ebs0000121
von Helversen, B., Wilke, A., Johnson, T., Schmid, G., & Klapp, B. (2011). Performance benefits of depression: Sequential decision making in a healthy sample and a clinically depressed sample. *Journal of Abnormal Psychology*, *120*, 962–968. https://doi.org/10.1037/a0023238

How to cite this article: Meyers EA, Koehler DJ. Individual differences in exploring versus exploiting and links to delay discounting. *J Behav Dec Making*. 2021;34:515–528. https://doi.org/10.1002/bdm.2226