Acta Psychologica 202 (2020) 102953

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Perceived patterns in decisions from experience and their influence on choice variability and policy diversification: A response to Ashby, Konstantinidis, & Yechiam, 2017

Ori Plonsky[1], Kinneret Teodorescu*,[1]

The Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Israel

A R T I C L E I N F O

Keywords:
Decisions from experience
Policy setting
Confidence
Pattern seeking
Wavy recency
Choice variability

A B S T R A C T

Searching for and acting upon perceived patterns of regularity is a fundamental learning process critical for adapting to changes in the environment. Yet in more artificial, static settings, in which patterns do not exist, this mechanism could interfere with choice maximization and manifest as unexplained choice variability in later trials. Recently however, Ashby et al. (2017) found that choice variability in later trials of a repeated choice setting is correlated with levels of diversification in policy tasks, in which patterns can never be exploited. They concluded that in repeated choice tasks, choice-variability in later trials is unlikely the result of following perceived patterns. Here, we demonstrate that correlations between choice variability and policy diversification can actually be the result of pattern seeking, rather than serving as evidence against it. We review evidence for the robustness of pattern seeking mechanisms in repeated choices and explain how such mechanisms could in fact create the results observed by Ashby et al. To examine our interpretation for their results, we conducted a sequential dependencies analysis of their data and find evidence that many participants behaved as if they believed trials are inter-dependent, even though they were explicitly instructed that the environment is stationary. The results of a new experiment in which sequential patterns are directly manipulated support our interpretation: Experiencing patterns affected both choice variability in later trials and policy diversification. Finally, we argue that decisions from experience tasks are a valid tool to examine the generation of preferences via fundamental learning processes.

1. Introduction

In the decisions from experience (DFE) paradigm, participants make repeated choices among options and obtain feedback following each choice. One robust finding across DFE studies is that even after a decent amount of experience with the incentive structure and toward the end of the task, participants tend to exhibit a stable amount of variability in their choices: For example, Gonzalez and Dutt (2011) report alternation rates of about 20% over trials 100–200 (reanalysis of data from Barron & Erev, 2003); Teodorescu, Amir, and Erev (2013) found in a similar binary task alternation rates of 12–17% over trials 51–100; and Weiss-Cohen, Konstantinidis, Speekenbrink, and Harvey (2018) found in the Iowa Gambling Task (IGT), which includes 4 alternatives, alternation rates of about 30% over trials 80–100. From an economic point of view, this behavior is suboptimal in the sense that after the best option has been identified, it should be selected in all trials. Why do participants keep alternating between options?

Ashby, Konstantinidis, and Yechiam (2017), henceforth referred to as AKY, recently suggested that the observed variability reflects the fact that decision makers cannot be certain that they identified the best option, and so their choice variability is meant to avoid putting all eggs in one basket. That is, given the uncertainty of the environment, decision makers do not feel confident in their understanding of the incentive structure, and therefore diversify their choices to hedge their bets. AKY contrast this under-confidence explanation with an alternative explanation stating that agents tend to search for patterns in the sequence of generated outcomes. Although in most DFE paradigms trials are all independent of one another, and thus no true patterns exist, agents may nonetheless encounter spurious patterns and follow them. For example, in a binary choice task, an agent may think that one option provides the best outcome every third trial, and the alternative provides the best outcome in all other trials. Such agent would then

* Corresponding author.
E-mail address: kinnerett@technion.ac.il (K. Teodorescu).
[1] Both authors contributed to this work equally.

https://doi.org/10.1016/j.actpsy.2019.102953
Received 14 February 2019; Received in revised form 3 November 2019; Accepted 4 November 2019
0001-6918/

exhibit variability in choice merely by attempting to exploit this pattern.

To compare these two explanations, AKY performed two experiments in which they let decision makers face first a standard DFE task in which agents repeatedly choose between unmarked options and then a policy task in which agents set in advance a distribution of choices between the same unmarked options. Importantly, in the policy task, a random device set which option is selected in each round (contingent on adhering to the prescribed policy), and participants could not prescribe the order by which their chosen distribution is allocated within the task. Therefore, in the policy task, participants could not have exploited any perceived patterns they may have thought existed. The results of the experiments showed high correlations between the individual choice proportions in the two tasks. AKY took their results as evidence that "variation in choice is related to a lack of confidence in knowing which option is best, rather than being driven primarily by exploration or failures to recognize choice independence" (Ashby et al., 2017, p. 65).

Here, we argue that a pattern search account remains a good explanation for choice-variability in DFE tasks. First, we shortly review the evidence for pattern search within DFE tasks. Second, we discuss AKY experiments in detail and raise an alternative, pattern seeking, interpretation to their findings. To examine our alternative interpretation, we conducted a sequential dependencies analysis of their data and find evidence that many participants failed to behave in a manner consistent with choice independence, even though the incructions explicitly noted that the environment is stationary. Finally, we demonstrate that AKY's main experimental results can be replicated also when we *know* agents follow sequential patterns in the DFE task. That is, we show that AKY's results cannot rule out a pattern search explanation for the variability in choice phenomenon. In so doing, we also demonstrate, to the best of our knowledge for the first time, that decision makers tend to diversify policy portfolios in accordance with a perceived sequential pattern even though the pattern can never be exploited. We take our results as further evidence for the claim that decision makers find it very hard to resist the urge to seek and exploit environmental patterns.

1.1. Searching for patterns

We live in a dynamic environment characterized by natural regularities (e.g. sun and moon cycles, seasons, etc.). It is therefore not surprising that humans have evolved to search for and recognize patterns of regularity. Searching for patterns is a critical mechanism for adapting to changes in the environment, yet it could harm performance in more artificial settings in which outcomes are randomly generated (with no specific pattern) such as gambling (Foster & Kokko, 2008; Beck & Forstmeier, 2007; Yu & Cohen, 2009). For example, Gaissmaier and Schooler (2008) found that in a probability learning task, the suboptimal result of probability matching could be attributed to pattern search in > 50% of the cases. Those 'pattern seekers' seemed to perform poorly in a classical version of the task in which patterns do not exist, but performed best in a modified version of the task that included patterns.

In a computational investigation of binary DFE tasks, Plonsky, Teodorescu, and Erev (2015) demonstrate both sides of the pattern-search coin. Whereas searching and following patterns is found to be generally near optimal in dynamic tasks (when the setting can change from trial to trial), when the environment is static (i.e., the probability to receive a given outcome is constant across trials), the same process leads to underweighting of rare events (Barron & Erev, 2003) and deviations from maximization. Moreover, they showed that in static settings that include rare events, this mechanism predicts a non-trivial *wavy recency effect*, a prediction that is corroborated by experimental data (and later replicated by Plonsky & Erev, 2017; Szollosi, Liang, Konstantinidis, Donkin, & Newell, 2019). Specifically, they found that immediately after observing an outcome, participants' choice patterns

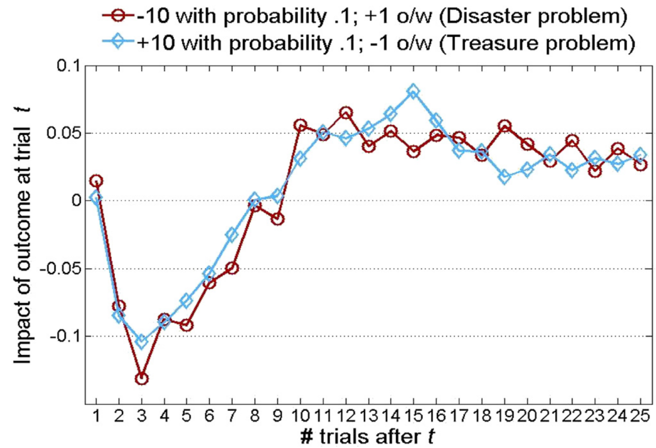

Fig. 1. The wavy recency effect in DFE tasks. Participants chose 100 times between a safe option generating 0 with certainty and a risky option with the payoff distribution given in the legend. The plot shows the impact of risky option outcomes generated at trial t on choice at trial $t + n$, which is the difference between the probability of choosing the risky option at trial $t + n$ when the outcome at trial t was positive and the probability of choosing the risky option at trial $t + n$ when the outcome at trial t was negative.
Data analyzed is originally presented in Plonsky et al. (2015). Figure is reprinted from Plonsky and Erev (2017).

reflect sensitivity to that outcome, but this sensitivity quickly drops, reaching a nadir three trials after the outcome occurs, and then it gradually increases, peaking 12–15 trials after the outcome occurs. The effect then slowly diminishes. Fig. 1 demonstrates this wavy pattern. It shows the mean aggregate impact of an outcome generated at trial t by a risky option that includes a rare event (in a binary choice experiment, with an alternative of 0 for sure) on the choices in trials $t + 1$ through $t + 25$. The impact of an outcome at trial t on a future choice is defined as the difference between the rate of risky choice contingent on the outcome at trial t being positive (better than the alternative) and the rate of risky choice contingent on that outcome being negative (worse than the alternative). This curve provides two indications that participants fail to acknowledge the different trials in these experiments are independent. First, a negative impact at trial $t + n$ (e.g. for $n = 3$) suggests decision makers behave as if the outcome at trial t is negatively correlated with the outcome expected at $t + n$: They are more likely to choose the risky option if the outcome at trial t was negative than if it was positive and vice versa. Second, a curve that increases with n (e.g. between trials $t + 3$ and $t + 12$) suggests decision makers behave as if they believe outcomes that are generated later are more likely to be similar (or relevant) to the outcome at trial t than outcomes that are provided earlier. Such patterns cannot be captured with most classical learning models that assume positive recency (more recent outcomes have a greater effect on choice). Yet, as Plonsky et al. (2015) show, they can be captured by a model that is based on pattern search.

Similar wavy recency effects were also observed in more complex partial feedback DFE tasks, and importantly, also in repeated decisions between fully described options (Plonsky & Erev, 2017). That is, it appears that even when agents are given the information that trials are independent, they tend to behave as if they search for patterns. One prominent example comes from a study in which participants were given a die to role on their own and were asked to predict the outcome of the next die role (Peterson & Ulehla, 1965). The results showed participants still exhibited sequential dependencies (see Beach & Swensson, 1967) and significant choice variability even in the last 100 out of 300 trials of experience. Specifically, only 24% of the participants did not diversify their predictions in these 100 trials.

To sum, considerable evidence suggests that in DFE tasks, participants search and try to follow (spurious) patterns in the sequence of outcomes. Such mechanism can then explain the choice variability observed even after many trials. If participants fail to acknowledge that each trial is independent of the others, there is no reason for them to settle on a single option.

1.2. AKY's experiments

AKY argue that while pattern search is a feasible mechanism for the way agents behave in DFE tasks, it is unlikely to explain the observed choice variability. They perform two experiments to substantiate their claim. In their experiments, participants play 100 trials in a DFE task that includes both safe and risky options. Participants receive trial-by-trial feedback regarding the outcome generated by the chosen alternative. In (their) Experiment 2, participants in addition also received feedback regarding the outcomes generated by the unchosen alternatives, and, importantly, were told that the options would not change in any way in the course of the study. After finishing the DFE task, participants were asked to play 100 additional trials of the same choice problem, but this time specify *in advance* how they would like to distribute their choices across the options. Importantly, they were not able to affect the order of choices within their policy allocation of the 100 trials. Their chosen policy (the number of trials allocated to each option) was then carried out by the computer and only the final result of this process was accessible to the participant. Moreover, in Experiment 2, after completing both the choice and the policy tasks, participants were also asked a few follow-up questions, including which option they think paid up the most on average, how confident they are in that assessment, and what they think should be the optimal strategy in either of the two tasks.

AKY concluded that choice variance in later trials reflects under-confidence in their knowledge of which option is better, rather than being driven by pattern seeking. Their conclusion is based mainly on a strong correlation between choices made in the trial-by-trial DFE task and allocations made in the policy task. AKY reasoned that since in the policy task patterns can never be exploited, an observed variability in their choice in that task (diversified policies) must reflect under-confidence in their knowledge of which option was best. The strong correlation with the trial-by-trial task thus led them to conclude that in that task as well variability in choice reflects under-confidence. As additional supporting evidence to their claim, AKY find a (weak) negative correlation between participants' reported under-confidence in knowing which option provided a higher payoff on average and the number of options they select in the final trials of the DFE experiment (taken as a measure of variability in choice).[2] They also find that only a small minority of participants reports that the best strategy in the DFE task is to choose according to the underlying patterns.

1.3. Alternative interpretation of AKY's findings

We believe AKY's main finding, a correlation between choices in the DFE task and choices in the policy task, can also emerge if most participants in the DFE task are pattern seekers. To see why, let us focus on AKY's binary DFE task (i.e. trial-by-trial choice between one safe and one risky option) and assume that the participants facing this task do indeed search for and try to follow-up on patterns in sequential outcomes. Trials are independent, and so any patterns participants think they spot are spurious. In particular, two agents playing the same task may each think they found a different pattern. For example, Amy may believe the safe option provides the best outcome in 4 of 5 trials whereas Brad may believe the safe option provides the best outcome in just 3 of 5 trials. Next, participants face AKY's policy task. What should they do? Normatively, since they cannot exploit the patterns they "found" in the DFE task, they should allocate all 100 trials to the option that provides the higher expected value (say, the safe option). Yet, we hypothesize that participants who think they identified a pattern in the environment are not likely to follow this normative prediction. Instead, they will aim to "exploit" the pattern even in the policy task. Possible reasons for such behavior are that they feel they should not deny themselves the possibility that the computer will use their allocation according to the pattern they believe exists or that they want to signal that they "solved the puzzle" and identified the underlying structure of the environment. Regardless of the reason, trying to "exploit" the pattern in the policy task would lead to allocations that reflect the patterns they think they found. For example, Amy may allocate 80 trials to the safe option whereas Brad would allocate only 60 trials to the safe option. The procedure we conjecture here would then lead to the correlation AKY find between the observed behavior in the two tasks, although participants may still be pattern seekers.

Moreover, we also believe AKY's finding of a weak negative correlation between the reported confidence in identifying the option that paid out the most on average and the number of options selected at the end of the experiment can be a result of the same mechanism. Again, think of a binary DFE task and assume that most participants search for patterns, however some of them do not find any. Those who think they found a pattern will tend to select both options (because either is "better" in different trials), while those who did not search for or did not find any patterns will be more likely to settle on a single option. Moreover, when asked which option is better on average, those who believe patterns exist will be less likely to have a good answer than those who do not, either because they did not follow payoff averages at all or because they will find this question confusing (both options are better, but on different trials). Therefore, participants who believe a pattern exists presumably will show on average more variability in choice, and will report lower confidence in their answer to the "which is better" question.[3]

2. Investigation of the pattern seeking account

2.1. A sequential dependencies analysis of AKY's data

We cannot know what participants in AKY experiments had in mind, and so a direct examination of the alternative interpretation presented above is impossible. However, we can analyze data from these experiments and examine whether participants choose in the DFE task in a manner consistent with the notion that trials are dependent on one another. If they do, it would lend support to the claim that the variability observed in these tasks is at least partially due to the tendency to search and follow spurious patterns. Our investigation of AKY's data[4] focuses on the simplest of all conditions they ran, a binary choice task with full feedback concerning both outcomes in each trial. Two reasons underlie this focus. First, identifying behavior consistent with choice dependence is much simpler in this condition. Second, AKY's main finding of correlation between choices in the DFE task and choices in the policy task is strongest in that condition[5] Moreover, we focus on the

[2] In a two-armed bandit tasks, Hertz, Bahrami, and Keramati (2018) also report correlations between choice variability and confidence ratings.

[3] As for the finding that only few participants report the best strategy is to look for patterns, we speculate this could be the result of a unique design choice AKY made, namely that they explicitly notified participants in the instructions that the two options do not change in any way over the course of the study. This design choice might signal to the participants that they should not answer that the best strategy is to look for patterns, even if this is indeed what they do. That is, we speculate participants may have taken this question to be a sort of test checking their attention in reading through instructions. Moreover, some participants may be looking for patterns implicitly rather than explicitly and so their stated preferred strategy and their actual strategy may not align. Admittedly though, all these are conjectures that we do not – and cannot – test here.

[4] We thank AKY for sharing their data with us.

[5] AKY's experiments included more complex conditions in which there were

second half of the DFE task, in which participants have already had a chance to learn which option is better and so alternations are less likely to be a result of exploration.[6]

To check for evidence for behavior consistent with inter-trial dependence, we compute the impact of an outcome generated by the Risky option (which provided 100 with probability 0.2 and 30 otherwise) at trial t on the choices made in subsequent trials. That is, we compute, for each participant (and for $n = 1, \ldots 15$), the difference between the risky rate at trial $t + n$ contingent on the Risky option providing a high (100) outcome at trial t and the risky rate at trial $t + n$ contingent on the Risky option providing a low (30) outcome at trial t. As mentioned above, a negative impact at trial $t + n$ implies that the participant behaves as if the outcome in that trial is likely to be negatively correlated with the outcome at trial t; an increasing impact (with an increase in n) implies that the participant behaves as if outcomes arriving further away from trial t are likely to be more similar to the outcome provided at trial t than outcomes arriving soon after trial t. Both types of behaviors violate the common assumption of positive recency but can be captured by assuming inter-trial dependence.

Fig. 2 shows the impact curve aggregated over the participants in the relevant condition.[7] Consistent with a pattern search account, the impact shows a wavy pattern: It starts strongly positive, quickly drops to a nadir at $t + 3$, then increases until $t + 6$ and then the effect diminishes.[8] Notably, however, the curve is far "less wavy" then the curve from Fig. 1. For example, it receives its maximum at $t + 1$ rather than much later, and it does not change as much after $t + 1$. We believe this reduced waviness is due to the payoff distributions in the current setting. The wavy recency effect is particularly pronounced in settings with rare events in which the options have similar expected values. In the current setting, the least frequent event occurs with probability 0.2, which is not especially rare, and the difference between the two options' expected values is large. Indeed, the model CAT (contingent average and trends) that Plonsky et al. (2015) developed to capture the wavy recency effect predicts that in this setting the wavy effect will be quite weak.

Nevertheless, the impact curve in Fig. 2 does show a wavy pattern, and specifically, it shows both indications for violation of choice independence discussed above: negative impact at $t + 3$ and a gradual increase in the impact after $t + 3$. The difference between zero and the negative mean impact at trial $t + 3$, $M = -0.0205$, 95% bootstrapped CI $[-0.045, 0.004]$, is only marginally significant, $p = .097$. However, the evidence for the gradual increase after $t + 3$, which seems to end at $t + 6$, appears somewhat stronger:, the mean difference between the impact at $t + 6$ and the impact at $t + 3$, $M = 0.0398$, 95% bootstrapped CI [0.005, 0.075], is statistically significant, $p = .028$.[9] Moreover,

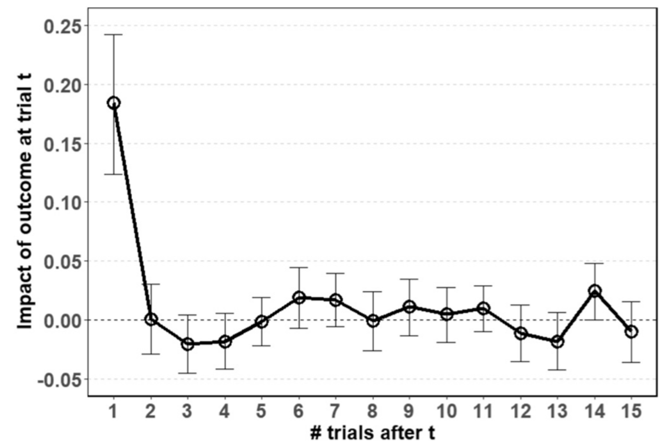

Fig. 2. The impact of risky option outcomes on subsequent choice in AKY's Study 2, binary condition. Impact of an outcome at trial t on choice at trial $t + n$ is the difference between the probability of choosing the risky option at trial $t + n$ when the outcome at trial t was a relative gain (higher than the outcomes from the safe option) and the probability of choosing the risky option at trial $t + n$ when the outcome at trial t was a relative loss (lower than the outcome from the safe option). Data is averaged across 150 participants facing the binary condition in AKY's study 2 and is analyzed for the second half of that experiment. Error bars indicate 95% bootstrapped CI.

among the group of individuals who showed any variability in their choice in the last 50 trials (109 of the participants), 68.8% either had a negative impact at trial $t + 3$ or had a positive difference between the impact at trial $t + 6$ and trial $t + 3$ (or both). That is, the majority of the participants that exhibited variability in choice in the DFE task ($p < .001$ in a sign test) showed some indication of behavior consistent with choice dependence, as predicted by a pattern search account.

Whereas for 31.2% of participants that exhibited choice variability we do not have evidence for violation of choice dependence, this does not necessarily mean they did not search and follow sequential patterns. In the current problem, following perceived patterns could result in a monotonically decreasing and strictly positive impact curve. The reasons for this go beyond the scope of this paper but are related both to the large difference between the options' expected values and to the relatively high probability for the infrequent event. Indeed, for a certain choice of parameters, the model CAT (Plonsky et al., 2015) mentioned above, which is heavily based on pattern seeking, predicts in this problem a positive and monotonically decreasing impact curve.[10]

2.1.1. Discussion

The setting used in AKY experiments poses a significant challenge to the effort to find evidence for sequential dependencies in choice. Nonetheless, although the evidence is not too strong, a sequential dependencies analysis of the simplest condition used in AKY, hints that on average, participants behaved as if they perceived trials to be

(*footnote continued*)
more options to choose from and/or there was no feedback concerning the forgone outcome. In these more complex situations, interactions of the attempt to search and exploit patterns with other behavioral phenomena related to the complex nature of the setting may emerge (e.g., Plonsky & Erev, 2017). It is outside the scope of this paper to investigate such patterns. Moreover, with respect to the main findings of AKY, differences between the different conditions were minor, and if anything, stronger in the simpler condition we analyze.

[6] We do not use a shorter time horizon because, in order to get reasonably stable estimates, our analysis requires that each participant will have observed each outcome generated by the Risky option at least a few times.

[7] While there were 151 such participants, this curve aggregates 150 participants. The curve for one participant could not be computed because the less frequent event was not observed by that participant in the trials analyzed. We therefore excluded this participant from all analyses.

[8] The nadir in the impact curve is almost always obtained at $t + 3$, but the end point for the gradual increase in the impact changes from dataset to dataset. It is normally obtained sooner for settings with "more frequent" rare events than settings with "more rare" rare events, so the fact it is obtained at $t + 6$ here is not surprising.

[9] The statistical properties of the impact at trial $t + n$ are unknown so to

(*footnote continued*)
estimate the uncertainty associated with the statistics, we employ a non-parametric bootstrap procedure, using package boot in R (Canty & Ripley, 2019), with a "basic" bootstrap interval. *P-values* are computed using a shifted bootstrapped distribution such that it is centered around zero.

[10] Broadly, CAT includes two independent mechanisms: a "contingent averages" mechanism which is designed to capture global patterns in the sequences of outcomes (i.e. sequences that repeat themselves), and a "trend" mechanism which is designed to capture local patterns of recent outcomes that appear to increase or decrease. Whereas the latter trend mechanism predicts a weak wavy effect similar to the one in Fig. 2, the former contingent averages mechanism predicts a monotonically decreasing positive impact in this setting. Therefore, for agents who only follow global sequences, we do not expect to see any wavy recency effect according to this model.

dependent of one another. This is despite explicit instructions to the contrary. As far as we know, this analysis presents the first evidence for violations of choice independence in DFE tasks in which participants are explicitly instructed that the setting is static. Moreover, our finding is in stark contrast to the fact very few (11%) participants reported they believed searching for patterns is the best strategy to use in this setting. This analysis provides evidence for the robustness of pattern seeking over instructions that patterns do not exist and suggests that explicit retrospective reports might not be a good measurement for the actual strategies people use.

Yet, the evidence in the current setting is relatively weak, particularly regarding the relationship between choice variability, confidence ratings, and policy diversification. Indeed, here we did not analyze the policy diversification at all. Because the current design does not allow us to determine which participants truly followed perceived patterns, and which pattern each participant followed, our stronger argument that perceived patterns can cause the observed correlations remains only a plausible speculation. To directly examine the effects of perceived patterns on choice variability, confidence ratings, and policy diversification, we therefore designed a simple experiment, in which the patterns participants follow are directly manipulated.

2.2. Experimental manipulation of patterns

The new experiment incorporates the main procedure of AKY's experiments in which a DFE task is followed by a policy task and confidence questions. However, in our new DFE task, the underlying payoff structure is dynamic rather than static, such that it actually includes a simple pattern of regularity in the outcomes of the risky option. Importantly, each participant faced one of two possible dynamic settings, each involving a different underlying sequential pattern. One pattern implied relatively many choices of the risky option and the other relatively few choices of the risky option. Assuming the majority of the participants will pick up on the simple pattern they encounter (as is reported in the literature for over 6 decades; see Galanter & Smith, 1958; Lee, 1971; Gaissmaier & Schooler, 2008), we can compare the policies set by each group in the second task and see if they correlate with the choice rates from the DFE task. If they do, it implies that the policies participants set are also influenced by the pattern they found in the DFE task even though the pattern cannot be exploited.

2.2.1. Method

2.2.1.1. Participants. 207 Prolific workers completed the current study[11] (56% men, $M_{age} = 32$). They were paid a show-up fee of £0.8 and a bonus payment according to one randomly selected trial (that could be drawn from both the DFE and the policy tasks). On average, participants completed the study in 10 min and earned a bonus of £0.27.

2.2.1.2. Procedure. Participants faced two tasks, first a DFE task and then a policy task. In the DFE task, participants were presented with two unmarked keys for 100 trials on the computer screen and received full feedback (both obtained and forgone outcomes) after each trial. Unbeknownst to participants, the underlying payoff structure included a simple pattern: As shown in Table 1, in one condition (Problem 1) the safe key was better every three trials and in the other condition (Problem 2) the risky key was better every three trials. This pattern of three outcomes repeated itself over and over. The payoff in very first trial was randomly selected for each participant from the implied distribution in each problem.

We chose these simple patterns, to increase the chances that most participants will learn to follow them. Conditions were manipulated between participants, so that each participant played only one problem. Notice that in both conditions the safe option yields 25 in each trial, the expected value of always selecting the risky option is around 24 and the expected value of playing the optimal strategy (according to the specific pattern in each condition) is 31.67.

After completion of the DFE task, participants were asked to set up a policy for the next 100 trials (played by the computer). It was clarified that their choice is between the same keys as in the DFE task, and that they cannot impact the order by which their chosen distribution will be played. They were then asked which option they think is the better option (global-better) and rated their confidence in their answer (global-better confidence). We also added a question at the end in which we provided participants with the outcomes of the risky option in the last 3 trials and asked them which option will provide the better outcome in the next trial (local-better) following by confidence rating in their answer (local-better confidence).

Hypotheses:

H1. Most participants will search and find the patterns in the DFE task. This will translate to choice variance also toward the end of the task. Specifically, we predict that toward the end of the task (when the pattern is discovered) participants will choose the risky option in about 67% of the trials in Problem 1 and in about 33% of the trials in Problem 2.

H2. Most participants will diversify their choices in the policy task as well, even though no patterns can be exploited in this task.

H3. Allocation of risky choices in policies will be higher in Problem 1 than in Problem 2. Moreover, the choice rate in the DFE task will positively correlate with the policy allocations in the policy task. This correlation will be driven primarily by participants who diversify their choices in the policy task indicating that diversification in the policy task is related to the patterns found in the DFE task.

H4. Confidence in the answer to the global-better question will be weakly negatively correlated with choice variability in the DFE task.

H5. Confidence in the answer to the global-better question will be lower on average than the confidence in the answer to the local-better question, indicating that in the presence of patterns, the global-better question is more difficult to answer.

2.2.2. Results

2.2.2.1. DFE task. On average, participants selected the risky option 59% of the trials in Problem 1 and only 37% of the trials in Problem 2, a significant difference, $t(162.94) = 13.21$, 95% CI [0.18, 0.25], $p < .001$, despite the fact that the expected value of the two options were very similar in the two problems (in fact, the expected value of the risky option was slightly smaller in Problem 1). To explicitly examine choice variability toward the end of the task, we calculated the risk rates and alternation rates in the last 16 trials.[12] The alternation rates were 62% and 64% in the two problems, demonstrating similar and significant choice variability in later trials. Risk rates in the last 16 trials were 62% and 34% in Problem 1 and Problem 2 respectively, a significant difference, $t(173.3) = 15.53$, 95% CI [0.24, 0.31], $p < .001$. These risk rates are very close to those expected under the assumption that participants searched for and discovered the simple pattern in each condition (H1).

To explicitly examine whether participants learned to follow the pattern in each condition, we calculated the rate of choices congruent

[11] 300 workers entered the link to the experiment, out of which 93 were not allowed to proceed because they were using a mobile phone (which was not allowed) and/or because the failed the attention check (in the instructions participants were asked to type the letters "NOYB" in the comment box presented at the end of the instructions). The analysis included all other 207 participants.

[12] This number was selected following AKY analyses choices.

Table 1
Problems faced by participants in the experiment.

Problem	Payoff structure		EV(risky)	EV (optimal)
	Safe	Risky		
1	Always 25	1 if the last two payoffs were 35; 35 otherwise (i.e. 1, 35, 35, 1, 35, 35…)	23.67	31.67 (optimal strategy: S,R,R,S,R,R, …)
2	Always 25	45 if the last two payoffs were 14; 14 otherwise (i.e. 14, 14, 45, 14, 14, 45…)	24	31.67 (optimal strategy: S,S,R,S,S,R, …)

Note. EV = expected value. S = safe, R = risky.

with the pattern in the last 16 trials of the task. Results show that participants' choices toward the end of the task were aligned with the optimal pattern strategy in 93% and 95% of the time in problem 1 and problem 2 respectively (difference insignificant, $p = .24$). Moreover, we classified participants who made the correct choice according to the pattern they face in at least 15 of the last 16 trials as pattern-followers. Results show that 85% of participants facing Problem 1 and 90% of participants facing Problem 2 were classified as pattern followers, showing that most participants indeed learn to follow the pattern of outcomes in each condition.

2.2.2.2. Policy setting and its relation to the DFE task. Recall that the correct (maximizing) answer to this question is to select the safe key in all choices, because the expected value of the safe key is larger in both conditions, and patterns cannot be exploited. However, > 75% of participants chose to diversify their policies (76% in Problem 1 and 80% in Problem 2; H2). More importantly, participants allocated significantly more choices to the risky option in Problem 1 than in Problem 2 (49.7% vs. 35.3%, $t(193.76) = 3.77$, 95% CI [0.07, 0.22], $p < .001$), in accordance with their selection of the risky option in the DFE task (H3). The correlation between the risk rate in the DFE task and the allocation to the risky option in the policy task was 0.43, 95% CI [0.31, 0.54], $p < .001$. If limiting the analysis to participants who diversified their allocation, this correlation increases to 0.61, 95% CI [0.51, 0.70], $p < .001$.

2.2.2.3. Learning, confidence ratings, and their relation to the DFE task. Most participants (85%) declared they thought there's a globally better option. Among these participants, most participants who faced Problem 2 (72%) but a minority of participants who faced Problem 1 (40%) chose the maximizing safe option is the globally better option. This difference may reflect the fact that in Problem 2, but not in Problem 1, the safe option was also better than the alternative in most trials (as dictated by the induced sequential patterns). While choice of the globally better option differed between problems, the average global confidence ratings were similar and moderate in both conditions (77% and 72% in Problem 1 and 2, respectively). The correlation between participants choice variability in the DFE task and global confidence ratings was negative and weak ($r = -0.12$, $p = .084$), as per H4. These results are very similar to those observed in AKY's study (e.g. the correlation in their study was also only marginally significant).

In our experiment, we also asked participants which of the two options will generate a higher payoff in the next trial, given a particular sequence of three outcomes, and rate their confidence in their answer. Again, most participants (92% in Problem 1 and 91% in Problem 2) stated it is possible to know which option will generate a higher payoff, and among those the majority (95% in Problem 1 and 79% in Problem 2) gave the correct answer. More importantly, in both conditions, the local confidence ratings participants gave for their answers were significantly higher than the global confidence ratings they provided (Problem 1: 86% vs. 77%, $t(199.99) = 3.77$, $p < .001$; Problem 2: 87% vs. 72%, $t(197.64) = 5.93$, $p < .001$), as per H5.

2.2.3. Discussion

The most striking result in our experiment is that participants' policies were diversified in line with their experienced pattern, even though these patterns could not be exploited. This result challenges AKY's conclusion that the observed correlation between choice variation and policy diversification suggests pattern seeking processes are unlikely drivers of the observed behavior. That is, our results provide evidence that searching for and acting upon perceived patterns affects both variation in on-going decisions and diversification of policies. In addition, the significant difference between participants' global and local confidence ratings suggests that global confidence ratings might not be a valid tool to examine participants' confidence in understanding the payoff structure. Had participants believed there is a pattern in outcomes occurrence (a false belief in AKY experiments and a correct belief in our experiment), they should not be confident which option is better overall (global confidence ratings), even if they think they completely understand the underlying payoff structure of the task. In AKY's experiment, stronger belief about the (false) existence of patterns could therefore have been translated to more choice-variance, less global confidence, and more policy diversification, creating correlations between these variables. In other words, our results lend support to the conjecture that these correlations can be caused by pattern seeking rather than being driven by under-confidence.

3. General discussion

Previous studies demonstrate that searching for patterns of regularity in the environment is an adaptive learning process facilitating efficient decision making in dynamic environments. Yet in more artificial static environments, where outcomes are randomly generated, this process can lead to deviations from optimality. Searching for patterns is such a fundamental learning process, that it seems to underlie people's behavior even when it leads to inferior performance: Previously described phenomena includes probability matching (Gaissmaier & Schooler, 2008), underweighting of rare events and the wavy recency effect (Plonsky et al., 2015). The results of our new experiment provide additional evidence: In policy setting, when one needs to allocate future choices between options and any allocation clearly cannot exploit on-going regularities, people nevertheless diversify their choices in accordance with their perceived patterns of regularity.

Our DFE task (played before setting up a policy) included simple patterns of outcomes, thus the belief that patterns exist was correct and choice variability in this task was objectively justified. However, even in static DFE tasks in which outcomes are randomly generated, some patterns of outcomes could occur accidently, strengthening the (now false) belief that objective patterns exist. In turn, choice variability emerges. Our results suggest that those who perceive patterns to exist, will exhibit choice variability, report only moderate confidence levels in knowing which option is better overall, and will diversify their policy in congruence with the perceived pattern. Therefore, perception of patterns is expected to result in correlations between choice-variability in a DFE task, confidence levels in which option is better overall, and policy diversification, even when the DFE task does not include patterns. AKY found these correlations, yet interpreted them as resulting from under-confidence, and as evidence against perceptions of patterns. Yet, a re-analysis of their data suggests that the majority of participants who diversify their choices behaved as if different trials were interdependent (despite explicit instructions to the contrary). Such perceived dependency between choices can emerge if decision makers search for and try to exploit patterns in the environment. Together with the results of our new experiment, this paper thus demonstrates that, rather than serving as evidence against perceptions of patterns, AKY results may actually be expected under the assumption that decision makers search for patterns. Moreover, since AKY explicitly told participants that the setting is static, our analysis might also demonstrate, to the best of our knowledge for the first time in DFE tasks, that participants tend to behave as if they are searching for sequential patterns, even when explicitly told these do not exist.

Nevertheless, the evidence we find for pattern seeking in AKY experiments is not very strong. This is likely partly due to the challenge in identifying such behavior when the rare event is not very rare and when the expected values of the available options are not very close. But it may also be partly because of the explicit instructions participants were given. Future studies should compare the relative importance of the different factors influencing the tendency to search for patterns when there are none.

We do not wish to claim that pattern search is the only mechanism that drives choice (or choice variability). Indeed, we believe the robust tendency of decision makers to search for and act upon perceived patterns in DFE tasks is implied by a more general tendency of people (and other animals) to do whatever worked best in similar situations in the past, an interpretation of Skinner's (1953) contingencies of reinforcement approach. Under this interpretation, perceiving patterns is merely an attempt to identify the situations that are most similar to the current decision (Plonsky et al., 2015). Naturally, under different conditions, the underlying similarity function may change, and so choice may be driven by mechanisms that are not directly related to sequential patterns. For example, direct manipulations of the agents' beliefs regarding the underlying structure of the environment, as well as external manipulation of switching costs can lead to different behaviors (Szollosi et al., 2019(Erev and Haruvy, 2013)(Ashby & Teodorescu, 2019)) since they imply changes to the underlying similarity function.

If searching for and acting upon perceived patterns underlies participants' behavior in DFE tasks, one could argue that DFE tasks are not a valid tool to examine preferences (Ashby et al., 2017). We would like to propose a more optimistic view. We believe evidence for such mechanisms in fact supports the validity of DFE tasks in revealing on-going preferences in the presence of feedback. We believe that DFE research enables a closer look at the basic learning processes that underlie many of our daily decisions, thereby connecting economic factors (e.g. monetary rewards, risk attitudes) with earlier research of learning in psychology. Perhaps the most basic learning principle is "the law of effect" (Thorndike, 1898) according to which good outcomes increase the likelihood of the reinforced behavior and bad outcomes decrease it. Indeed, participants facing a repeated decision making task in which one option is deterministically better than the other (e.g. one option always yields 11 and the other option always yields 10) show fast learning and increased preference toward the better paying option (e.g., Haruvy & Erev, 2002; (Erev and Haruvy, 2013)). Later principal learning phenomena such as Partial Reinforcement Extinction Effect (Humphreys, 1939), Learned Helplessness (Maier & Seligman, 1976) and the role of surprise in learning processes (Kamin, 1969; Rescorla & Wagner, 1972) were also replicated by Erev and colleagues in DFE settings (Hochman & Erev, 2013; Teodorescu & Erev, 2014; Nevo & Erev, 2012, respectively), supporting the validity of DFE tasks as an appropriate tool to examine fundamental learning processes. Our results provide additional evidence that DFE tasks are a valid tool to investigate the generation of preferences via learning processes, showing that one of the most fundamental learning processes, seeking for and following environmental patterns, also underlies behavior in DFE tasks.

Funding

This research was supported by the I-CORE program of the Planning and Budgeting Committee and the Israel Science Foundation (grant no. 1821/12).

References

Ashby, N. J., Konstantinidis, E., & Yechiam, E. (2017). Choice in experiential learning: True preferences or experimental artifacts? *Acta Psychologica, 174*, 59–67.
Ashby, N., & Teodorescu, K. (2019). The effect of switching costs on choice-inertia and its consequences. *PloS one, 14*(3), e0214098.
Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making, 16*(3), 215–233.
Beach, L. R., & Swensson, R. G. (1967). Instructions about randomness and run dependency in two-choice learning. *Journal of Experimental Psychology, 75*(2), 279.
Beck, J., & Forstmeier, W. (2007). Superstition and belief as inevitable by-products of an adaptive learning strategy. *Human Nature, 18*(1), 35–46.
Canty, A., & Ripley, B. (2019). boot: Bootstrap R (S-Plus) functions. *R Package Version, 1*, 3–23.
Erev, I., & Haruvy, E. (2013). Learning and the economics of small decisions. *The handbook of experimental economics*. *Vol. 2* (9781400883172-011).
Foster, K. R., & Kokko, H. (2008). The evolution of superstitious and superstition-like behaviour. *Proceedings of the Royal Society B: Biological Sciences, 276*(1654), 31–37.
Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition, 109*(3), 416–422.
Galanter, E. H., & Smith, W. A. (1958). Some experiments on a simple thought-problem. *The American Journal of Psychology, 71*(2), 359–366.
Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118*(4), 523.
Haruvy, E., & Erev, I. (2002). On the application and interpretation of learning models. *Experimental business research* (pp. 285–300). Boston, MA: Springer.
Hertz, U., Bahrami, B., & Keramati, M. (2018). Stochastic satisficing account of confidence in uncertain value-based decisions. *PLoS One, 13*(4), e0195399.
Hochman, G., & Erev, I. (2013). The partial-reinforcement extinction effect and the contingent-sampling hypothesis. *Psychonomic Bulletin & Review, 20*(6), 1336–1342.
Humphreys, L. G. (1939). The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *Journal of Experimental Psychology. General, 25*, 141–158.
Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell, & R. M. Church (Eds.). *Punishment and aversive behavior* (pp. 279–296). New York: Appleton-Century-Crofts.
Lee, W. (1971). *Decision theory and human behavior.* New York, NY: John Wiley & Sons.
Maier, S. F., & Seligman, M. E. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General, 105*(1), 3.
Nevo, I., & Erev, I. (2012). On surprise, change, and the effect of recent outcomes. *Frontiers in Psychology, 3*, 24.
Peterson, C. R., & Ulehla, Z. J. (1965). Sequential patterns and maximizing. *Journal of Experimental Psychology, 69*(1), 1.
Plonsky, O., & Erev, I. (2017). Learning in settings with partial feedback and the wavy recency effect of rare events. *Cognitive Psychology, 93*, 18–43.
Plonsky, O., Teodorescu, K., & Erev, I. (2015). Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychological Review, 122*(4), 621.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*. *Vol. 2*. *Classical conditioning II: Current research and theory* (pp. 64–99).
Skinner, B. F. (1953). *Science and human behavior.* New York, NY: Free Press.
April 29Szollosi, A., Liang, G., Konstantinidis, E., Donkin, C., & Newell, B. R (2019). Simultaneous Underweighting and Overestimation of Rare Events: Unpacking a Paradox. *Journal of Experimental Psychology: General. Advance online publication.*. https://doi.org/10.1037/xge0000603.
Teodorescu, K., Amir, M., & Erev, I. (2013). The experience–description gap and the role of the inter decision interval. *Progress in brain research*. *Vol. 202*. *Progress in brain research* (pp. 99–115). Elsevier.
Teodorescu, K., & Erev, I. (2014). Learned helplessness and learned prevalence: Exploring the causal relations among perceived controllability, reward prevalence, and exploration. *Psychological Science, 25*(10), 1861–1869.
Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs: General and Applied, 2*(4), i–109.
Weiss-Cohen, L., Konstantinidis, E., Speekenbrink, M., & Harvey, N. (2018). Task complexity moderates the influence of descriptions in decisions from experience. *Cognition, 170*, 209–227.
Yu, A. J., & Cohen, J. D (2009). Sequential effects: superstition or rational behavior? In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Vol. Eds.), *Advances in neural information processing Systems: 21*, (pp. 1873–1880). MIT Press.