Article:

Ashby, NJS, Konstantinidis, E orcid.org/0000-0002-4782-0749 and Yechiam, E (2017) Choice in experiential learning: True preferences or experimental artifacts? Acta Psychologica, 174. pp. 59-67. ISSN 0001-6918

https://doi.org/10.1016/j.actpsy.2017.01.010

Choice in experiential learning: True preferences or experimental artifacts?

Nathaniel J. S. Ashby[1*]

Emmanouil Konstantinidis[2]

Eldad Yechiam[1]

[1]Technion – Israel Institute of Technology

[2]University of New South Wales

Word Count: 8,601 (all inclusive)

[*]Correspondence concerning this article should be addressed to: Nathaniel J. S. Ashby (E-mail: nathaniel.js.ashby@gmail.com, Phone: +972 587950970), Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Technion City, Haifa 32000, Israel.

Abstract

The rate of selecting different options in the decisions-from-feedback paradigm is commonly used to measure preferences resulting from experiential learning. While convergence to a single option increases with experience, some variance in choice remains even when options are static and offer fixed rewards. Employing a decisions-from-feedback paradigm followed by a policy-setting task, we examined whether the observed variance in choice is driven by factors related to the paradigm itself: Continued exploration (e.g., believing options are non-stationary) or exploitation of perceived outcome patterns (i.e., a belief that sequential choices are not independent). Across two studies, participants showed variance in their choices, which was related (i.e., proportional) to the policies they set. In addition, in Study 2, participants' reported under-confidence was associated with the amount of choice variance in later choices and policies. These results suggest that variance in choice is better explained by participants lacking confidence in knowing which option is better, rather than methodological artifacts (i.e., exploration or failures to recognize outcome independence). As such, the current studies provide evidence for the decisions-from-feedback paradigm's validity as a behavioral research method for assessing learned preferences.

Keywords: confidence; uncertainty; preference; learning; decisions-from-feedback.

In the *decisions-from-feedback* (DF) paradigm the values of options are learned by making repeated decisions (choices) and receiving outcome feedback. DF tasks have been used extensively as a model for understanding the formation of preferences from experience (e.g., Barron & Erev, 2003; Erev, Ert, & Yechiam, 2008; Ludwig & Spetch, 2011; Roth & Erev, 1995; Weber, Shafir, & Blais, 2004) and of individual differences (e.g., Bechara, Damasio, Damasio, & Anderson, 1994; Lejuez et al., 2003). In DF tasks the proportion of selections (choices) from different options is often taken as an index of the relative preference for each of the options. A prominent example is probability learning, where individuals typically match their rate of selections from a given option to the probability they have experienced it as being rewarding. For example, if option A is rewarded 80% and option B 20% of the time, option A is typically selected around 80% of the time. The rate of selection from each option in a DF task is thus presumed to represent people's relative preferences which are contingent on reward probabilities (e.g., Bereby-Meyer & Erev, 1998). Similarly, in the Iowa Gambling task (Bechara et al., 1994), a four-alternative DF task, the proportion of selections from riskier – high reward but large loss - options is considered an index of individuals' risk taking preferences. Furthermore, Pleskac (2008) reported that there was a relationship between the proportion of risky choices in the Balloon Analog Risk Task (a task where participants learn to pump while not overfilling virtual balloons) and general risk taking attitudes, providing evidence that choosing behavior is reflective of risk preference. Still, choice proportions presumably reflect not only the relative preference for each option, but also factors driven by the paradigm itself, such as exploitation/perseveration versus exploration, and search for, or exploitation of, perceived outcome patterns (Busemeyer & Stout, 2002; Plonsky, Teodorescu, & Erev, 2015; Worthy, Pang, & Byrne, 2013; Yechiam, Busemeyer, Stout, & Bechara, 2005). This notion has led to a criticism of the usage of choice proportions (Maia & McClelland, 2004; 2005): Because choice proportions might be heavily influenced by these factors, they may be a poor index of how much each option is actually preferred. In two experimental studies we examine whether choice proportions in DF tasks represent one's preferences by comparing them to overt policies for dividing one's choices across multiple future trials – an elicitation method whereby exploration and perceived pattern exploitation are irrelevant.

The question of whether choice proportions represent preference-related (or task related) factors seems particularly pertinent in light of a curious empirical phenomenon in DF tasks: Participants rarely converge to selecting the same option exclusively. Specifically, while participants in DF tasks often learn to select options of higher expected value (EV) more frequently with experience (Ashby & Rakow, 2015; Jessup, Bishara, & Busemeyer, 2008), some variability in choice remains even after a great deal of experience (Konstantinidis, Ashby, & Gonzalez, 2015), and even occurs when options are fully described (Barron & Lieder, 2010; Lejaraga & Gonzallez, 2011; Weiss-Cohen, Konstantinidis, Speekenbrink, & Harvey, 2016; Yechiam, Barron, & Erev, 2005). As noted above, this pattern was observed in probability learning tasks (Bereby-Meyer & Erev, 1998; Erev, Bereby-Meyer, & Roth, 1999) and in a DF task where one option deterministically dominated another option: Variance in choice remained after 200 trials of choosing between one option providing 11 points with certainty and another providing 10 points with certainty (Haruvy & Erev, 2002).

In the current study we examine two potential explanations for this variance. The absence of convergence can be readily explained by factors related to the paradigm itself. As noted above, two such factors are exploration and pattern search/exploitation. First, decision makers might rightly or wrongly engage in continued exploration or monitoring of stationary environments (Mehlhorn et al., 2015; Teodorescu & Erev, 2014). Another possibility is that decision makers do not recognize the independence of their decisions and switch between options to find and exploit perceived patterns in the rewards provided (Gal & Baron, 1996; Gassmaier & Schooler, 2008; Plonsky et al., 2015). For example, Gal and Baron (1996) explored why decision makers preferred to choose from more than one option when doing so resulted in a smaller payout. They reported that some participants thought that picking from lower value options on some occasions would in fact lead to greater earnings, failing to recognize the outcome independence of successive choices. Interestingly, female participants held this belief more frequently than males. Similarly, Gassmaier and Schooler (2008) found that some participants searched for patterns, while others seemed to employ win-stay-lose-shift strategies. It therefore seems possible that much of the choice variance seen in DF tasks is due to processes that reflect a continued need to monitor the environment or to exploit perceived outcome patterns, processes that are not necessarily related to the perceived attractiveness of the options.

An alternative interpretation is that variance in choice *accurately* represent one's preferences. Under this notion, individuals might lack confidence in their assessment of which option is best (i.e., most rewarding), resulting in some degree of choice variance (Payne, Bettman, & Johnson, 1992; Slovic, 1995). Specifically, when faced with several options participants might learn from experience that one option rewards more on average, but not be entirely confident that what they have learned is correct. As a result, they might select from less preferred options as a way of hedging their bets in the event they have in fact not learned which option is better. Thus, variation in choice might better be understood as subjective uncertainty as to which option is best.

In the current studies we examined whether variance in choice in the DF paradigm is driven primarily by factors related to the paradigm (e.g., exploration or a belief that choices are not independent). This distinction is critical to determining the efficacy of using the DF paradigm as a measure of learned preferences: If experimental factors unrelated to preference drive choice behavior, the DF paradigm would be limited in its ability to elicit and infer preferences. In order to achieve this, we gauged participants' preferences in a task where unchangeable choice policies were set following experience (for similar experimental designs, see Ben-Zion, Erev, Haruvy, & Shavit, 2010; Dilla & Steinbart, 2005). Such policy questions circumvent the issue of exploration and pattern search/exploitation since participants cannot gain additional information by observing the outcomes of their policies or design their policies to exploit perceived patterns in outcomes. Policies therefore represent an index of preferential choice that is independent of exploration and pattern search/exploitation. We tested how close choice proportions in policies resembled choice proportions elicited in a DF task. If policies approximate choice proportions well, then choice proportions during the DF task are unlikely to be driven primarily by factors related to the paradigm itself (i.e., exploration, failure to recognize outcome independence, or exploitation of perceived patterns).

In addition, in Study 2 we explicitly asked participants to indicate which option they thought was more rewarding, if any, and their confidence in that assessment, as well as their preferred decision strategy (e.g., always picking the most rewarding option or attempting to find patterns in payouts). This allowed us to determine whether a lack of confidence in knowing which option is best could account for observed choice variance in DF and policy setting, providing a clearer indication of the choice strategies employed during DF tasks.

1. Study 1

We administered a DF task with outcome feedback followed by a policy setting task where the order of choices could not be controlled and where no outcome feedback was provided (both were fully incentivized). We predicted that: a) Participants would show increased maximization with experience as found previously (Bereby-Meyer & Erev, 1998), b) there would nevertheless be considerable variance in choices (Ashby & Rakow, 2015; Konstantinidis et al., 2015). Additionally, we examined the contradicting predictions concerning the relation between repeated choices and policies detailed in the introduction: If choice proportions in the DF task are strongly related to choice proportions in policy setting this suggests that choice variance is not the result of exploration, perceived pattern exploitation, or a failure to recognize outcome independence. Lastly, while not the central focus of our investigations, we examine whether variance in choice is related to gender, with females showing more variation in choice as reported by Gal and Baron (1996).

1.1 Methods

1.1.1 Participants.

One-hundred and two participants (M_{age} = 31.89; 29% Female) were recruited from Amazon Mechanical Turk and completed the study: We aimed for 100 participants to provide sufficient power to detect small-to-medium sized effects. An attention check was included – participants had to click an invisible box rather than “continue” - before the study began to screen out inattentive individuals. Those who failed the attention check were not allowed to participate. Participants received $0.25 and an additional amount contingent on their decisions (specified below). Completion time averaged 15 minutes.

1.1.2 Materials.

The first part of our task was a *decisions-from-feedback* paradigm (DF: Barron & Erev, 2003) with participants making 100 incentivized choices in each of four conditions (within-subjects and encountered in random order): Involving two, four, eight, or 16 options[1]. Options appeared as buttons labeled alphabetically ("Option A" through "Option P") and were randomly paired with the safe (two moderate outcomes occurring with equal probability) and risky (low probability of a larger outcome, but a higher chance of a smaller outcome) gambles presented in *Table 1*. Specifically, in the two-option condition gambles Safe 1 (S1; a safer higher EV option) and Risky 1 (R1; a riskier lower EV option) were randomly assigned to "Option A" and "Option B". The four, eight, and 16-option conditions included S1 and R1 as well, but also additional gambles as indicated by a "+" in *Table 1*. Thus, each condition contained equal numbers of safe and risky options, and always had the same EV maximizing option (S1). Following each choice, participants received feedback about the outcome of their decision: To minimize familiarity with the gambles over conditions either 0, 3, 7, or 11 points were added to all outcomes in a given condition (randomly determined for each participant).

[1] Options were displayed in a diamond configuration.

Table 1. Safe and risky gamble pairs outcomes in points (OS1- OS2 and OR1- OR2), probabilities (50%-50% and 20%-80%), and EVs in points (EV_{Safe} and EV_{Risky}) in all experimental conditions in Study 1. Condition (Number of Options refers) refers to whether 2, 4, 8, or 16 options were available to select from. The + signs in each column indicate which pairs were included in each set of options.

Condition (Number of Options)					Safe			Risky		
				Gamble	OS1	OS2		OR1	OR2	
2	4	8	16	Pair	50%	50%	EV_{Safe}	20%	80%	EV_{Risky}
+	+	+	+	1	70	60	65	100	30	44
	+	+	+	2	57	47	52	110	20	38
		+	+	3	52	42	47	120	10	32
		+	+	4	50	40	45	130	0	26
			+	5	56	46	51	105	25	41
			+	6	53	43	48	115	15	35
			+	7	51	41	46	125	5	29
			+	8	49	40	45	135	0	27

1.1.3 Procedure.

Participants provided consent, answered demographic questions, and were informed that they would be presented with two, four, eight, or 16 options and that they would have to play the options in order to learn what outcomes were possible. Participants were told that their goal was to earn points which would be converted to money (100 points = $0.01). After each choice participants were shown what outcome occurred (i.e., the amount they had won). After participants made 100 choices in a condition, they were told that they would make another 100 choices from the same set of options but would need to indicate up front how they would like to distribute them across the available options (i.e., assigning a number of choices between 0 and

100 to each option)[2]. Importantly, the order that policy choices would be played out could not be controlled, preventing participants from setting policies attempting to exploit perceived outcome patterns. In addition, the outcomes of policies were not revealed to participants so that they would not alter (inform) subsequent decisions: Policies were carried out and the outcomes added to a participant's earnings.

1.2 Results

1.2.1 Choices.

The left panel of *Figure 1* plots the maximization rate (i.e., the rate of selecting the EV maximizing option - S1) across the 100 choices by condition. EV maximization appears to increase over the course of the task, though to a lesser extent when more options were present. A logit regression controlling for repeated measurement confirmed these observations and is described in the Appendix.

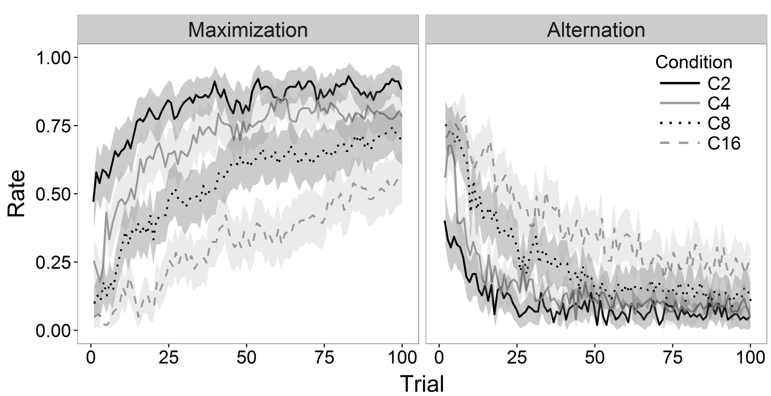

Figure 1. The left panel shows the proportion of EV maximizing choices while the right panel shows the rate of alternation. Both plotted by trial number in Study 1 over two (C2), four (C4), eight (C8), and sixteen option (C16) conditions. Shading represents 95% confidence intervals.

[2] There was no open, or other, option.

One way to examine choice variance (consistency) is to look at how frequently participants chose different options on consecutive trials (alternations). The right panel of *Figure 1* plots the alternation rate across trials and conditions and indicates that while participants became more consistent in their choices some variability remained in later trials. To examine the factors affecting variance, we performed a logit regression predicting alternations (coded 0 for no alternation, 1 for an alternation) by condition, choice number, their interactions (both coded linearly and centered), as well as gender, and clustering by participant to control for repeated measurement (Rogers, 1994). We find that the likelihood of alternating decreased with experience (b = -.02, z = -13, p < .001) and that when more options were present the likelihood of alternation was higher (b = .64, z = 13.12, p < .001) - the interaction between the two was not significant, p = .87. Gender was not a significant predictor, p = .14.

To better understand participants' choice variance in later choices we looked at the last 16 trials, as this allowed for all options to be selected in the 16 option condition (using the last 10 or 25 trials returned similar findings), and counted how many unique options were selected. *Figure 2* displays the number of participants choosing from a given number of options in each condition for their first and last 16 choices and indicates that in all but the 16 option condition the majority of participants chose the same option for each of their final 16 choices (i.e., showed no variance in choice)[3]. Nevertheless, over a third of participants in each condition made varied choices, and this variance in choice increased as the number of options available to choose from

[3] *Figure 2* also indicates that choice variance was lower in the final 16 choices than in the first 16 choices, an observation confirmed by a Poisson regression, b = -.74, z = -12.19, p < .001. A similar pattern was found in Study 2, b = -.94, z = -14.34, p < .001.

increased, as confirmed by a Poisson regression clustering by subject to control for repeated measurement, $b = .68$, $z = 9.28$, $p < .001$[4].

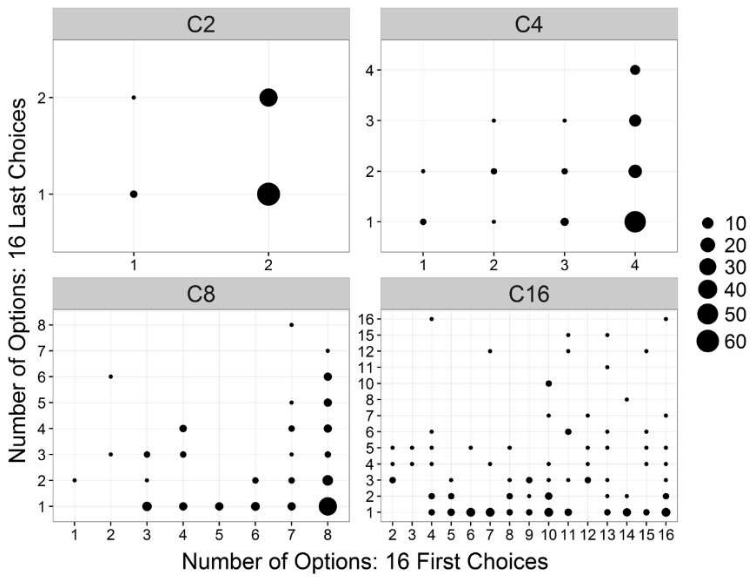

Figure 2. Scatter plot with points indicating the number of options in the first and last 16 choices a paticipant chose from in Study 1 by condition: Two (C2; top left panel), four (C4; top right panel), eight (C8; bottom left panel), and sixteen options (C16; bottom right panel). The size of the points indicate how many observations each point represents.

1.2.2 Policies.

Figure 3 plots the proportion of policies that selected only one option – set all 100 policy choices to one option – by condition and shows that about half of participants selected one option exclusively in each condition (only 30% of participants never diversified when setting their policies). However, it also indicates that many participants chose to spread their policy choices

[4] Gender was not found to be a significant predictor of choice variance in the DF task in either study, *ps* > .19.

over options: In 46% ($CI_{95\%}$ [.38, .54]) of the policies set, participants opted to spread their selections over more than one option. Nevertheless, when participants indicated a diversifying policy they selected from the EV maximizing option more than other options 87% of the time. This indicates that they had some feeling as to which option was objectively best and were not simply spreading their choices randomly across options.

To examine what influenced the decision to diversify when setting a policy, we predicted whether a policy was diversifying (coded 0 for only selecting one option, coded 1 for diversifying choices across options) by gender, condition, and a variable coding whether a participant chose only one option (or not) in their final 16 choices to examine whether variance in DF choices foreshadowed policy decisions.[5] The effect of condition was not significant, p = .57. Those who selected from more than one option in their final 16 choices were more likely (69% vs. 27%) to diversify when setting their policies, indicating that variance in later choices was predictive of diversity in policy setting, $b = 1.83$, $z = 7.71$, $p < .001$. In addition, in line with the results reported by Gal and Baron (1996), we found that females (62%) were more likely to diversify when setting policies than males (39%), $b = -.89$, $z = -2.61$, $p < .001$.

[5] We also examined whether the rate of alternation was predictive in Studies 1 and 2 and find it to be a positive predictor in each with those who alternated more being more likely to diversify when setting their policy. This suggests that the rate of alternation might be indicative of a participant's preferences in decisions-from-feedback as it is in decisions-from-samples (Hills & Hertwig, 2010).

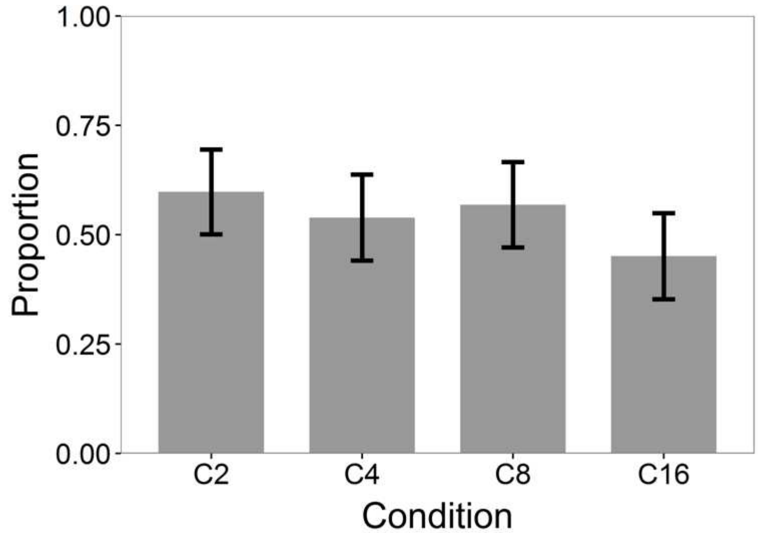

Figure 3 Proportion of participants selecting only one option when setting a policy in Study 1 over the two (C2), four (C4), eight (C8), and sixteen option (C16) conditions. Error bars represent 95% confidence intervals.

1.2.3 Choices versus policies.

To examine the relationship between choice rates and policies, we created a variable coding the proportion of choices from each option across the 100 choices in the DF and policy tasks separately for each subject in each condition. *Figure 4* presents the average correlation (and variance) between these proportions and indicates that there is a strong relationship between the proportion of choices made from each option in the DF tasks and how choices were distributed over the options when setting a policy, with the correlations ranging between .75 and .90 across conditions. The relationship was generally stable across conditions, though there was a significant decrease going from condition 2 or 4 to condition 8 and 16 (paired *t*s > 2.73, *p*s < .01).

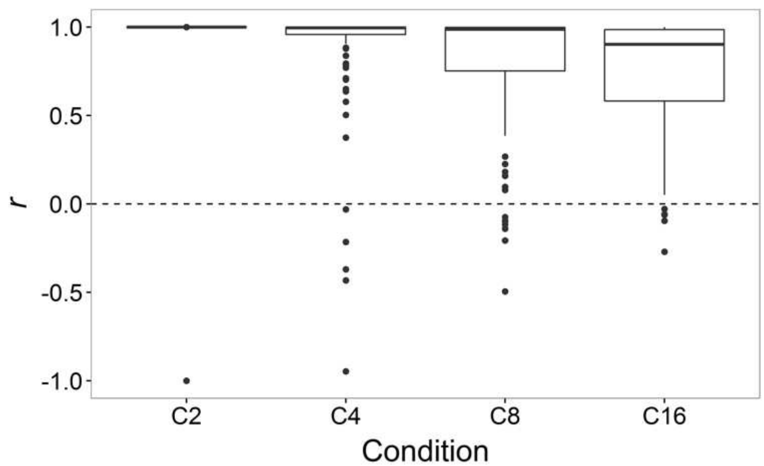

Figure 4. Boxplots displaying the average participant level correlations, and spread of correlations, between the choices proprtions in the decisions-from-feedback and policy tasks over two (C2), four (C4), eight (C8), and sixteen options (C16) in Study 1.

1.3 Discussion

We replicate the finding that while many participants drifted towards the EV-maximizing option with experience, a large proportion of participants continued to show considerable variance in choice (Ashby & Rakow, 2015). Moreover, this variance in choice carried over to policy decisions: First, about half of the policy decisions were diversified. Secondly, participants who demonstrated choice variance in their final choices in the DF task were more likely to diversify when setting a policy. Third, we found a strong positive relation between the proportion of times a participant selected from each option in the DF task and the frequency they chose to play it when setting a policy. These findings suggest that while choices may not become consistent in DF that does not necessarily indicate continued exploration or monitoring for change, nor does it imply that participants were engaged in perceived pattern exploitation. Nevertheless, the current findings do not provide insight into why this choice variation is observed.

2. Study 2

The results of Study 1 might have been influenced by the experimental design itself. Specifically, participants needed to select options in order to see what outcomes they rewarded, thus some of the choice variance in the DF task is likely reflective of exploration. Exploration which might have weakened the relationships observed between choice proportions in the DF and policy tasks. Study 2 was designed to exclude this potential confound by including foregone payoffs (payoffs of all options – realized and unrealized – on each trial). In addition, we asked participants if they thought one option was better than the others and their confidence in that assessment. We also examined their insight into what led them to show variance in choices and policies (i.e., their preferred choosing strategies) to more directly assess what influences choice variability.

2.1 Methods

2.1.1 Participants.

Five-hundred and eighty-five new participants (M_{age} = 35.92; 58% Female) were recruited from Amazon Mechanical Turk and completed the study in full. The study was run in combination with unrelated studies and the sample size was determined by the needs of those studies. The same attention check used in Study 1 was used to screen out inattentive individuals before the study began. Participants received $1.25 as well as incentivization contingent on their decisions (as specified below). Completion time averaged 30 minutes.

2.1.2 Materials and Procedure.

The design and procedure was identical to Study 1 except for the following changes: Participants only made decisions in one condition (between-subjects - randomly assigned) to reduce fatigue and order effects. The gambles were altered to make the additional risky options (risky gambles other than R1) less appealing by reducing their outcomes (see *Table 2*). The probability of the larger outcome occurring in the safer options was increased to 55% (rather than 50%) to increase its observed desirability. Points were converted with 250 points equaling $0.01. On each choice participants received outcome feedback (green font) as well as foregone outcome feedback for each of the other non-selected options (black font). In addition, participants were explicitly told that the options would not change in anyway over the course of the study. These two changes were made to ensure that if a participant was found to alternate between options in the DF task it could not be attributed to exploration or a perceived need to monitor the options for a potential change in their payoffs.

After participants made their 100 choices and set their policies, they were asked to indicate which option they thought paid out the most on average, or indicate that no option paid out more on average, as well as how confident they were in that assessment (on a slider 1000 pixels wide with ends labeled "Not at all Confident" and "Very Confident"). Lastly, participants were asked what an optimal strategy would be when making repeated choices and setting policies. For the repeated choice strategy participants were asked: "To earn the most money possible when making repeated decisions a person should": and were able to select from the following options: i) "Always pick the option which provides the highest average value.", ii) "Select the option that pays out the most on average most of the time but sometimes play other options as well.", iii) "Select options based on patterns in the outcomes (e.g., if one option wins a few times in a row its unlikely to win again).", or iv) "Distribute choices across options in proportion to their value." Similarly, for the policy setting strategy participants were asked, "To earn the most money possible when setting an unchangeable policy a person should indicate they want to play:" i) "Only the option that pays out the most on average.", ii) "Mostly the option that pays out the most on average but other options as well.", iii) "Choose from options in proportion to their value", or iv) "Choose from options at random." Thus, if participants endorsed the first option for either question it would suggest that they knew that an EV maximization strategy was optimal, while any other response would indicate that they thought inefficient strategies were superior. Note that endorsement of the second option suggests some perceived benefit of varying choices whereas endorsement of the third option in the repeated choice question would suggest a participant did not think outcomes were independent, as suggested by Gal and Baron (1996).

Table 2. Safe and risky gamble pairs outcomes in points (OS1- OS2 and OR1- OR2), probabilities (55%-45% and 20%-80%), and EVs in points (EV_{safe} and EV_{risky}) in all experimental conditions in Study 2. Number of Options refers to each condition where either 2, 4, 8, or 16 options were available to select from, while the + signs in each column indicate which pairs were included in each set of options.

Condition (Number of Options)					Safe			Risky		
2	4	8	16	Gamble Pair	OS1 55%	OS2 45%	EV_{Safe}	OR1 20%	OR2 80%	EV_{Risky}
+	+	+	+	1	70	60	65.50	100	30	44.00
	+	+	+	2	57	47	52.50	87	27	39.00
		+	+	3	55	45	50.50	85	25	37.00
		+	+	4	53	43	48.50	83	23	35.00
			+	5	51	41	46.50	81	21	33.00
			+	6	49	39	44.50	79	19	31.00
			+	7	47	37	42.50	77	17	29.00
			+	8	45	35	40.50	75	15	27.00

2.2 Results

2.2.1 Strategy Endorsement.

Table 3 reports the percentage of participants endorsing a given strategy in the repeated choice and policy questions. Interestingly, while EV maximization seemed to be the dominant strategy on the aggregate, about a third of participants endorsed a strategy where the maximizing option is chosen most frequently but other options are also selected, indicating that for many participants some kind of choice variability was attractive. Furthermore, there seems to be consistency in strategy endorsements over the conditions indicating that the number of options did not impact the perceived efficacy of each strategy. Importantly, very few participants indicated strategies reflective of a failure to recognize choice independence. This suggests that while EV maximization was seen by many as being optimal in at least some situations, a number of participants saw value in diversifying across options, and subjectively, this diversification did not seem to stem from failures to recognize outcome independence or attempts to exploit perceived outcome patterns. There was also a gender effect, with males being more likely than females to endorse the EV maximization strategy in each decision situation in line with the results of Gal and Baron (1996; $ts > 2.63$, $ps < .01$).

Table 3. Proportion of participants endorsing a given decision strategy by task, gender, and condition (C2 - C16).

Decisions-from-feedback	**Male**	**Female**	**C2**	**C4**	**C8**	**C16**
Pick only the option providing the highest average pay out.	0.56	0.35	0.44	0.47	0.40	0.45
Pick mostly the option with highest average payout but pick others sometimes.	0.28	0.42	0.32	0.38	0.38	0.36
Pick options based on patterns.	0.07	0.12	0.11	0.08	0.12	0.10
Choose from options in proportion to their value.	0.09	0.10	0.14	0.07	0.10	0.08
Policy						
Pick only the option providing the highest average pay out.	0.61	0.50	0.56	0.52	0.53	0.58
Pick mostly the option with highest average payout but pick others sometimes.	0.29	0.34	0.29	0.33	0.32	0.34
Choose from options at random.	0.01	0.01	0.13	0.15	0.14	0.07
Choose from options in proportion to their value.	0.09	0.15	0.01	0.00	0.01	0.02

2.2.2 Choices.

The left panel of *Figure 5* plots the rate of maximization across the 100 choices by condition. As in the previous study, maximization appeared to increase over the course of the task, though to a lesser extent when more options were present. A logit regression, controlling for repeated measurement, supports these observations and appears in the Appendix.

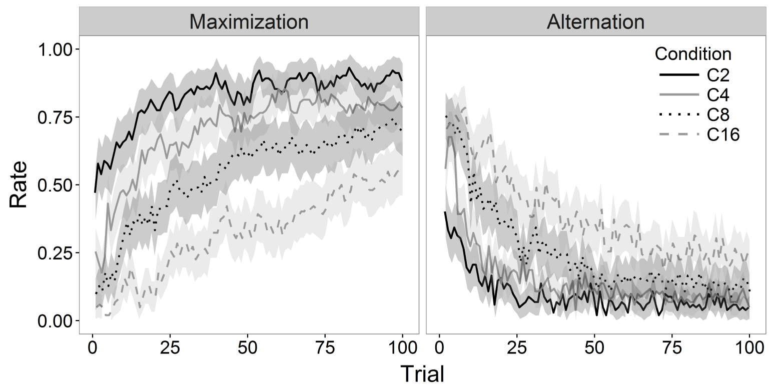

Figure 5. The left panel shows the proportion of EV maximizing choices while the right panel shows the rate of alternation. Both plotted by trial number in Study 2 over two (C2), four the (C4), eight (C8), and sixteen option (C16) conditions. Shading represents 95% confidence intervals.

The right panel of *Figure 5* plots the alternation rate across trials and conditions (left panel). Predicting alternations as in Study 1, but including a dummy variable coding for whether a participant endorsed a diversifying strategy when making repeated choices, we find that the likelihood of alternating decreased with experience (b = -.01, z = -13.74, p < .001) and that when more options were present the likelihood of alternating was higher (b = .58, z = 14.37, p < .001) - an increase that showed a stronger decrease over the course of trials relative to when fewer options were present as reflected by the interaction between trial number and condition, b = -.001, z = -4.44, p < .001. Those who endorsed an EV maximization strategy were less likely to alternate, indicating a link between the rate of alternation and the type of decision strategy favored, b = -.94, z = -10.01, p < .001. As in Study 1 there was not a significant effect of gender, p = .50

Figure 6 displays the proportion of participants choosing from a given number of options in the first and last 16 trials as in Study 1. Counter to what might have been predicted given the presence of foregone outcome feedback, the increase in the attractiveness of the maximizing option, and decrease in attractiveness of the added options, we find that participants displayed considerable choice variance. As in Study 1 the degree of variance increased as more options became available, $b = .63$, $z = 20$, $p < .001$. Furthermore, as with the rate of alternation, we find that participants who endorsed an EV maximization strategy for repeated choice showed less variance in later choices suggesting an additional link between reported decision startegy and behavior, $b = -.61$, $z = -9.15$, $p < .001$.

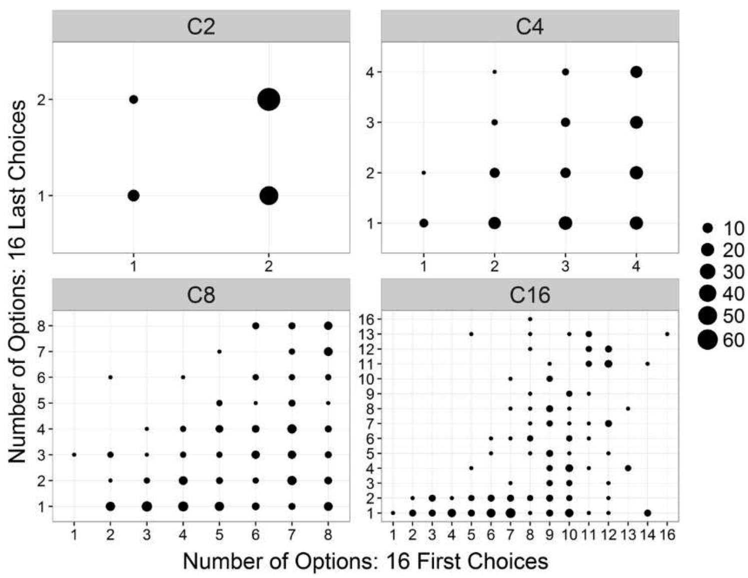

Figure 6. Scatter plot with points indicating the number of options in the first and last 16 choices a paticipant chose from in Study 2 by condition: Two (C2; top left panel), four (C4; top right panel), eight (C8; bottom left panel), and sixteen option (C16; bottom right panel). The size of the points indicate how many observations each point represents.

2.2.3 Learning and Confidence.

One reason participants might have shown variance in choice is that they did not think one option was of greater value than the others. This does not appear to be the case as 92% of participants indicated that they thought one option was better than all others. In fact, 72% of participants correctly identified the maximizing option as the best option: 81%, 84%, 69%, and 53% in the 2, 4, 8, and 16 option conditions respectively.

In terms of their confidence in their ability to identify the higher value (maximizing) option the majority of participants (84%) indicated a level of confidence greater than the mid-point (500) of the scale (M = 712.90; $CI_{95\%}$ [693.57, 732.24]) suggesting that most participants were moderately confident in their assessments. Participants who indicated the EV maximizing option as being best (M = 756.19; $CI_{95\%}$ [735.94, 776.44]) were more confident than those who did not (M = 600.82; $CI_{95\%}$ [560.03, 641.61]), and males (M = 735.92; $CI_{95\%}$ [706.67, 765.17]) reported greater confidence than females (M = 696.31; $CI_{95\%}$ [670.68, 721.94]). The level of confidence in the 16 option condition was less than in the 2, 4, and 8 option conditions (which did not differ from one another) suggesting that increasing the number of options only reduced confidence when the number of options was large, $ts > 2.5$, $ps = .01$. Nevertheless, only 25% of participants indicated strong confidence (e.g., 900 or greater). This suggests that while most participants were fairly confident in their assessments many had at least some doubt that they had learned which option was best, a factor that likely contributed to the variance in their choices. This assertion is supported by the fact that in all conditions (marginal in the eight option condition) there was a negative relationship between the number of options selected from in the final 16 choices (the variance in choice) in the DF task and the levels of confidence reported (r's $< -.14$, p's $< .10$). In other words, the more confident participants were that they had correctly identified the maximizing option the less they chose from other options. This suggests that a lack of confidence in knowing which option was best played a role in the level of choice variance observed.

2.2.4 Policies.

Figure 7 plots the proportion of policies that selected only one option for all 100 policy choices by condition and shows that in each condition most participants diversified when setting their policies (71%; $CI_{95\%}$ [.68, .75]). When participants chose to diversify, the majority (61%) selected from the maximizing option more than others.

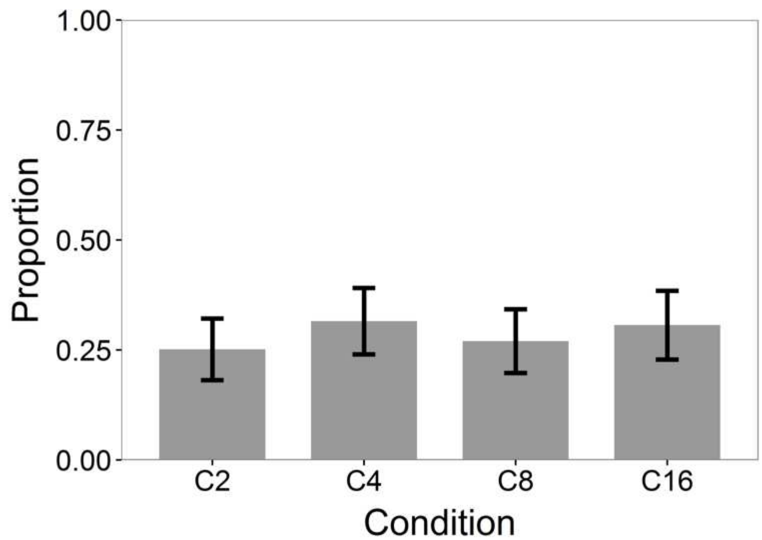

Figure 7. Proportion of participants selecting only one option when setting a policy in Study 2 over the two (C2), four (C4), eight (C8), and sixteen option (C16) conditions. Error bars represent 95% confidence intervals.

We examined what led participants to diversify by performing a logit regression predicting the decision to diversify when setting a policy as in Study 1, but including a dummy variable coding whether they had endorsed a policy strategy which favored maximization as well as the level of confidence they had reported. As in Study 1 the effect of condition was not found to be significant, $b = -.16$, $z = 1.66$, $p = .097$. We replicate the finding that those who selected from more than one option in their final 16 choices were substantially more likely (85% vs. 48%) to diversify when setting policies, $b = 1.55$, $z = 6.99$, $p < .001$. In addition, we found in line with the results of Study 1, and those reported by Gal and Baron (1996), that females (79%) were more likely to diversify when setting policies than males (60%), $b = -.80$, $z = -3.70$, $p < .001$. Perhaps unsurprisingly those who endorsed a maximization strategy (57%) were less likely to diversify when setting a policy than those who endorsed a diversifying strategy (89%), $b = -1.39$, $z = -5.70$, $p < .001$. Lastly, those who reported higher confidence that they knew which option was the best were less likely to diversify, $b = -.002$, $z = -2.84$, $p = .005$. Thus, while most participants chose to diversify when setting a policy, several factors such as confidence, gender, strategy endorsement, and prior behavior were related to this decision.

2.2.5 Choices versus Policies.

Figure 8 presents the average correlation (and variance) between the proportion of choices each option received in the DF task and the policy setting task by condition. As in Study 1, there was a strong relation between the proportion of choices made in DF and how choices were distributed over the options when setting a policy: A relationship that held across conditions, though decreased when comparing condition 2 or 4 to condition 8 and 16 (ts > 3.40, ps < .01). Interestingly, reported confidence was a positive predictor of the strength of this relationship, indicating that participants who were more confident that they had learned which option was better were more likely to employ similar choice proportions across tasks, $b = .001$, $t(553)$, $p < .001$.

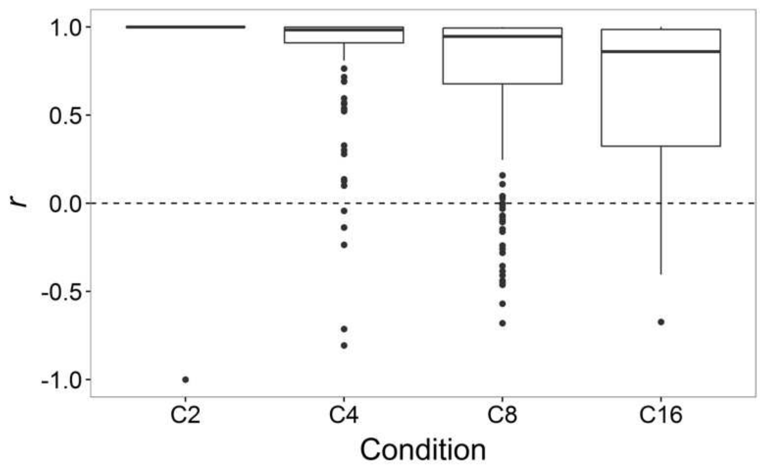

Figure 8. Boxplots displaying the average participant level correlations, and spread of correlations, between the choices proprtions in the decisions-from-feedback and policy tasks over two (C2), four (C4), eight (C8), and sixteen options (C16) in Study 2.

2.3 Discussion

In Study 2 we replicate Study 1 and provide support for the supposition 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. Specifically, in addition to finding that participants showed related choice variance in later choices and when setting policies, participants were not entirely confident in their assessment of which option was best. And, confidence was directly related to the amount of choice variation shown. Furthermore, most participants endorsed EV maximization strategies as being superior both for repeated choice and policy setting, with few indicating that they did not understand the independence of the outcomes in contrast to the findings of Gal and Barron (1996).

While not central to the motivations of the current studies it is worth noting that there were marked differences between Study 1 and Study 2: The rate of alternation and the level of variability in choice when setting policies was higher in Study 2. There are two factors that may have contributed to these differences: First, the increased experience in the task in Study 1 resulting from its within-subjects design (i.e., encountering all conditions) may have decreased alternation and choice variability as is often seen in studies involving several DF tasks (Ashby & Rakow, 2015). Analyses looking at the order conditions were encountered and the rate of switching as well as the degree of choice variance in the final 16 choices and in policies in Study 1 suggests that repeated trials was a likely contributor - there were significant decreases in alternation as well as choice variance in both the final choices and in policies in later trials, all *p*s < .001. Second, the introduction of foregone outcome feedback may have increased the likelihood that a participant would chase high value outcomes in other options or risk seeking behavior (e.g., Yechiam & Busemeyer, 2005). Comparing the overall risk rate in Study 1 (M = .26; $CI_{95\%}$ [.23, .29]) to Study 2 (M = .28; $CI_{95\%}$ [.26, .29]) does suggest a slight increase in risk taking in Study 2. Lastly, foregone outcome information may have decreased confidence about which option was best, perhaps due to memory constraints (Ashby & Rakow, 2014). As such, future research into the efficacy of the DF paradigm for eliciting preferences should be sure to take into account the impact of both repeated tasks using the same paradigm as well as the impact of foregone feedback on behavior.

3. General Discussion

We set out to uncover whether the variance in choice found in decisions-from-feedback (DF) is primarily the result of the paradigm itself (e.g., exploration or exploitation of perceived patterns). Across two studies we found that the proportion of choices from each option in the DF task was strongly related to the proportion of choices from each option when setting a policy – policies that could not be used to explore options or exploit perceived patterns. Thus, the current studies provide much needed clarification as to what can be inferred about one's preferences from choices in the DF paradigm by supporting the view that choice proportions are indeed related to preferences rather than being driven primarily by factors related to the experimental paradigm itself.

In addition, in Study 2 we found that the majority of participants knew which option was better and had confidence in this assessment (Konstantinidis & Shanks, 2014; Maia & McClelland, 2004). Nevertheless, participants' confidence in their ability to correctly identify the best option varied and this variation was predictive of the degree of choice variance they showed in the DF and policy setting tasks. Put simply, participants had a general idea that one option was better than the others, but were not entirely confident in their assessment, and therefore selected from other options to "hedge their bets". This interpretation is further supported by the fact that the majority of participants indicated that the correct strategy for maximizing one's earnings was to only pick the option paying out the most on average, or to select it most the time, rather than attempting to exploit perceived patterns in the outcomes provided by options.

It is also of interest that males were less likely to show choice variance – but only in policy setting (as in Gal & Barron, 1996), not in repeated choices. Thus, the two tasks appeared to differentially trigger traits associated with gender differences. Also, females showed less confidence in their ability to correctly identify the maximizing option, though they were no different in their ability to identify it. This latter gap between confidence and accuracy in decision performance is in line with previous research suggesting robust differences in confidence reported between the sexes (Campbell & Hackett, 1986; Lundeberg, Fox, & Punćcohaŕ, 1994). While we do not wish to make sweeping generalizations, the current results do seem to warrant closer inspection of the role individuals differences - such as gender - might play in experiential learning and confidence as such examinations should help to advance theory and the prediction of behavior both in and outside the lab.

Our results also partially address a previous concern (Steingroever et al., 2013) that healthy individuals do not maximize or have a clear learning process (e.g., exploration followed by exploitation) in the Iowa Gambling task (Bechara et al., 1994), a popular DF task used to assess individual differences. We have shown that regardless of the specific adaptation processes, individuals' decisions in DF tasks were consistent with their policy setting, suggesting that the patterns of behavior in these types of tasks are generalizable to other contexts (see also Pleskac, 2008). It should also be noted that Steingroever et al. (2013) focused on non-incentivized tasks. Without incentivization participants may indeed have lower motivation to perform and this may increase error and reduce generalizability of individual differences into other tasks.

In sum, the current investigation suggests that while there are several different factors affecting the rate of selecting different options over trials in DF tasks, there is a strong link between this rate and the rate options are selected when policies are set. This suggests that choice rates in DF are not just artifacts induced by paradigm-specific processes. Our findings thus provide support for the use of the DF paradigm to investigate preferences learned from experience, and more generally for the common usage of choice rates as indices of relative preference (e.g., Bechara et al., 1994; Bereby-Meyer & Erev, 1998) compared to the traditional usage of modal choices (i.e., which choice is preferred: Brandstätter, Gigerezner, & Hertwig, 2006; Diecidue & Wakker, 2001; Kahneman & Tverksy, 1979; Tversky & Kahneman, 1992).

Acknowledgements: E. Yechiam was supported in part by the Max Wertheimer Minerva Center for Cognitive Studies and by the I-CORE program of the Planning and Budgeting Committee and the Israel Science Foundation (Grants no. 1821/12, 199/12).

References

Ashby, N. J., & Rakow, T. (2014). Forgetting the past: Individual differences in recency in subjective valuations from experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1153.

Ashby, N. J. S., & Rakow, T. (2015). Eyes on the prize? Evidence of diminishing attention to experienced and foregone outcomes in repeated experiential choice. *Journal of Behavioral Decision Making*.

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.

Barron, G., & Leider, S. (2010). The role of experience in the gambler's fallacy. Journal of Behavioral Decision Making, 23(1), 117-129.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7-15.

Ben-Zion, U., Erev, I., Haruvy, E., & Shavit, T. (2010). Adaptive behavior leads to under-diversification. *Journal of Economic Psychology*, *31*(6), 985-995.

Bereby-Meyer, Y., & Erev, I. (1998). On learning to become a successful loser: A comparison of alternative abstractions of learning processes in the loss domain. *Journal of Mathematical Psychology*, *42*(2), 266-286.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409-432.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. *Psychological assessment*, *14*(3), 253.

Campbell, N. K., & Hackett, G. (1986). The effects of mathematics task performance on math self-efficacy and task interest. *Journal of Vocational Behavior*, *28*(2), 149-162.

Diecidue, E., & Wakker, P.P. (2001). On the intuition of rank-dependent utility. *Journal of Risk and Uncertainty*, *23*(3), 281-298.

Dilla, W. N., & Steinbart, P. J. (2005). Relative weighting of common and unique balanced scorecard measures by knowledgeable decision makers. *Behavioral Research in Accounting*, *17*(1), 43-53.

Erev, I., Bereby-Meyer, Y., & Roth, A. (1999). The effect of adding a constant to all payoffs: Experimental investigation, and implications for reinforcement learning models. *Journal of Economic Behavior and Organization*, *39*(1), 111-128.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., Hertwig, R., Stewart, T., West, R., & Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making*, *23*(1), 15-47.

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making*, *21*(5), 575-597.

Gal, I., & Baron, J. (1996). Understanding repeated simple choices. *Thinking & Reasoning*, *2*(1), 81-98.

Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, *109*(3), 416-422.

Haruvy, E., & Erev, I. (2002). Interpreting parameters in learning models. In R. Zwick & A. Rapoport (Eds.), *Experimental Business Research* (pp. 285-300). Kluwer Academic Publishers.

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.

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological Science*, *19*(10), 1015-1022.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 263-291.

Konstantinidis, E., Ashby, N. J. S., & Gonzalez, C. (2015). Exploring complexity in decisions from experience: Same minds, same strategy. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 1177–1182). Austin, TX: Cognitive Science Society.

Konstantinidis, E., & Shanks, D. R. (2014). Don't bet on it! Wagering as a measure of awareness in decision making under uncertainty. *Journal of Experimental Psychology: General*, *143*(6), 2111-2134.

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, *116*(2), 286–295.

Lejuez, C.W., Aklin, W.M., Jones, H.A., Richards, J.R., Strong, D.R., Kahler, C.W., & Read, J.P. (2003). The Balloon Analogue Risk Task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology*, *11*(1), 26-33.

Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description-experience gap in risky choice limited to rare events? *PLoS ONE*, *6*, e20262.

Lundeberg, M. A., Fox, P. W., & Punćcohaŕ, J. (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal of educational psychology*, *86*(1), 114.

Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences*, *101*(45), 16075-16080.

Maia, T. V., & McClelland, J. L. (2005). The somatic marker hypothesis: still many questions but no answers. *Trends in Cogntive Science*, *9*(4), 162-164.

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., & Gonzalez, C. (2015). Unpacking the exploration-exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, *2*(3), 191-215.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual review of psychology*, *43*(1), 87-131.

Pleskac, T. J. (2008). Decision making and learning while taking sequential risks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 167.

Plonsky, O., Teodorescu, K. & Erev, I. (2015). Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psycholgical Review*, *122*(4), 621-647.

Rogers, W. (1994). Regression standard errors in clustered samples. *Stata technical bulletin*, *3*(13).

Roth, A., & Erev, I. (1995). Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, *8*(1), 164-212.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*, 364-371.

Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., & Wagenmakers, E-J. Performance of healthy participants on the Iowa Gambling Task. *Psychological Assessment*, *25*, 180-193.

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.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297-323.

Von Neumann, J., & Morgenstern, O. (2007). *Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition)*. Princeton university press.

Weber, E.U., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, *111*(2), 430-445.

Weiss-Cohen, L., Konstantinidis, E., Speekenbrink, M., & Harvey, N. (2016). Incorporating conflicting descriptions into decisions from experience. *Organizational Behavior and Human Decision Processes, 135,* 55-69.

Worthy, D.A., Pang, B., & Byrne, K.A. (2013). Decomposing the roles of perseveration and expected value representation in models of the Iowa gambling task. *Frontiers in Psychology*, *4*, 640.

Yechiam, E., Barron, G., & Erev, I. (2005). The role of personal experience in contributing to different patterns of response to rare terrorist attacks. Journal of Conflict Resolution, 49(3), 430-439.

Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review*, *12*, 387–402.

Yechiam, E., Busemeyer, J. R, Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision making deficits. *Psychological Science*, *16*(12), 973-978.

Appendix

A1. Learning Patterns

To examine participants' learning patterns we conducted a logit regression separately for each study predicting selection of the maximizing option (coded 0 for not selecting the maximizing option, 1 for selecting it) by condition and choice number (both coded linearly and centered) and their interaction, gender, whether a maximization strategy was endorsed for repeated choice (Study 2), and cluster on over subjects to control for repeated measurement (Rogers, 1994).

In Study 1 the likelihood of maximization increased with experience, $b = .02$, $z = 12.65$, $p < .001$. Increases in the number of options decreased the likelihood of maximization ($b = -.82$, $z = -16.40$, $p < .001$). The interaction between experience and condition was not significant, $p = .15$, nor was the effect of gender, $p = .26$.

In Study 2 the rate of maximization increased with experience, $b = .01$, $z = 12.82$, $p < .001$. In addition, increases in the number of options decreased the likelihood of maximization ($b = -.64$, $z = -13.93$, $p < .001$) and led to slower learning of which option was superior as indicated by the significant interaction, $b = .002$ $z = 2.71$, $p = .007$. The main effect of gender was not significant ($p = .39$). Lastly, participants who endorsed an EV maximization strategy for repeated choice were more likely to maximize, indicating that participants had insight into the type of decision strategy they preferred, $b = .97$, $z = 8.85$, $p < .001$.