The Evaluative Advantage of Novel Alternatives: An Information Sampling Account

Running Head: Information Sampling and Novelty

Gaël Le Mens [corresponding author]

Universitat Pompeu Fabra, Department of Economics and Business, Spain

Address: Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain

Email: gael.le-mens@upf.edu

Yaakov Kareev

The Federmann Center for the Study of Rationality, The Hebrew University, Jerusalem, Israel

Judith Avrahami

The Federmann Center for the Study of Rationality, The Hebrew University, Jerusalem, Israel

March 11, 2015

Manuscript length: 1914 words (excluding methods and results sections)

ABSTRACT

New products, services and ideas are often evaluated more favorably than similar but older items. Although several explanations exist, we identify an overlooked asymmetry in information about new and old items that emerges when people seek positive experiences and learn about the qualities of (noisy) alternatives by experiencing them. We analyze a simple learning model and demonstrate that in such settings most people will tend to evaluate a new alternative more positively than an older alternative with the same payoff distribution. The reason is that, when people seek positive experiences and thus avoid selecting again alternatives that led to poor payoffs, this precludes additional feedback on their qualities. Negative quality estimates, even when caused by noise thus tend to persist. This negative bias takes time to develop, and affects old alternatives more strongly than similar but newer alternatives. An experimental study with 769 participants supports the predictions of our model. [150 words]

Keywords: Judgment, Learning, Attitudes, Novelty, Information Sampling.

INTRODUCTION

Why do people like new items, practices, gadgets and ideas? Prior work has advanced two main classes of explanations. The first proposes that people like the new because it solves some problems that the old could not address (Boyd & Richerson, 1988; Schumpeter, 1939), or it is a better fit for changing tastes (Peterson & Berger, 1975; Lieberson, 2000). The second focuses on how things catch on. This research has generally relied on some imitation mechanism to explain how new artists, albums, restaurants or technological gadgets become popular while older items fall out of favor (Bass, 1969; Rogers, 1995; Salganik, Dodds, & Watts, 2006; Wasserman & Faust, 1994).

In this paper, we propose a different explanation for the evaluative advantage of novel alternatives. Our proposal builds on the adaptive sampling model of attitude formation, initially proposed by Fazio, Eiser & Shook (2004) and Denrell (2005; see also Einhorn & Hogarth, 1978, and March, 1996). This perspective focuses on situations in which people form attitudes from their own experiences with the alternatives and the alternatives have uncertain, noisy, payoffs. The crucial assumption is that people tend to sample again alternatives that led to positive experiences and to avoid alternatives that led to negative experiences. This avoidance behavior precludes feedback on the qualities of alternatives that led to poor payoffs. Negative evaluations thus tend to persist. This, in turn, leads people to frequently underestimate the qualities of the experienced alternatives (Fazio, Eiser & Shook, 2004; Denrell, 2005; Fetchenhauer & Dunning, 2010; Le Mens & Denrell, 2011; Smith & Collins, 2009).

Our explanation for the evaluative advantage of novel alternatives relies on the fact that this underestimation tendency is an *emergent* phenomenon that results from the sequential nature of the sampling process. In fact, the phenomenon becomes more pronounced the longer the experience with the original alternatives. When a new alternative becomes available, it has not yet been subject to adaptive sampling. Thus, the consequent systematic underestimation has not emerged yet. The new alternative will tend to be evaluated more positively than existing alternatives of the same quality, even if information about the new alternative is not processed more positively than information about the existing alternatives. The mechanism we propose is relevant to understanding human judgment even if information about new alternatives and existing alternatives is processed differently, because it focuses on a different level of analysis: the information on which cognitive processes operate. Our theory relies on few assumptions, and therefore has broad applicability. It can provide new insights on human judgments about a variety of attitude targets such as other people, artists, musical and movie genres, restaurants, political ideas or research streams.

We first provide a formal analysis of a simple learning model. Then we use computer simulations to show that our main result still holds if we relax some of our modeling assumptions. Finally, we report an experiment designed to test the prediction of our theory.

THE QUALITY ESTIMATION MODEL – FORMAL ANALYSIS

We consider a task environment where an individual faces a multi-armed bandit problem (Sutton & Barto, 1998). The individual makes a sequence of decisions between K alternatives with unknown (and noisy) payoff distributions. Her total payoff is the sum of the payoffs obtained in each period. The individual seeks to obtain positive payoffs and 'learns by doing': she updates her quality estimates for the alternatives on the basis of the sampled information. We assume that the individual is more likely to sample again an alternative that led to positive outcomes than an alternative that led to poor outcomes. This is the *adaptive sampling assumption* (see Fazio, Eiser & Shook, 2004; Denrell, 2005). Adaptive sampling is a reasonable choice heuristics when the individual tries to maximize the sum of her payoffs (Sutton & Barto, 1998; Denrell, 2005; Le Mens & Denrell, 2011). Besides, it is consistent with existing evidence about behavior in this kind of tasks (Erev & Barron, 2005; Thorndike, 1927; Hull, 1930; for a review see Erev & Roth, 2014).

Where our setting differs from the standard bandit problem is that, after some time – at period τ – a new alternative, Alt. N, becomes available. The payoff distribution of the new alternative is the same as that of one of the other alternatives. Without loss of generality, we assume this 'comparison' alternative is Alt. 1.We focus on how quality estimates about the new alternative, Alt. N, compare with the quality estimates of Alt. 1.

Model [METHODS]

We keep the model simple in order to derive formal results and illustrate the main intuition. Unless otherwise noted, we denote random variables by capital letters, and their instantiations by corresponding lower-case letters.

Payoff Distributions

We denote the density of the payoff distribution of Alt. k by f_k. The payoff distributions are continuous with positive variance. Although we assume continuity for simplicity, our proofs can easily be adapted to cases where the payoff distribution is discrete. Our proofs are also valid if one or several alternatives other than Alt. 1 have deterministic payoffs.

Initial Quality Estimates

The quality estimate for Alt. k at the beginning of period t is denoted by $\hat{Q}_{k,t}$. The initial quality estimate of every alternative is a random draw from its payoff distribution; it is therefore unbiased.

Estimate Updating

In all periods, $\hat{Q}_{k,t}$ is equal to the last observed payoff of Alt. k or to its initial quality estimate if the alternative has not yet been selected. Although we make this assumption for analytical convenience, recent experimental evidence on sequential choice between uncertain alternatives shows that choices are well predicted by a model that focuses only on the most recent outcome (Avrahami & Kareev, 2010; Kareev, Avrahami, & Fiedler, 2014).

Sampling Rule

Let $pS^k(\hat{q}_{1,t}, \dots, \hat{q}_{K,t})$ denote the likelihood the individual samples Alt. k in period t given her quality estimates for the K alternatives. We implement the adaptive sampling assumption by assuming that pS^k is increasing in $\hat{q}_{k,t}$.

Main Formal Result [RESULTS]

The assumptions about sampling and estimate updating imply Alt. 1 is likely to be believed inferior to Alt. *N* when Alt. *N* becomes available. That is, the probability that Alt. *N* is believed to be better than Alt. 1 is higher than the probability that it is believed to be worse than Alt. 1. This is formalized in the following theorem:

Theorem 1: $P\left[\hat{Q}_{N,\tau} > \hat{Q}_{1,\tau}\right] \geq P\left[\hat{Q}_{N,\tau} < \hat{Q}_{1,\tau}\right]$.

Proof: See Supplemental Material available online.

Discussion

This result is explained by the emergence of a systematic underestimation tendency for Alt. 1 (see Lemma 2 in the proof presented in the Supplemental Material available online – see also Denrell (2005) for a different formulation). This, together with the assumption that the initial estimate for the new alternative is unbiased (it is a random draw from the payoff distribution), implies an evaluative advantage for the new alternative.

Theorem 1 holds for almost any payoff distribution and any sampling rule that is increasing in the quality estimates. In order to prove this result, we assumed that the quality estimate for an alternative was equal to its last observed payoff (without this assumption, the proof becomes intractable). In the next section, we use computer simulations to show that our result continues to hold even if this assumption is relaxed and quality estimates are weighted averages of observed payoffs.

THE QUALITY ESTIMATION MODEL – NUMERICAL SIMULATIONS

The task environment is a version of the bandit setting described in the previous section, with initially just two alternatives. After some time, a third alternative is introduced. The third alternative has the same payoff distribution as Alt. 1. This simple task environment will also be used for the experiment described in the next section. Ancillary analyses show that similar results hold if there are more alternatives.

Model [METHODS]

Payoff Distributions

The payoff distributions of both Alt. 1 and Alt. *N* are uniform between 0 and 130. Alt. 2 is less risky, with a uniform payoff distribution between 62 and 68. There are 11 periods. The novel alternative, Alt. *N* becomes available at the beginning of the last period. Our focus is on the comparison between the quality estimates for the two alternatives with the same payoff distribution but different times of entry: Alt. 1 and *N*. We focus on estimates for alternatives with high variability because the underestimation tendency for the old alternatives tends to be stronger when payoffs are highly variable (see March, 1996; Denrell, 2005).

Estimate Updating

The updated quality estimate is a weighted average of the past estimate and the last observation. When the individual selects Alt. k, her quality estimate is updated as follows: $\hat{Q}_{k,t+1} = (1-b)\hat{Q}_{k,t} + bq_{k,t}$, where $q_{k,t}$ is the sampled payoff of Alt. k in period t, and b is the weight of the last observation ($0 < b \leq 1$). When an alternative is not sampled, its quality estimate does not change.

Sampling Rule

We assume that the sampling rule is a logistic choice rule and the individual updates her quality estimates using the ‘delta rule' (Busemeyer & Myung, 1992). Prior research has shown that this simple model provides a good fit to experimental data on sequential choice under uncertainty (Denrell, 2005). In each period before the introduction of *N*, the individual selects either Alt. 1 or Alt. 2. The sampling likelihoods are:

$$pS^1(\hat{q}_{1,t}, \hat{q}_{2,t}) = \frac{e^{s\hat{q}_{1,t}}}{e^{s\hat{q}_{1,t}} + e^{s\hat{q}_{2,t}}}, pS^2(\hat{q}_{1,t}, \hat{q}_{2,t}) = 1 - pS^1(\hat{q}_{1,t}, \hat{q}_{2,t}),$$

where *s* is a parameter that regulates the sensitivity of the sampling probability to the quality estimates ($s>0$). This equation implies that the individual is most likely to sample the alternative she believes to have the higher quality. Once *N* becomes available, the sampling likelihoods are adapted to the three-alternatives choice:

$$pS^1(\hat{q}_{1,t}, \hat{q}_{2,t}, \hat{q}_{N,t}) = \frac{e^{s\hat{q}_{1,t}}}{e^{s\hat{q}_{1,t}} + e^{s\hat{q}_{2,t}} + e^{s\hat{q}_{N,t}}},$$

$$pS^2(\hat{q}_{1,t}, \hat{q}_{2,t}, \hat{q}_{N,t}) = \frac{e^{s\hat{q}_{2,t}}}{e^{s\hat{q}_{1,t}} + e^{s\hat{q}_{2,t}} + e^{s\hat{q}_{N,t}}},$$

$$pS^N(\hat{q}_{1,t}, \hat{q}_{2,t}, \hat{q}_{N,t}) = \frac{e^{s\hat{q}_{N,t}}}{e^{s\hat{q}_{1,t}} + e^{s\hat{q}_{2,t}} + e^{s\hat{q}_{N,t}}}.$$

Results: Sensitivity to model parameters [RESULTS]

Alt. *N* has a clear evaluative advantage as compared to Alt. 1 at the time of introduction (beginning of the last period) and at the end of the last period (see Table 1). It is worth noting that it is not necessary that *N* be introduced many periods after Alt. 1 for it to have an evaluative advantage. Figure 1 illustrates what happens when *N* is introduced shortly after Alt. 1 (with $\tau = 3$). In this case, *N*'s evaluative advantage is not as strong as when *N* is introduced much later than Alt. 1. This is because the underestimation tendency for Alt. 1 tends to increase over time, and thus the later the time of entry of *N*, the stronger its evaluative advantage. Figure 1 also shows that Alt. *N*'s evaluative advantage diminishes with time. This is not surprising, because Alt. *N* is also subject to adaptive sampling and to the consequent underestimation tendency.

At the time of introduction, the evaluative advantage of *N* tends to be stronger when the weight of new evidence is high (*b* is close to 1). It is also stronger when the sensitivity of the sampling rule to quality estimates is higher (*s* is higher). This is because the underestimation tendency for Alt. 1 implied by adaptive sampling becomes stronger with *b* and *s*. The effect of *b* declines with time, and even reverts after enough time has elapsed since the entry of the novel alternative (see Figure 1). The reason is that when *b* is high, the learning process about the new alternative is faster than when *b* is low. This implies, in turn, that the emerging underestimation tendency implied by adaptive sampling affects *N*'s quality estimate faster when *b* is high. Similarly the evaluative advantage of *N* diminishes faster when *s* is higher. When *s* is high, *N* is almost always chosen when it is introduced because it is evaluated more positively and the choice rule is sensitive to differences in evaluations. If the sampled outcome is poor, then *N* is immediately avoided. In other words, when *s* is high, most of the learning is about *N*, and adaptive sampling quickly leads to an underestimation tendency for the new alternative as well.

Table 1: Probability that Alt. N is evaluated more favorably than Alt. 1 ($P[\hat{Q}_{N,t} > \hat{Q}_{1,t}]$) as a function of the slope parameter of the logistic choice rule (s) and the weight of new evidence (b). When $s = \infty$, the individual always selects the alternative with the highest quality estimate. For all the combinations of parameter values Alt. N tends to be evaluated more positively than Alt. 1: $P[\hat{Q}_{N,t} > \hat{Q}_{1,t}] > 0.5$. There are 11 periods and $\tau = 11$. Based on 50 000 simulations.

	s=∞		*s=0.1*		*s=0.02*		*s=0.01*	
	t=11	*After last period*	*t=11*	*After last period*	*t=11*	*After last period*	*t=11*	*After last period*
b=0.25	0.63	0.62	0.61	0.61	0.54	0.54	0.52	0.52
b=0.5	0.67	0.64	0.67	0.64	0.55	0.53	0.53	0.52
b=0.75	0.71	0.63	0.72	0.68	0.57	0.54	0.54	0.52
b=0.1	0.76	0.65	0.80	0.72	0.61	0.58	0.56	0.54

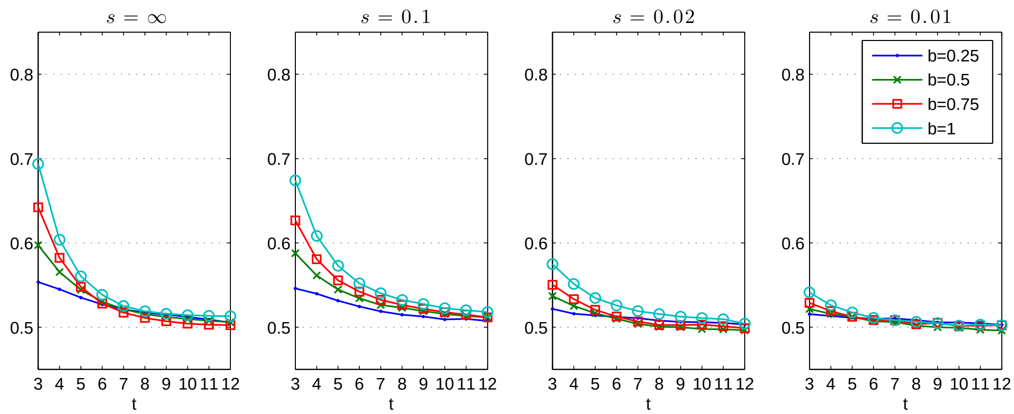

Figure 1: Same as Table 1, but with $\tau = 3$.

Discussion

We have demonstrated that adaptive sampling yields an evaluative advantage for a new alternative over an alternative with the same payoff distribution that has been available for longer. Our simulations relied on a number of modeling assumptions regarding prior estimates, estimate updating, and sampling behavior. To verify the realism of these assumptions, we designed and ran an experimental study.

EXPERIMENTAL STUDY

Participants faced the same task environment as in the simulations of the previous section. They were instructed to maximize the total number of points obtained in the experiment while learning about the payoff distributions of the alternatives and were compensated accordingly.

Our focus is on the comparison between the quality estimates for the two alternatives with the same payoff distribution but different times of entry: Alt. 1 and Alt. N. Let est_1 and est_N denote the subjective quality estimates for Alt.1 and Alt. N at the end of the experiment. Based on the analyses of the previous sections, we predict that participants will be more likely to believe Alt. N to be superior to Alt. 1 than to believe it to be inferior to Alt. 1: $P[est_N > est_1] > P[est_N < est_1]$.

Design [METHODS]

769 participants recruited via Amazon Turk completed an online experiment. They were paid in proportion to the number of points obtained in addition to a fixed minimal payment. The task environment was the same as in the simulations of the previous section, with the parameters used for Table 1. There were 11 periods and Alt. *N* became available at the beginning of the last period. The payoff distribution of Alt. *N* was the same as for Alt. 1: a uniform distribution between 0 and 130. The payoff distribution for Alt. 2 was uniform between 62 and 68.

The only information participants knew about the payoff distributions was that they were not changing over time (as implied by the cover story – see Supplemental Material available online). In order to ensure that participants had at least some unbiased prior information about all the alternatives, they were provided with a 'free' observation of one random draw of each alternative just before it was introduced. More specifically, participants saw one random draw of Alt. 1 and Alt. 2 before period 1 and one random draw of Alt. *N* between period 10 and period 11.

At the end of the study, participants were asked to state what payoff they would expect if they selected each alternative one more time.

Sample Size

Based on pre-tests, we predicted the parameter values of the learning model would be around .02 for *s* and .75 for *b*. Simulations of our model suggest that with these parameters the proportions of participants that would evaluate Alt. *N* more positively than Alt. 1 would be about 0.55. Calculations suggest a sample size of at least 545 participants for this proportion to be statistically different from the null hypothesis of .50 ($p=0.01$). The data was collected over several online sessions and we stopped collection of data for the study at the end of the session that number was achieved.

Results [RESULTS]

Sampling Behavior

In order to assess the extent to which participants engaged in adaptive sampling, we estimated the parameters of the model discussed in the previous section. When $s>0$ and $b>0$, participants could be regarded as behaving according to the adaptive sampling assumption.

To estimate s and b, Maximum Likelihood Estimations were implemented, using the unconstrained optimization routine *fminsearch* in Matlab R2013a. Standard errors are estimated using the BHHH estimator (see Greene, 2003, p. 481). The best fitting values are s=0.021, 95% CI [0.019,0.022], and b=0.74, 95% CI [0.68,0.80]. The positive estimates indicate that participants engaged in adaptive sampling, as expected. The high value for b suggests that participants were subject to a strong recency effect, in line with existing empirical evidence (Avrahami & Kareev, 2010).

The fact that participants' behavior can be explained by the adaptive sampling model implies that there will be an information asymmetry in favor of the new alternative. 100,000 simulations of the model of the previous section with the estimated parameters (s=0.021 and b=0.74) predict that 54% of the participants will have a quality estimate higher for Alt. N than for Alt. 1.

Final Quality Estimate: Based on Information Sampled by the Participants

In order to verify our prediction, we computed what would have been the participants' quality estimates, had they integrated the information they actually sampled according to the delta rule (see model of the previous section). We used the estimated parameter value of b=0.74. Let $\hat{Q}_{1,F}$ and $\hat{Q}_{N,F}$ denote the final quality estimates implied by this estimate-updating model. These estimates are recency-weighted averages of the payoffs actually obtained by the participants.

Consistent with our prediction, the implied quality estimate for the new is higher than the implied quality estimate for the old in more than 50% of the cases. The proportion of participants for whom $\hat{Q}_{N,F} > \hat{Q}_{1,F}$ is 0.57, 95% CI [0.54,0.61]. (The confidence intervals on proportions are constructed using

Beta(α+1,β+1) distributions where α is the number of participants for which the focal inequality (strictly) holds and β is the number of participants for which the opposite inequality (strictly) holds.) Furthermore, the implied quality estimate for the old is lower than its true mean more frequently than the quality estimate for the new: $\hat{Q}_{1,F}$ is below 65 (the true mean payoff) in 66% of the cases (95% CI [0.62,0.69]), whereas $\hat{Q}_{N,F}$ is below 65 in 53% of the cases (95% CI [0.50,0.57]).

In summary, there is a clear asymmetry in favor of Alt. N as compared to Alt. 1. Ancillary analyses show that there is also an asymmetry in favor of Alt. *N* in terms of last observed payoffs and in terms of non-weighted average observed payoffs (see Supplemental Material available online).

Final Quality Estimates: Subjective Estimates Given by the Participants

The information asymmetry in favor of *N* should translate into an evaluative advantage for this alternative as expressed by participants' final quality estimates (the value they expected to get with one additional draw). These subjective estimates, denoted by est_1 and est_N, favor Alt. *N*: Among the 735 participants with different estimates for the two alternatives, the proportion of participants with $est_N > est_1$ is 0.53, 95% CI [0.50, 0.57]. In addition, Alt. 1 tends to be underestimated as compared to the true common mean of 65. The proportion of participants with $est_1 < 65$ is 0.58, 95% CI [0.54,0.61]. Alt. *N* is also subject to a tendency for underestimation, but it is (unsurprisingly) much weaker than for Alt. 1. The proportion of participants with $est_N < 65$ is 0.51, 95% CI [0.47,0.55].

Individual Differences

Participants' choices and estimate updating behavior conformed to varying degrees to the assumptions of our model. We estimated the model parameters *s* and *b* individually for each participant. Not surprisingly, the evaluative advantage in favor of Alt. *N* was stronger for those who engaged more clearly in adaptive sampling. For example, among the participants with *s* higher than the median (*m* = 0.027), the proportion of participants with $est_N > est_1$ is 0.58, 95% CI [0.53,0.63]. This proportion is 0.49, 95% CI [0.44,0.54], for those participants with *s* lower than the median.

Discussion

Participants obtained samples of information about a new alternative that were generally more positive than the samples of information they obtained about an identical alternative that had been available since the beginning of the experiment. They also came to evaluate the new alternative more positively than the 'older' alternative despite the fact that these two alternatives had the same payoff distribution. The asymmetry in subjective estimates (est_1 and est_N) is not as strong as the asymmetry in estimates implied by the sampled information ($\hat{Q}_{1,F}$ and $\hat{Q}_{N,F}$). Nevertheless, the fact that the asymmetry emerges in terms of the subjective estimates is important, because it demonstrates that people do not fully correct for the information bias induced by their own sampling behavior. This is consistent with the claim that people lack the meta-cognitive ability to correct for sampling biases (Fiedler, 2012; Le Mens & Denrell, 2011; Kareev, Arnon, & Zeliger, 2000).

GENERAL DISCUSSION AND CONCLUSION

Our results apparently run against the evidence that supports preferences for familiar items such as the numerous studies on mere exposure effects (e.g. Zajonc, 1968; 2001). However, our approach neither contradicts empirical evidence in support of mere exposure effects nor questions the relevance of these effects. Instead, our results suggest a different mechanism that might play a complementary role in explaining attitudes and preferences in naturally occurring environments. Although the old benefits from prior exposure, our theory suggests conditions under which information about the value of an alternative might be subject to a systematic negative bias that, in turn, helps the new.

It is straightforward to adapt our simulations to a setting where there is a positive effect of exposure on evaluations. Assume, for example, that the true qualities of the alternatives increase by 2 every period since introduction.[1] Simulations with s=0.021, b=0.74 and τ =11 show that at the end of 11

[1] The payoff distribution of Alt. 1 and N is Uniform(2*(t-1),2*(t-1)+130) and the payoff distribution of Alt. 2 is Uniform(2*(t- τ)+62, 2*(t- τ)+68).

periods, the likelihood Alt. *N* is evaluated more positively than Alt. 1 is 0.42. When there is no adaptive sampling (i.e. the available alternatives have equal sampling probabilities), this likelihood is 0.33. In this case as well, adaptive sampling has a positive effect on relative evaluations of the new as compared to the old. A natural follow-up to this project would explore empirically how the two mechanisms – mere exposure and adaptive sampling – jointly affect evaluative judgments.

Our findings also seem to run against the ample evidence that supports ambiguity aversion (e.g. Ellsberg, 1961). After all, people know very little about the payoff distribution of the new alternative (just one or two realizations of the payoff distribution in our experiment) whereas they have much more information about the payoff distributions of the old alternatives. Ambiguity aversion suggests that people would avoid the new. Nevertheless, participants came to evaluate the new more positively than the old. Our result thus suggests a preference reversal in favor of the ambiguous alternative. Most of the existing evidence in support of ambiguity aversion uses a decision-from-description paradigm where participants choose between options whose payoff distributions are described in the form of sets of possible outcomes and associated probabilities. By contrast, in our setup, participants had to learn the payoff distributions from their own experiences. Existing research has found that risk preferences in decision-from-experience settings tend to differ from risk preferences in decision-from-description settings (Hertwig, Barron, Weber, & Erev, 2004). Our results tentatively suggest that there might be a similar phenomenon at play regarding preferences for or against ambiguity.

Our model and findings can help explain why trends catch on but rarely last. The leading explanations for why items become popular rely on some social influence mechanism whereby people tend to select alternatives that others in their social networks have selected (e.g. Bikhchandani, Hirshleifer, & Welch, 1992). Our theory suggests conditions under which some people may prefer new items in the first place. If others in their social network imitate these early adopters, a diffusion process might be triggered, and the new might become popular. As a result of its popularity, the new becomes available to other agents. Our model predicts that it becomes subject to declining evaluations as time passes. Soon enough, the new is subject to an evaluative disadvantage as compared to even newer items.

Settings where our mechanism likely contributes to short-lived fashions include attitudes toward restaurants, culinary trends, filmmakers, movie genres, actors, musicians or book writers. Analyzing how our individual-level mechanism can contribute to theories of fads and fashion that rely on models of collective behavior is a promising avenue for future research.

AUTHOR CONTRIBUTIONS

G. Le Mens, Y. Kareev, and J. Avrahami contributed to the study design, data collection and analysis. G. Le Mens developed the proofs the formal results and drafted the manuscript. Y. Kareev, and J. Avrahami provided critical revisions. All authors approved the final version of the manuscript for submission.

REFERENCES

Avrahami, J., & Kareev, Y. (2010). The role of impulses in shaping decisions. *Journal of Behavioral Decision Making*, *24*(5), 515–529. doi:10.1002/bdm.707.

Bass, Frank (1969). A new product growth model for consumer durables. *Management Science, 15* (5): 215–227.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy, 100(5)*, 992–1026.

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology General*, *121*(2), 177–194.

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review*, *112*(4), 951–978. doi:10.1037/0033-295X.112.4.951

DiMaggio, P. (1997). Culture and cognition. *Annual Review of Sociology*, 263-287.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The QuarterlyJjournal ofEeconomics*, 643-669.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, *85*(5), 395.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological review*, *112*(4), 912.

Erev, I., & Roth, A. E. (2014). Maximization, learning and economic behavior. *Proceedings of the National Academy of Sciences, 111(Supplement 3)*, 10818-10825.

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology*, *87*(3), 293–311. doi:10.1037/0022-3514.87.3.293

Fetchenhauer, D., & Dunning, D. (2010). Why so cynical?: Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others. *Psychological Science*, *21*(2), 189–193. doi:10.1177/0956797609358586.

Fiedler, K. (2012). Meta-cognitive myopia and the dilemmas of inductive-statistical inference. In B. H. Ross, *The psychology of learning and motivation.* (Vol. 57, pp. 1–46). London: Academic Press.

Greene, W. H. (2003). *Econometric analysis, 5th Ed*. New York: MacMillan.

Hertwig, R., Barron, G., Weber, E., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15(8)*, 534–539.

Hull, C. L. (1930). Simple trial and error learning: A study in psychological theory. *Psychological Review, 37*, 241-256.

Kareev, Y., Arnon, S., & Zeliger, R. (2002). On the misperception of variability. *Journal of Experimental Psychology: General, 131(2)*, 287-297.

Kareev, Y., Avrahami, J., & Fiedler, K. (2014). Strategic interactions, affective reactions, and fast adaptations. *Journal of Experimental Psychology: General*, *143*(3), 1112–1126. doi:10.1037/a0034641

Karlin, S., & Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *Journal of Multivariate Analysis*, *10*(4), 467-498.

Le Mens, G., & Denrell, J. (2011). Rational learning and information sampling: On the “naivety” assumption in sampling explanations of judgment biases. *Psychological Review*, *118*(2), 379–392. doi:10.1037/a0023010.

Lieberson, S. (2000). *A matter of taste: How names, fashions, and culture change*. New Haven: Yale University Press.

March, J. G. (1996). Learning to be risk averse. *Psychological Review*, *103*(2), 309-319.

Peterson, R. A., & Berger, D. G. (1975). Cycles in symbol production: The case of popular music. *American Sociological Review*, *40*(2), 158–173. doi:10.2307/2094343

Rogers, E. M. (1995). *Diffusion of innovations*. New York: The Free Press.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*(5762), 854–856. doi:10.1126/science.1121066

Schumpeter, J. A. (1939). *Business cycles: A theoretical, historical and statistical analysis of the capitalist process.* New York: McGraw-Hill.

Smith, E. R., & Collins, E. C. (2009). Contextualizing person perception: Distributed social cognition. *Psychological Review*, *116*(2), 343–364. doi:10.1037/a0015072.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. Cambridge, MA: MIT Press.

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology, 39(1/4)*, 212-222.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, MA: Cambridge University Press.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9(2), 1-27.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, 10(6), 224-228. doi:10.1111/1467-8721.00154.

SUPPLEMENTAL ONLINE MATERIAL: PROOF OF THEOREM 1

Our proof build on a notion of stochastic ordering described by Karlin and Rinott (1980). Let g and g' be two densities defined on the real line (denoted $\mathbb{R}$). According to the definition of the $\succ_{TP_2}$ relation in Karlin and Rinott (1980, p.472-473), $g \succ_{TP_2} g'$ if and only if for all $x, y \in \mathbb{R}$,

$$g(\max(x,y))g'(\min(x,y)) \geq g'(x)g(y).$$

We start by proving a lemma that claims that in all periods, the distribution of quality estimates is lower than the payoff distribution according to the $\succ_{TP_2}$ relation. For all $t \geq 1$, let $f_{1,t}$ denote the distribution of $\hat{Q}_{1,t}$, the quality estimates of Alt. 1 at the beginning of period t. For all $t \geq \tau$, let $f_{N,t}$ denote the distribution of $\hat{Q}_{N,t}$.

Lemma 2: For all $t \geq 1$, $f \succ_{TP_2} f_{1,t}$.

Proof: For all t, let $\Delta_t(x,y) = f(\max(x,y))f_{1,t}(\min(x,y)) - f_{1,t}(x)f(y)$

Proving the lemma is equivalent to showing that for all $t \geq 1$ and for all $x, y \in \mathbb{R}$, $\Delta_t(x,y) \geq 0$. We use a proof by induction.

<u>Base case:</u> Let $t = 1$. We need to show that for all $x, y \in \mathbb{R}$, $\Delta_1(x,y) \geq 0$. Since $f_{1,1} = f$, we have $\Delta_1(x,y) = 0 \geq 0$.

<u>Inductive step:</u> Let $t \geq 1$. Assume that for all $x, y \in \mathbb{R}$, $\Delta_t(x,y) \geq 0$. We need to show that for all $x, y \in \mathbb{R}$, $\Delta_{t+1}(x,y) \geq 0$.

Let $pS^1(\hat{Q}_{1,t})$ denote the marginal sampling probability of Alt. 1 in period t. We have:

$$f(\max(x,y))f_{1,t+1}(\min(x,y))$$

$$= f(\max(x,y))f(\min(x,y))E\left[pS^1\left(\hat{Q}_{1,t}\right)\right]$$

$$+ \left(1 - pS^1(\min(x,y))\right)f_{1,t}(\min(x,y))f(\max(x,y))$$

$$= f(y)f(x)E\left[pS^1\left(\hat{Q}_{1,t}\right)\right] + \left(1 - pS^1(\min(x,y))\right)f_{1,t}(\min(x,y))f(\max(x,y)).$$

The last equality follows from the fact that $f(\max(x,y))f(\min(x,y))= f(y)f(x)$.

We also have:

$$f_{1,t+1}(x)f(y) = f(y)f(x)E\left[pS^1\left(\hat{Q}_{1,t}\right)\right] + \left(1 - pS^1(x)\right)f_{1,t}(x)f(y),$$

Then,

$$\Delta_{t+1}(x,y) = \left(1 - pS^1(\min(x,y))\right)f_{1,t}(\min(x,y))f(\max(x,y)) - \left(1 - pS^1(x)\right)f_{1,t}(x)f(y).$$

By application of the induction hypothesis, we have

$$\Delta_{t+1}(x,y) \geq \left(1 - pS^1(\min(x,y))\right)f_{1,t}(\mathrm{x})f(\mathrm{y}) - \left(1 - pS^1(x)\right)f_{1,t}(x)f(y)$$

$$= \left(pS^1(x) - pS^1(\min(x,y))\right)f_{1,t}(\mathrm{x})f(\mathrm{y}).$$

The fact that $pS^1\left(\hat{q}_{1,t}, \ldots, \hat{q}_{K,t}\right)$ is increasing in $\hat{q}_{1,t}$ implies that the marginal sampling probability $pS^1\left(\hat{q}_{1,t}\right)$ is non-decreasing in $\hat{q}_{1,t}$. This implies, in turn, that $pS^1(x) - pS^1(\min(x,y)) \geq 0$. From this and the above equation, we have $\Delta_{t+1}(x,y) \geq 0$. This concludes the proof of the inductive step.

We have proven the base case and the inductive step. The principle of mathematical induction implies that, for all $t \geq 1$ and for all $x, y \in \mathbb{R}$, $\Delta_t(x,y) \geq 0$. That is, for all $t \geq 1$, $f \succ_{TP_2} f_{1,t}$. *QED.*

We assumed that the initial quality estimate for Alt. N was a random draw of the payoff distribution f common to Alt. 1 and Alt. N. In other words, $f_{N,\tau} = f$. This and Lemma 2 thus imply the following corollary:

Corollary 3: We have: $f_{N,\tau} \succ_{TP_2} f_{1,\tau}$.

This result implies that $\hat{Q}_{1,\tau}$ is lower than $\hat{Q}_{N,\tau}$ per the usual stochastic order. This is formalized as follows:

Corollary 4: For all $q \in \mathbb{R}, P[\hat{Q}_{1,\tau} < q] \geq P[\hat{Q}_{N,\tau} < q]$.

Proof: The definition of the $\succ_{TP_2}$ relation, Corollary 3 and Theorem 2.2 in Karlin and Rinott (1980, p. 477) imply that for any bounded and increasing function φ,

$$\int \varphi(x) f_{1,\tau}(x)dx \leq \int \varphi(x) f_{N,\tau}(x)dx\,.$$

This is a standard characterization of the usual stochastic order and thus implies that for all $q \in \mathbb{R}$, $P[\hat{Q}_{1,\tau} < q] \geq P[\hat{Q}_{N,\tau} < q]$ (to see this, it is enough to take φ equal to an indicator function equal to 1 for values lower than q and equal to 0 otherwise). *QED.*

We are now ready to conclude the proof of Theorem 1. Let $t \geq 1$. We have:

$$P[\hat{Q}_{N,\tau} > \hat{Q}_{1,\tau}] = \int P[q > \hat{Q}_{1,\tau}]\, f_{N,t}(q)dq.$$

Corollary 4 states that for all $q \in \mathbb{R}, P[q > \hat{Q}_{1,\tau}] \geq P[q > \hat{Q}_{N,\tau}]$. This implies

$$P[\hat{Q}_{N,\tau} > \hat{Q}_{1,\tau}] \geq \int P[q > \hat{Q}_{N,\tau}]\, f_{N,t}(q)dq.$$

Note that $\int P[q > \hat{Q}_{N,\tau}]\, f_{N,t}(q)dq$ is the probability that when picking two independent observations of a random variable with (continuous) distribution $f_{N,t}$ the first observation is lower than the second observation. This is equal to 0.5. Therefore, we have $P[\hat{Q}_{N,\tau} > \hat{Q}_{1,\tau}] \geq 0.5$. This immediately implies $P[\hat{Q}_{N,\tau} > \hat{Q}_{1,\tau}] \geq P[\hat{Q}_{N,\tau} < \hat{Q}_{1,\tau}]$. *QED.*

SUPPLEMENTAL ONLINE MATERIAL: METHODOLOGICAL DETAILS ON THE EXPERIMENTAL STUDY

In all online sessions, the task discussed in the paper was the first of a series of 2 or 3 short experimental tasks. In a typical session, participants worked for about 5 minutes. They spent on average about 1 minute and 13s to make the 11 choices and enter their quality estimates at the end of the choice sequence.

Ensuring data quality

We removed the data of 4 participants who missed a choice (probably because of letting more than the maximum of 90s elapse before making a choice). We also removed the data of 65 participants whose final quality estimates were outside of the range of observed payoffs (0 to 130). We believe that these responses signal that the participants did not pay adequate attention to the observed payoffs. Our main results remain almost the same if we consider all 834 participants who completed the experiment. The proportion of participants for whom $\hat{Q}_{N,F} > \hat{Q}_{1,F}$ is 0.58, 95% CI [0.55,0.62]. Among the 793 participants with different estimates for the two alternatives, the proportion of participants with $est_1 < est_N$ is 0.53, 95% CI [0.50, 0.57].

Game Instructions

Your goal is to collect as many points as possible by making a series of choices between two or more buttons. At the end of the game, the points you accumulated will be converted to US dollars at a rate of 10000 points = 1 USD. Behind each button, there is a large stack of chips with values marked on them. Some buttons can be BETTER than others:

- BETTER buttons cover stacks of chips with HIGH values.
- WORSE buttons cover stacks of chips with LOW values.

On every round, you have to select one button by clicking on it with the mouse. A chip is randomly picked from the stack behind the button. The button you selected becomes ORANGE and a number is DISPLAYED on the button: This is the VALUE of the chip that was picked from the stack. (if the number you see is X, you will collect X points; if X is negative, you give back X points). After you collect the points from your chip, the chip is put back in the stack from which you took it and you move to the next round.

You start with an initial endowment of 500 points. There are about 11 rounds. The values you see on the buttons BEFORE STARTING are the values of ONE randomly picked chip from the corresponding stack. Please note that you have up to 90 seconds to make each choice.

Instructions for eliciting the quality estimates

For each button, please indicate the number of points you expect to obtain if you were to select it one more time. If your answer does not make sense given what you have seen, your HIT might be rejected. For example, if the numbers you have seen are in the hundreds, but you respond in the thousands, your hit will likely be rejected.

SUPPLEMENTAL ONLINE MATERIAL: ADDITIONAL ANALYSES OF EXPERIMENTAL DATA

To evaluate the contribution of information sampling to the asymmetry in quality estimates, we complement the recency-weighted analysis presented in the body of the paper by two additional analyses.

Last sampled payoff

Let $L_{1,F}$ denote the last observation for Alt. 1 and $L_{N,F}$ denote the last observation for Alt. *N*. Adaptive sampling implies that the last observations of the alternatives will tend to favor Alt. *N* rather then Alt. 1. We find that this is the case (see Table 1): the last observation for *N* is higher than the last observation for Alt.1 in 58% of the cases. Besides, the last observation for Alt. 1 is frequently below the mean of the payoff distribution (in 66% of the cases). This underestimation tendency is stronger than that for Alt. *N* (underestimation in 55% of the cases).

Average sampled payoff

It is also possible to characterize the asymmetry in terms of the average sampled payoff for each alternative. Let $Av_{1,F}$ denote the average observation for Alt. 1and $Av_{N,F}$ denote the average observation for Alt. *N*. The average sampled payoff for Alt. 1 tends to be lower than the average sampled payoff for Alt. *N* (in 53% of the cases).

$P[L_{1,F} < L_{N,F} \vert L_{1,F} \neq L_{N,F}]$	.58 [.55,62] (n=763)
$P[L_{1,F} < 65 \vert L_{1,F} \neq 65]$	.66 [.63,.69] (n=766)
$P[L_{N,F} < 65 \vert L_{N,F} \neq 65]$	.55 [.51,.58] (n=762)
$P[Av_{1,F} < Av_{N,F} \vert Av_{1,F} \neq Av_{N,F}]$	.53 [.50,.57] (n=769)
$P[Av_{1,F} < 65 \vert Av_{1,F} \neq 65]$	.54 [.51,.58] (n=767)
$P[Av_{N,F} < 65 \vert Av_{N,F} \neq 65]$	.51 [.48,.55] (n=761)

TABLE S1: Summary of the final quality estimates and sampled information. Based on 769 participants. 95% CI are in brackets. The number of participants with estimates different from the comparison value is indicated in parentheses.