Mental Models and Transfer Learning*

Ignacio Esponda Emanuel Vespa Sevgi Yuksel

Behavioral biases may persist in the presence of transparent feedback for at least two reasons. One reason is that confidence in some initial misconception or incorrect mental model can lead both to sluggish response and inattentiveness to new information. We recently documented this phenomenon for the case of base-rate neglect (Esponda, Vespa & Yuksel (2022), henceforth EVY). A second reason is that, even if people learn to correct their mistakes over time in a particular setting, they may still fail to learn general principles that are applicable to other, related settings. So, when facing a similar situation but with different parameter values or slight modifications to other features, people may not be able to entirely benefit from a previous learning experience. That is, learning in one environment might not *transfer* to another.[1]

In this paper, using a laboratory experiment, we investigate the extent to which learning can be transferred between related problems in the context of a simple updating task. Our focus is on settings in which deviations from Bayesianism are very stark. Specifically, in these settings participants state posterior beliefs that are incompatible with *any* (correct or incorrect) interpretation of the signal. That is, participants violate a simple updating principle, which requires the range of posterior beliefs to include the prior. In a binary

*We are grateful for financial support from the UCSB Academic Senate. Esponda: UC Santa Barbara, Economics Department, iesponda@ucsb.edu. Vespa: UC San Diego, Economics Department, evespa@ucsd.edu. Yuksel: UC Santa Barbara, Economics Department, sevgi.yuksel@ucsb.edu.

[1] Erev & Haruvy (2016) in their review of the literature on learning from feedback define *transfer* "as the effect of learning in one task on the performance of another task".

setting, this boils down to updating positively after a positive signal and negatively after a negative signal. Violation of the principle indicates subjects do not properly account for the prior (e.g., base-rate neglect).[2] Our baseline is one in which subjects may initially fail to satisfy the principle, but after substantial feedback eventually adjust. The challenge is to then present them with a different parameterization that is also known to generate the same deviation from Bayesianism. Thus, by examining cross transfer of this principle, we are checking whether subjects who adjust their initial answers learn to account for the prior in the new parameterization, even if they haven't learned how to compute correct posteriors.

The initial parameterization we implement is well-known to generate incorrect answers from a majority of participants. In particular, initial answers typically fail the updating principle described above. After we collect initial answers, we provide people with experience and repeated feedback under the initial parameterization, inducing learning (i.e., movement towards the Bayesian answer) for this particular parameterization. Finally, we change the parameter values of the problem. The final parameterization is one that also known to generate incorrect answers, as the initial parameterization. This change lets us investigate the extent to which learning was tailored specifically to the original context (i.e., if subjects still fail the principle under the new primitives) or whether people learned more general updating principles that help with other, related decisions.

A broad literature in psychology and economics studies behavioral spillovers: how past experiences impact future behavior.[3] Our focus is specifically on transfer learning: how experience in one setting can generate insights that can help an agent make better decisions in a different, but related, setting. Most work in economics stressed and studied the question of transfer learning in the context of games.[4] As pointed out by Fudenberg & Kreps (1988),

[2]Under or over reaction to the signal, for example, (the former is a commonly documented bias in updating, see Benjamin (2019) for a review), do not generate a violations of this principle.

[3]See Erev & Haruvy (2016) and Dolan & Galizzi (2015) spillover paper for a review.

[4]In the case of decision problems, there is a literature in psychology that studies transfer of behavior within the clicking paradigm. Subjects click one of two buttons and observe a payoff in each of several periods, but are not given any information about the way in which payoffs are generated. After substantial experience with one fixed environment (which shifts behavior to the payoff optimal option), subjects face an

because it is unreasonable to expect the exact same game to be played over and over again, learning foundations for equilibrium are more credible in situations where people are able to transfer their learning across similar games.

The experimental literature has studied cases in which the rules of the game are fixed, but the parameters of the game change within a session, and cases in which subjects play two different but related games.[5] An early example of the latter is Kagel (1995). That paper investigates cross-game learning in common value auctions, specifically between the first-price, sealed bid and the English auctions. He finds that experience in the former improves performance in the latter, but not the other way around.

In general, learning across games can happen in two ways by impacting equilibrium selection and strategic sophistication. There is mixed evidence on whether past experiences influence equilibrium selection in a new environment.[6] Evidence on transfer of strategic sophistication is also limited. Li & Schipper (2020) studies strategic sophistication in a class of persuasion games that involve different levels of complexity and finds weak evidence of transfer learning. A related set of papers focus on epiphany (sometimes referred to as Eureka) learning. Dufwenberg, Sundaram & Butler (2010) focus on the games of 6 or 21 to study whether subjects learn to play optimally (allowing them to guarantee a win). They find some (but limited) evidence that playing the simple game of 6 first helps subject play more optimally in the large game of 21.[7]

Overall, while transfer of insights across games is often difficult, in a synthesis of the

environment in which clicking the same buttons leads to different payoffs. A typical pattern is that repeating a choice that was successful under the initial parameterization can temporarily lead to worse payoffs in a new parameterization, but that eventually most subjects learn to adjust. For further details on this literature, see Erev & Haruvy (2016).

[5]As an illustration of the former approach, Grimm & Mengel (2012) presents subjects with three-action two-player normal-form games in which subjects gain experience for a fixed parameterization, but face several parameterizations within a session.

[6]Duffy & Fehr (2018) find limited evidence of spillovers in equilibrium selection between related games. Peysakhovich & Rand (2016) show that behavioral norms that support cooperation are carried over into atypical situations beyond the reach of the institution.

[7]Huck, Jehiel & Rutter (2011), Mengel (2012), and Grimm & Mengel (2012) are other examples studying learning across different games. For a survey of learning in games, including discussions of cross-game extrapolation, see Fudenberg & Levine (2009).

early work in this literature, Cooper & Kagel (2003) document that cross-game learning is more likely to happen when there is meaningful context (relative to an abstract setting) and when the role of a player in a game is taken on by a team of human participants.

Properly summarizing the literature on cross-game learning is beyond the scope of this paper, but understanding why transfer works or does not work across games is challenging because several aspects change as we move from one game to the other. For instance, it could be that similarities between games (with different rules and parametrizations) are challenging for people to identify. But, in addition, it could also be difficult to anticipate the degree to which other players identify these similarities. That is, a player who sees two games similarly might not expect other players to behave similarly in them.

The environment we study in this paper abstracts from these challenges. We focus on a very simple decision task in which beliefs over others' actions play no role. The rules in our environment are kept unchanged throughout the session. In the updating problem we study the only aspect that changes from one environment to the other is the parameterization. The parameterizations we use, however, are specifically selected to learn whether insights transfer. For both parameterizations it is well established in the literature that most subjects completely ignore the prior. Hence, even if we keep the rules unchanged, we can study whether subjects who discover to adjust in the right direction with feedback (i.e., get an insight that their initial answer was not correct) adjust when they face another parameterization that absent any previous experience is known to lead to the same mistake.

The updating task

The updating task and data comes from the experiment presented in EVY and is as follows.[8] There is a binary state of the world, success or failure. There is also a signal that is informative about the state of the world. In the initial parameterization, the probability of success is

[8]The data comes from EVY and will be available at https://doi.org/10.3886/E183963V1.

.15 and the reliability of the signal is .8, meaning that the probability of a positive (negative) signal conditional on the project being a success (failure) is .8. We gave this information to subjects and asked them to report the probabilities that the state is a success conditional on the signal being positive and conditional on the signal being negative. We refer to the initial answers as the R1 answers. Subjects have all the necessary information to provide the correct answers: applying Bayes' rule, the correct answers are .41 and .04, respectively.

We then provided feedback to subjects by repeatedly sampling from the state space. In each round, we sampled a state and then drew a signal conditional on the state. We informed subjects of the signal and subsequently of the true state. We did this for 200 rounds. We refer to the answers in the last round as the R200 answers. We then changed the primitives of the updating task to a probability of success of .95 and a reliability of .85, and asked them one final time to report the conditional probabilities. We refer to the answers in the new parameterization as the R1' answers. For more details of this experiment, including instructions, information about the subjects, and payments, see EVY.[9]

In EVY, we focused on the evolution of answers from rounds 1 through 200. We found that learning was hindered relative to a control treatment where subjects were not given any information about the primitives of the problem. Then, aided by a series of additional treatments, we concluded that confidence in initial responses led to both a sluggish response and inattentiveness to new information, resulting in an incomplete learning experience.

In EVY, we documented that the most common mistake in R1, made by slightly more than half the subjects, is to say that the probabilities are .8 and .2. This is known as (perfect) base-rate neglect (Kahneman & Tversky 1972), because the answers completely neglect the prior. But then, over the course of 200 rounds, most subjects adjust their responses, indicating that they learn that their initial responses were incorrect. In addition, for the final parameterization, subjects are much less likely to report the corresponding

[9]The experiment uses the strategy method so that for each answer (R1, R200, R1') we obtain from each subject posteriors for both possible signals.

base-rate neglect answers, .85 and .15.[10] But this begs the question of whether subjects are learning principles on how to update from the first task that they are then applying to the other. For example, it could be that they learn they should account for the prior information in their responses, but they are not sure how or why.

In this paper we focus on a specific principle that is important in these kind of updating tasks. Specifically, in binary tasks with informative signals, an important concept is that the posterior should be higher than the prior after a positive signal and lower than the prior under a negative signal. A subject who fully ignores the prior (perfect base-rate neglect) clearly violates this principle. (For example, in the first parameterization, both .8 and .2 are above the prior of .15!). But there are many other responses that also violate this principle.

To assess the extent of learning and transfer learning of this basic updating principle, we will focus on transitions between R1, the first round of the first parameterization, R200, the last round of the first parameterization, and R1', the first (and only) round of the second parameterization of the updating task. We refer to a response as consistent if it satisfies the updating principle (that is, the direction of updating is correct, even if the actual numbers are not), and inconsistent otherwise.

Results

As demonstrated in the literature on base-rate neglect (see Benjamin (2019) for a survey), very few subjects provide the correct response in R1 (only 3 out of 64 subjects). Because Bayesian updating is known to be hard, we focus on a simple updating principle that is

[10]In EVY there is a control treatment in which subjects are not told the specific values of the initial parameterization. That is, when they provide the R1 answer they do not know that the prior is .15 and that the accuracy of the signal is .8. However, they do get the same 200 rounds of feedback. Thus, subjects in the control treatment can learn only from feedback but cannot provide an initial base-rate neglect answer because they do not know the base rate. For the final primitives, subjects in this control treatment *are* provided with the primitives. Thus, subjects did not have the opportunity to make an initial base-rate neglect mistake that they subsequently learn to correct over the course of 200 rounds of feedback. As reported in EVY, subjects in this control treatment select the base-rate neglect answers in R1' more frequently than subjects in the main treatment.

Table 1: Transition between consistent and inconsistent responses

		R200		R1'		Total
		Consistent	Inconsistent	Consistent	Inconsistent	
R1	Consistent	8	2	5	5	10
	Inconsistent	23	31	12	42	54
Total		31	33	17	47	64

Notes: (i) R1: Answers for initial parameterization (prior: .15; accuracy of signal: .8). (ii) R200: Answers for initial parameterization after 200 rounds of feedback. (iii) R1': Answers for final parameterization (prior: .95; accuracy of signal: .85). (iv) Consistent: The subject provides answers such that the posterior is higher after a positive signal and lower under a negative signal.

necessary for Bayesian updating but only requires people to update positively after positive signals and negatively after negative signals.

Table 1 presents the main results, documenting the transitions between responses that are consistent and inconsistent with the simple updating principle out of a total of 64 subjects who participated in the experiment. In R1, only 10 out of 64 subjects update in the right direction (i.e., are consistent with the updating principle). But, by R200, this number increases to 31 subjects. As shown by the table, this is mostly due to subjects who failed the principle in R1 but now satisfy it by R200.

The increase in the number of subjects who update in the right direction from R1 to R200 is a result of the experience and feedback that subjects receive, but it does not necessarily indicate that subjects understand that they must update positively for positive signals and negatively for negative signals. In fact, as the table indicated, only 17 subjects abide by the principle in the new parameterization, R1'. In fact, of the 23 subjects who learned to update in the right direction from R1 to R200, only 12 are able to update in the right direction with the new parameterization, in R1'. This is striking, because the nature of the problems in R1 and R1' are very similar, and many subjects become proficient in solving the first problem. But then, when faced with the second problem, only half of these subjects are able to update in the right direction. Moreover, this proportion is an upper bound to the proportion of subjects who learn to update in the right direction, since some of those subjects may be updating in the right direction by chance in R1'.

Finally, compared to subjects in the control treatment described earlier, subjects in R1' are more likely to update in the right direction (17 out of 64, compared to only 4 out of 64 in the control treatment). This comparison suggests that experience with a similar problem can triple the amount of people who update in the right direction in a new problem.

Discussion

Understanding the persistence of behavioral biases over time is important to predict which biases are resilient to learning, and therefore more likely to play an important role in explaining economic phenomena. Understanding the reasons why biases are persistent is also important to devise policy interventions to mitigate these biases. For example, in EVY we established that base-rate neglect persists in large part because people are confident in their initial answers, and this confidence leads both to a sluggish response and inattentiveness to new information. An implication is that biases may persist even in information-rich settings, and that to mitigate these biases we need to find ways for people to engage with the data.

A different but related reason why biases may persist is that people are unlikely to face the same exact problem over and over again. Instead, people may face variations of a specific type of problem. Anecdotal evidence suggests that transferring learning from one problem to another is challenging. For example, students sometimes do poorly in exams when we make slight changes to a problem they have already seen.

In this short paper we have briefly reviewed some of the evidence for transfer learning in the economics literature and presented new findings in the context of a simple updating task. While we find that learning in one task triples the amount of subjects who update in the right direction, we also find that about half of the subjects who learn to update in the right direction in the context of one problem do not benefit from this learning experience when facing a new problem. While more research is required to understand the reasons for these findings, one conjecture is that many people engage in a fairly unsophisticated form of learning, where they simply try to find in the data what works best in a particular situation. For example, in the first parameterization of the updating task discussed in this paper, some subjects are likely to be responding to the frequencies obtained in the data, rather than trying to learn something more fundamental about the updating task.

More research, however, is needed to understand why transfer learning is difficult and also the types of situations where we can expect people to struggle most. For example, it seems plausible that if people knew they would be facing repetitions of the same problem but with different parameterizations, then they would have an incentive to engage with data in a different way, perhaps trying to draw some general lessons from their experience. Of course, it is not always the case that people anticipate facing similar situations in the past, and there is also the issue of discounting and impatience. In any case, our results raise questions that we hope motivate further work in this area.

References

Benjamin, D. J. (2019), 'Errors in probabilistic reasoning and judgment biases', *Handbook of Behavioral Economics: Applications and Foundations 1* **2**, 69–186.

Cooper, D. J. & Kagel, J. H. (2003), 'Lessons learned: generalizing learning across games', *American Economic Review* **93**(2), 202–207.

Dolan, P. & Galizzi, M. M. (2015), 'Like ripples on a pond: behavioral spillovers and their implications for research and policy', *Journal of Economic Psychology* **47**, 1–16.

Duffy, J. & Fehr, D. (2018), 'Equilibrium selection in similar repeated games: experimental evidence on the role of precedents', *Experimental Economics* **21**(3), 573–600.

Dufwenberg, M., Sundaram, R. & Butler, D. J. (2010), 'Epiphany in the game of 21', *Journal of Economic Behavior & Organization* **75**(2), 132–143.

Erev, I. & Haruvy, E. (2016), 'Learning and the economics of small decisions', *in The handbook of experimental economics, Volume 2, J Kagel and A. Roth eds.* pp. 638–700.

Esponda, I., Vespa, E. & Yuksel, S. (2022), 'Mental models and learning: The case of base-rate neglect', *Working Paper* .

Fudenberg, D. & Kreps, D. (1988), 'Learning, experimentation and equilibrium in games', *Working Paper* .

Fudenberg, D. & Levine, D. K. (2009), 'Learning and equilibrium', *Annual Review of Economics* **1**(1), 385–420.

Grimm, V. & Mengel, F. (2012), 'An experiment on learning in a multiple games environment', *Journal of Economic Theory* **147**(6), 2220–2259.

Huck, S., Jehiel, P. & Rutter, T. (2011), 'Feedback spillover and analogy-based expectations: A multi-game experiment', *Games and Economic Behavior* **71**(2), 351–365.

Kagel, J. H. (1995), 'Cross-game learning: Experimental evidence from first-price and english common value auctions', *Economics Letters* **49**(2), 163–170.

Kahneman, D. & Tversky, A. (1972), 'On prediction and judgement', *ORI Research Monograph* **12**(4).

Li, Y. X. & Schipper, B. C. (2020), 'Strategic reasoning in persuasion games: An experiment', *Games and Economic Behavior* **121**, 329–367.

Mengel, F. (2012), 'Learning across games', *Games and Economic Behavior* **74**(2), 601–619.

Peysakhovich, A. & Rand, D. G. (2016), 'Habits of virtue: Creating norms of cooperation and defection in the laboratory', *Management Science* **62**(3), 631–647.