Effect of Conditional Feedback on Learning

Navid Ghaffarzadegan
navidg@gmail.com
Rockefeller College, the State University of New York at Albany, USA

Abstract

Formal studies of decision threshold learning assume full feedback conditions, that is, no matter what the decision is (positive or negative), the feedback will be provided. However, in the real world feedback may be conditional on the decision made. For example, in college admissions decisions, there is no feedback available for the students who are not admitted. In this paper, we investigate how conditional feedback can result in biased decisions. First, based on signal detection theory, a dynamic model of threshold learning is proposed. Then the model is adjusted to examine effects of conditional feedback on learning and decision making. Finally, the model is used to replicate some empirical findings. The results suggest conditional feedback can be a barrier to learning. Further, this study warns about problems with the current assumption of full feedback condition in most dynamic decision-making studies.

Keywords: threshold learning, conditional feedback, signal detection theory

1. Introduction

In a dynamic decision making environment, there are many barriers to learning from feedback. In fact, not all feedback is clear and understandable. Complexity of the environment (Gonzalez 2005), misperception of delays (Rahmandad et al. 2007, Rahmandad 2008), feedback asymmetry (Denrell and March 2001), the existence of noise in feedback (Bereby-Meyer and Roth 2006), and problems of mental models (Senge 1996) make it very difficult to learn from feedback. As results, sometimes, people ignore feedback and sometimes they misperceive it (Sterman 1989a, Sterman 1989b).

Most studies on learning from feedback are founded on a common theme: a decision maker (individual, group, or organization) makes a decision and receives a payoff (with or without delay); then the question is whether or not the decision maker is capable of interpreting and learning from the information. While the formal assumption is that information on payoff always exists and is clear, i.e. full feedback condition (e.g. Erev 1998), few studies have examined other assumptions about feedback.

Full feedback is not common in the real world. For example, a human resources manager will know true performance of a candidate if he decides to recruit the applicant. A police officer, who decides to search a suspect, will know whether or not the suspect is a drug dealer; otherwise he will not be informed about the true status of the suspect. This is the same for the cases of admission decisions in universities, strategic decisions in companies, most medical decisions, etc. In all of these situations, and in many other real world conditions, there is dependency between one's decision and whether or not he receives a clear feedback (Elwin et. al 2007). Usually for positive decisions (e.g. admitting a candidate, or deciding to search a suspect), we receive feedback, otherwise we lack a clear feedback, or at least it is very difficult to interpret the results of negative decisions. This kind of feedback is referred as conditional (or selective) feedback.

Studying effects of conditionality of feedback can give a new explanation about barriers to learning in the real world. In one of the few studies about conditional feedback, Elwin et. al (2007) investigate empirically the effects of conditional feedback on decision making. While observing that people underestimate the base rate (the ratio of signals to total observations), they argue that people assume their negative decisions, for which they do not receive feedback, are true. While their results are very valuable and provocative, they have not studied the effects of base rate, accuracy of signal detection, and initial threshold on the final results.

In the current study, we focus on signal detection framework as a classical judgment and decision making framework, and expand the few studies of conditional feedback by building a simulation model and observing effects of different parameters on biases. We examine the dynamics of learning and the effects of conditional feedback on decision results. This new insight is crucial as it can warn about the underestimation of one of the common assumptions in dynamic decision making and learning studies.

In following, based on a brief review of signal detection framework (section 2), we build a simulation model of full feedback (section 3) and conditional feedback systems and examine effects of different ways of coding negative decisions on learning optimal thresholds (section 4). Then, using data from a published empirical work in this area, we replicate the results for different scenarios by the developed model (section 5). Finally we discuss possible implications of simulation results (section 6).

2. Signal detection framework

From signal detection perspective (Green and Swets 1966; Swets 1991; Swets, Dawes and Monahan 2000, Arke and Mellers 2002), decision makers try to differentiate signals from noise (e.g. guilty from innocent persons, capable from incapable candidates). In order to do that, they make judgments based on different cues, and make decisions based on those judgments. A police officer judges how suspicious a suspect appears, and then decides if the person should be searched or not. A human resources manager judges how capable an applicant is, and then makes a decision about him. An admission committee judges a candidate based on their perception of the candidate's capability, and then decides whether or not to offer admission.

In the real world, making proper decisions is very difficult because evidence is often ambiguous, and there is uncertainty in the environment (Hammond 1996, Stewart 2000). This means we are not always able to differentiate signals from noise based on our judgment, and errors will be made. For example, the police officer may search some innocent people, and may let some guilty persons go.

Let's focus on the police officer situation. The probability distribution in Fig. 1 shows what might occur over an infinite number of trials from signal detection perspective. The Y-axis is the chance that the value of the random variable *x* (officer's judge) could arise from a distribution of innocents or a distribution of guilty persons. The distributions are normal, and guilty persons (signals) are, on average, more culpable than innocent persons (noise). As the figure shows, due to uncertainty, the distributions overlap.

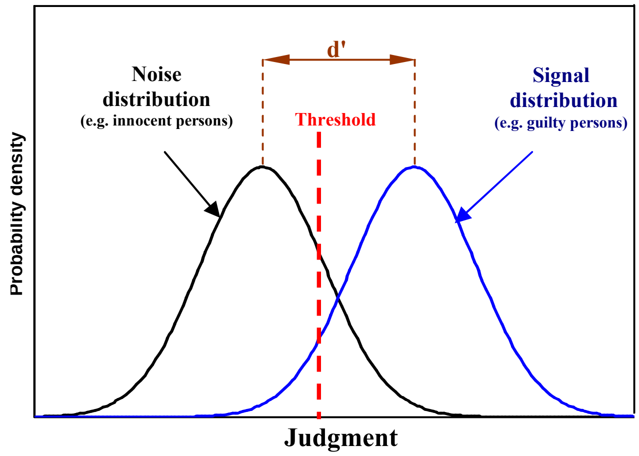


Fig.1: Distribution of noise and signal and an example of decision threshold location

A common assumption is that decision makers use a threshold (cutoff) in making a decision based on their judgment. So, for any *x* more than their threshold they decide "yes" (e.g. search, recruit, …), and for any *x* less than their threshold, they make a "no" decision (e.g. not search, reject, …).

Therefore, in any yes-no decision making situation, there are four possible decision outcomes. You can say "yes" and be right or wrong or you can say "no" and be right or wrong. We can name these outcomes as true and false positives and true and false negatives. As Fig.2 shows, a police officer can decide to search a person (a positive decision) and the person maybe guilty (true positive) or innocent (false positive). The police officer can also decide not to search the person (a negative decision). And again the person can be guilty (false negative) or innocent (false positive). Thus, there are two kinds of errors: false positives and false negatives.

Decision

State of the world	NO (not search)	YES (search)
YES (guilty)	False negative	True positive
NO (innocent)	True negative	False positive

Fig.2 : Four possible outcomes

An important point is that different threshold locations impose different error rates, and as the probability of one error decreases, the probability of the other error increases (see figure 1). Unless the distributions can be moved further apart, it is impossible to simultaneously decrease both errors by changing the threshold. This means without increasing in d' (the ability of the observer to discriminate signals from noise), changing the threshold does not decrease the uncertainty.

In this framework, the ratio of positive decisions to total trials is called selection rate (e.g. if 50 percent of people are selected for searching, selection rate is 0.5). On the other hand, the ratio of number of "Yes" in the state of the world to the total number of trials is called base rate (e.g. if 50 percent of people are guilty, base rate is 0.5).

Obviously, there is an optimal location for threshold, which depends on decision makers' value system. For each cell in Fig. 2, each decision maker can assign a different value, and the difference in the value systems results in different payoffs and, therefore, different optimal thresholds.

3. Full Feedback Model

In a dynamic decision making environment, we get more information as we make more decisions. The information may help us to learn more about the environment and to amend our decision rules. From a signal-detection perspective, we learn the optimal threshold. We can also learn about cues and cue weights. In this paper, we focus on the first part – threshold learning.

Although, the optimal threshold for any normal distribution of signals and noise can be calculated, we may doubt if people can discover this threshold. People do not know about theories of decision sciences and are not always rational and coherent in decision making, but they learn through experiments. Thus, we can assume that a person may require many trials to learn a threshold. In each trial, he will receive information about his performance and will try to correct his threshold, in order to increase the performance. For example, a human resources manager will find what are the minimum characteristics of an applicant, (e.g. education and experience) to be capable of doing their desired task. Fig. 3 shows a diagram of threshold learning, as well as its implication in a signal detection framework.

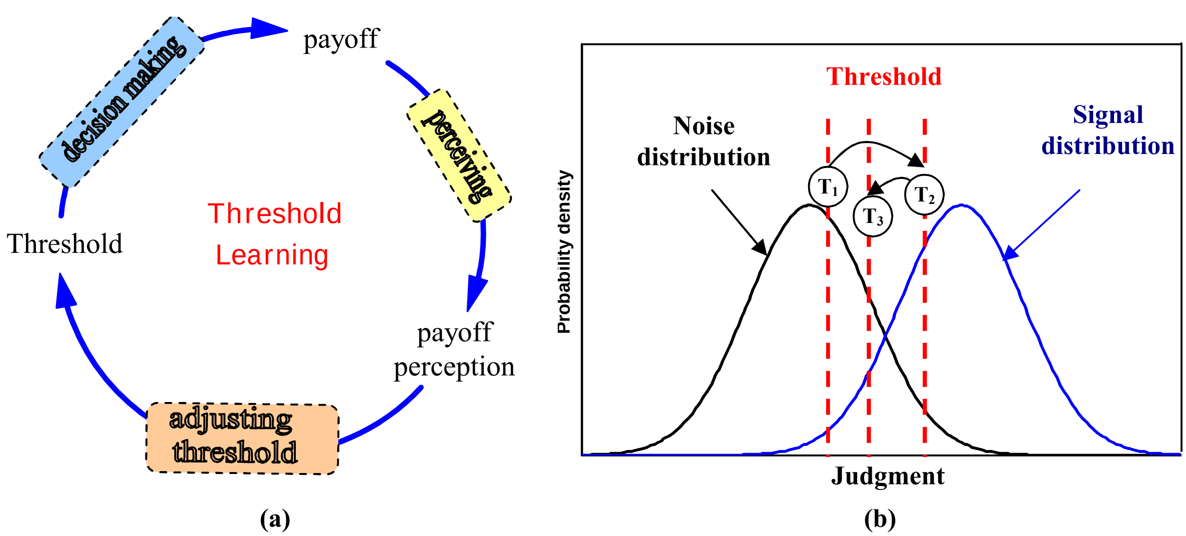


Fig. 3: Threshold learning in a full feedback condition: a) a dynamic model b) in signal detection framework

As it is depicted in Fig.3, from signal detection perspective, there can be three main processes in threshold learning: decision making, perceiving the outcomes, and adjusting (correcting) threshold. Actually, in this paper, we assume people use their current threshold as an anchor and adjust it to a new level using the new piece of information they have received (Tversky and Kahneman 1974). Psychologically, it means that we have an answer (a threshold) in our mind and we try to shift it toward the best answer through experiments. This assumption is consistent with many studies of decision science on anchoring and adjustment (e.g. Epley and Gilovich 2001), as well as many system dynamics models of decision making (e.g. Sterman 1989.b). In following, we model each of these three processes for a full feedback condition.

3.1. Decision Making Process

From signal detection perspective, an experimenter has a threshold and makes his decision by comparing his observation with the threshold. If the observation is greater than the threshold he judges it as a signal, otherwise as a noise. We assume the existence of a single threshold which can be formulated by an if-then-else decision rule:

$$d=0 \quad \text{if } x<threshold$$
$$d=1 \quad \text{if } x\geq threshold \qquad \text{(eq. 1)}$$

whereby d represents a decision, and is 1 for positive decisions and zero for negative decisions. x is the subject's judgment. Let us show the true state of the world by Q which will be either 1 or zero. By comparison of d and Q we can find the payoff. The following formula does the same:

$$Payoff\ (Q,d) = (1-Q)*(1-d)*V_{tn} + Q*(1-d)*V_{fn} + (1-Q)*d*V_{fp} + Q*d*V_{tp} \qquad \text{(eq. 2)}$$

whereby payoff will be equal to V_{tn}, V_{tp}, V_{fn} and V_{fp}, called values, in true negative, false negative, true positive, and false positive decisions respectively.

In this paper we assume $V_{tn}= V_{tp}=1$ and $V_{fn}= V_{fp}=-1$. This symmetry in values helps us to examine simulation results much easier. As a result of symmetry in values, in base rate equal to 0.5, when threshold is equal to the optimal threshold, the selection rate is equal to the base rate. However, this simplification is, only, used to make the paper easier to follow.

3.2. Perceiving results

Different learning algorithms can be assumed in this stage. Basically, in most of the algorithms, we try to increase our payoff, by changing the threshold in different directions and interpreting the results.

Here, we assume a more intuitive process of learning from results: as a subject gets information about the true value of the previous observation (Q), he can judge the payoff shortfall. Payoff shortfall is the difference between the maximum possible payoff for Q and the current payoff. We can formulate the process as following:

$$Payoff\ shortfall= maximum\ possible\ payoff\ (Q) - payoff\ (Q, d) \qquad \text{(eq. 3)}$$
$$maximum\ possible\ payoff\ (Q)= V_{tn}+ Q*(V_{tp}-V_{tn}) \qquad \text{(eq. 4)}$$

Maximum possible payoff is the maximum value that a person can receive from a decision, and as we assumed higher values for correct decisions it can be calculated by a linear function of V_{tp} and V_{tn}.

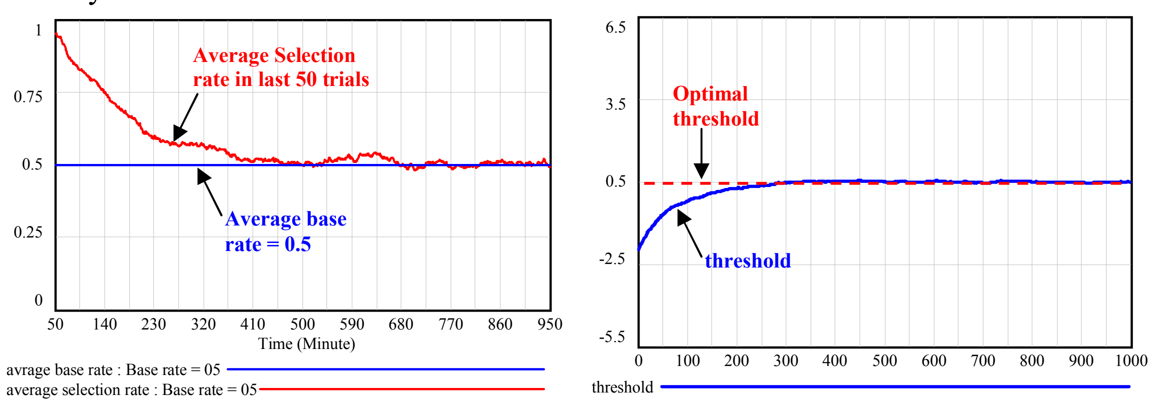
3.3. Adjusting threshold

Knowing that we have made a wrong decision (payoff shortfall > 0), the model assumes that the decision threshold will be amended toward the observation. In the real world, one observation can not change the whole assumptions and the subject's mental model, but, in fact, it takes time for a person to change his threshold. Considering such a process, we can say:

$$\text{Change in threshold} = (x\text{-threshold})/\tau \qquad (\text{if } \textit{payoff shortfall} > 0) \qquad (eq.\ 5)$$

where τ is the time to change threshold, which can depend on many factors, such as the personal characteristics of the decision maker and his confidence, and the latter can change dynamically in the system.

So far, in addition to the threshold adjustment loop (threshold→change in threshold→threshold) we have introduced one simple loop that formulates a full feedback system (threshold→decision→ payoff→payoff shortfall→change in threshold→threshold). As it is clear from the formulation, the full feedback loop is a first order loop with only one stock, i.e. threshold. This feedback leads the subject toward the optimal threshold, without any need to learn about the theories of how to find optimal thresholds in the signal-detection framework. We produce a set of random signals and noise, consistent with the signal detection condition (*noise ~ N(0,1)* and *signal ~ N(d',1)*), and choose randomly from them with a ratio that creates the desired base rate.[1]



(a) **(b)**

Fig. 4. Moving average of selection rate in last 50 trials (a) and threshold (b) for base rate of 0.5

Now, we can examine simulation results of this simple full feedback system. In Fig.4-a, we see how the model adjusts its selection rate to the base rate when the base rate is

[1] System dynamics suggests use of pink noise in modeling which is more similar to what happens in the real world (Sterman 2000). As one of our main goals in this paper is to test our model with data from a laboratory experiment, in which signals and noise are generated totally randomly, without any correlation among data points, we avoid pink noise generation, and use the simple normal random generator of Vensim. Also, this simplification makes the model easier to follow.

equal to 0.5. The figure illustrates average selection rate in last 50 trials, for 50<t<950. In this run d' is 1. Fig.4-b shows how the model is able to find the optimal threshold which in the base rate of 0.5 is equal to d'/2, i.e. 0.5. The experiment starts from an initial threshold of -2.

The speed of approaching depends on the time to change threshold (τ). Small changes in the selection rate graphs after t=450 relate to the randomness of experiments.[2] The detailed formulation of this model is illustrated in Appendix 1.

4. Conditional Feedback Model

As we discussed before, in the real world, whether you make a positive or negative decision can determine whether or not you receive (or at least perceive) a feedback. Back to our first examples, a police officer will know whether or not a suspect is a drug dealer, only if he decides to investigate him. A human resources manager will know about the true performance of an applicant, only if he hires him, and the applicant will know how good the job offer is after accepting and experiencing it. Otherwise, feedback is not clear, and in many cases, impossible to interpret. This concern leads us to activate a causal link from our decision rule to our perception about payoff (Fig. 5). In a simple word, increasing threshold will decrease positive decision rates (selection rate) and therefore, an experimenter will receive less feedback about payoff. The new introduced loop can have a major effect on the final results. Here, for simplification, we assume an immediate payoff perception, and keep the conditional loop a first degree non-liner loop. The rest of the paper will investigate the effect of such a link in threshold learning and the relevance of the ignorance of that link in formal studies.

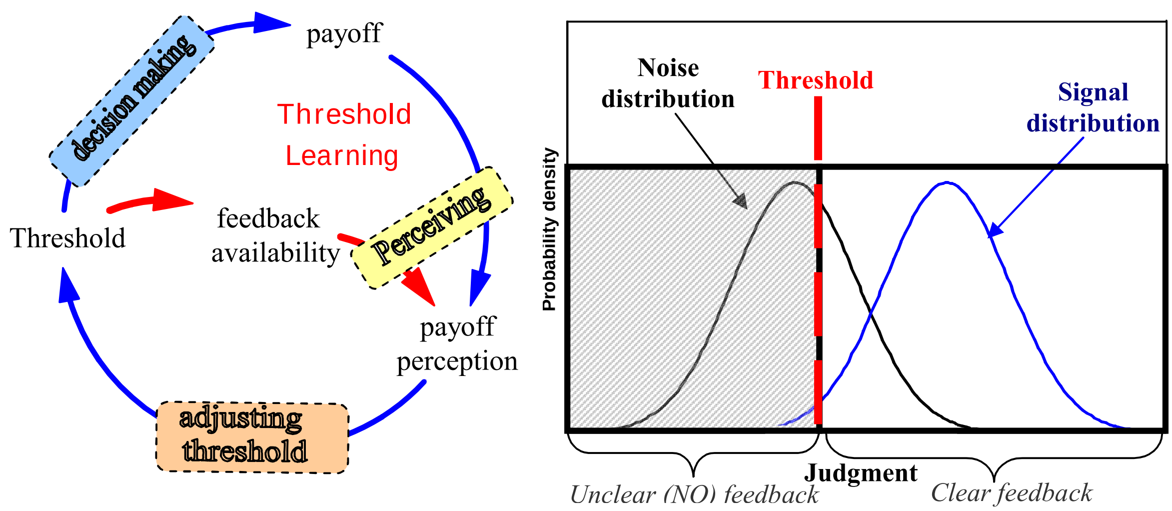


Fig. 5: Threshold learning in a conditional feedback situation: a) a dynamic model b) in signal detection framework

[2] As initial threshold is lower than the optimal threshold the selection rate's dynamics starts form 1, however, the model is not qualitatively sensitive to initial threshold. Further, our sensitivity analysis shows the model is not sensitive to random seeds, and it is able to find the optimal threshold.

4.1. Constructivist Coding

The important issue in modeling conditional feedback is about how people judge (code) the result of negative decisions. For the human resources manager, would he judge that all of his negative decisions about last year candidates were 100 percent correct? What about the police officer: will he believe that some portion of people who were not searched by him were actually drug dealers?

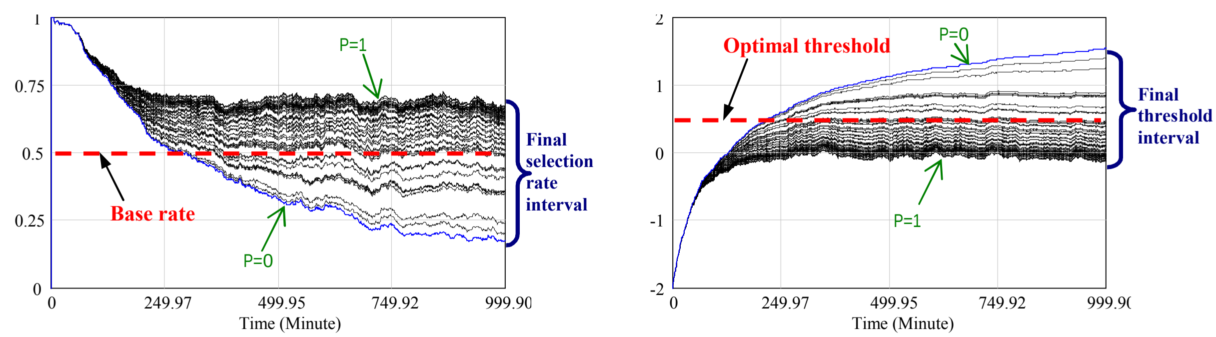
Constructivist coding is defined as a coding that represents what one believes is true (Elwin et al. 2007). We define *p*, proportion of coding absent feedback as signals, as a parameter to use for payoff estimation. So, when *p* is 0, the model assumes there is no wrong negative decision and when is equal to 1 the model assumes all of its negative decisions were wrong. Payoff estimation in conditional feedback can be calculated using eq.2. For positive decisions (*d=1*), we have

$$perceived\ payoff = payoff\ (Q,1) = (1-Q)*V_{fp} + Q*V_{tp} \qquad (eq.\ 6)$$

and for negative decisions (*d=0*):

$$perceived\ payoff = payoff\ (p,0) = (1-p)*V_{tn} + p*V_{fn}. \qquad (eq.7)$$

We simulate the model for the base rate of 0.5, d' of 1, and for a wide range of *p* ($0 \le p \le 1$). As we see in Fig. 6, the model is sensitive to the value of *p*, which means the way that people interpret their negative decisions can substantially influence their results. At two extremes, people who believe their negative decisions were always right or wrong end up with a considerable bias. This raises the importance of investigating how people really judge their negative decisions' performance.



(a) **(b)**

Fig. 6: Possible selection rates (a) and thresholds (b), for different strategies (different *P*s) of constructivist coding

There are three points about why there is a possibility for different coding of absent feedback. First, different people have different personalities; some are more conservative, presumably, coding more false for their negative decisions. Second, there are some state variables, like confidence that can change dynamically through the process of judgment, and create a different *p*. Third, a second loop learning process, if it exists, can lead to a more realistic perception of false negatives. If a person does not limit his learning to the feedback he receives from current false positives, but also, sometimes, questions the current threshold, and tests other areas to have some new experiences, he may be able to learn more about the hidden area under his negative decisions.[3] However, existence of second loop learning is an empirical question.

Most empirical studies of conditional feedback suggest there is a tendency to underestimate the optimal selection rate or, in another word, to overestimate the threshold (Elwin et. al 2007, Stewart et. al 2007). Elwin and his colleagues argue that, in conditional feedback situations, people tend to code negative decisions, the ones they don't receive feedback for, as totally correct ones. We call these individuals, confident constructivists. For this scenario, we have: $p=0$.

Fig. 7 shows simulation results for the base rate of 0.5 for a confident constructivist. This figure compares simulation results from full feedback condition with conditional feedback. Other parameters for conditional feedback are the same as full feedback condition (section 2). In Fig. 7-a, we see selection rate moves lower than the average base rate. Also, threshold moves higher than the optimal threshold, in Fig. 7-b.

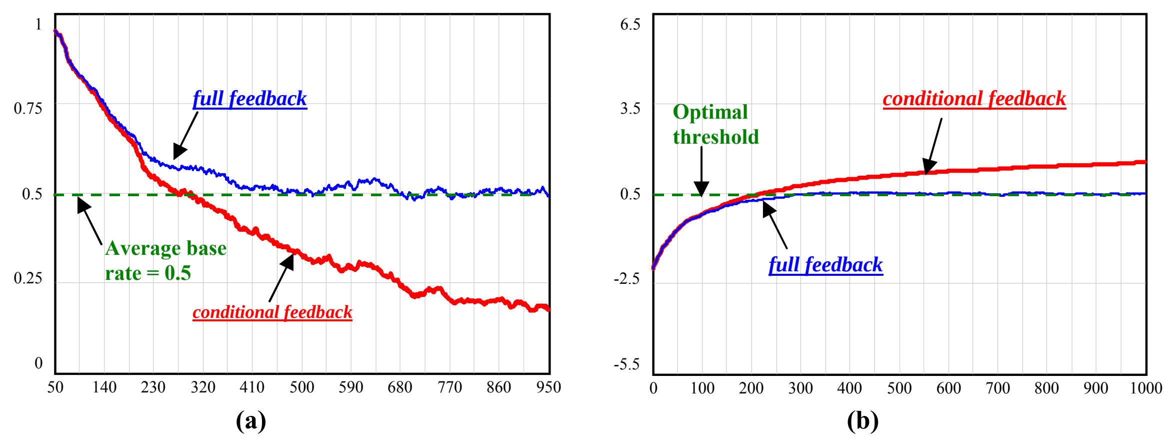


Fig. 7. Moving average of selection rate in last 50 trials (a) and threshold (b) for base rate of 0.5

But what do these simulation results really mean? Basically, in a full feedback situation, false positive decisions increase the threshold, and false negative ones decrease it. As in conditional feedback, the confident constructivist assumes all negative decisions are correct, there is only one adjustment force, and that is from false positive results. Therefore, forces are always toward increasing the threshold (decreasing selection rate), and it continues until no noise is perceived.

So far, we have shown how relaxing the assumption of full feedback, in the lack of second loop learning, can influence the final results. Particularly, considering suggestions of Elwin et. al, (2007) and Stewart et. al (2007) we see how people can underestimate the optimal selection rate. But the question is what is the real value for p, or how do people really code their negative decisions? Don't they learn any thing about the performance of their negative decisions? Later on, we use data from Elwin et al. (2007) and narrow the possible values of p to find more about people's behavior.

[3] In this paper, we do not attempt to model second loop learning, and leave it for further research; however, a non-zero p can be interpreted as a parameter to represent a person who may have tried to find more about the true performance of his negative decisions by some explorations.

5- Replications of an empirical investigation

Elwin et al. (2007) conduct an experiment including binary and continuous decision making situations. Sixty four subjects performed a computerized task of predicting economic outcomes for companies varying on four continuous cues (e.g. number of staffs) with values ranging from 0 to 10. Outcome was an additive function of the values of the four cues, with assignment of the cue weights of 4, 3, 2, and 1 to different concrete cue labels. The base rate of profitable companies was 0.5. In the binary set of experiments the subjects were supposed to select the companies for which they predict a positive profit.

The experiment had two major phases: First any subject had a series of training trials, and then entered the test phase. In the training part, a group of subjects performed 120 trials of full feedback decision making, while the other group performed 240 trials of conditional feedback. In the test phase, 60 judgments were made without feedback. They find that the subjects, who had the conditional feedback training, ended up with much lower selection rate in the test phase (0.33) in comparison with the other group (0.52). The authors propose a model of constructivist individuals that code true for all negative decisions (absent feedback) in the training phase and their model fits the data. Table-1 shows a summary of their experiment and results.

	Full feedback training	Conditional feedback training
Trials in training phase	120	240
Trials in test phase	60	60
Maximum d'	No maximum	No maximum
Number of subjects	32	32
Base rate	0.5	0.5
Selection rate	0.52	0.33
95% CI	0.44-0.60	0.26-0.41
Result of constructivist coding	0.48	0.34
95% CI	NA	Smaller than the interval of the selection rate.

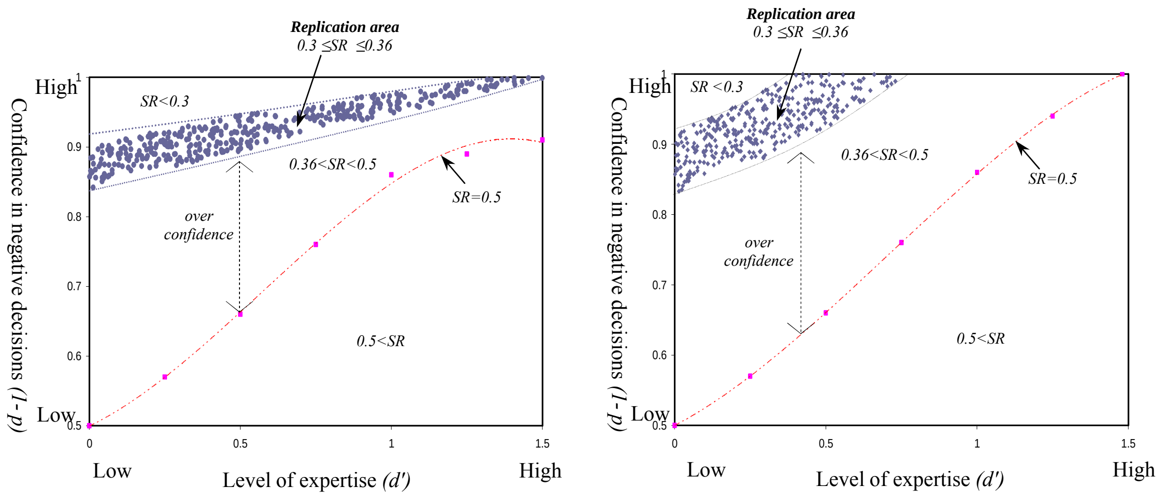
Table-1: Available data on Elwin et. al.'s work

Replicating the data by our model can be interesting for several reasons[4]. It can help us to learn more about the dynamics of Elwin et al.'s argument and check whether or not their results can be replicated. Further, we can test new possible explanations for the data, other than what is expressed by the original paper.

Two of the important parameters in the model are the level of expertise (*d'*) and the level of confidence in coding absent feedback (*1-p*). To investigate the effect, we conduct a sensitivity analysis for these parameters. Fig.10 shows the results. For each of the figures we have conducted 2000 simulation experiments to find the area that can replicate the reported data. Illustrated points in this figure represent the experiments that ended with the selection rate in the interval of [0.3, 0.36]. The first figure (8.a) is for τ=20 and

[4] Generally, calibration is the proper way of finding unknown parameters. But as the available data is limited to the final results of the test phase and does not include the dynamic behavior of subjects in the training phase, we believe calibration will suffer from an extensive number of possible solutions. This concern is consistent with one of the main concerns of Forrester (2007) in his talk at the 50th anniversary of system dynamics.

the second one (8.b) is for τ=100. In the figures, the areas that result in higher and lower selection rates are illustrated.



(a) τ=20 **(b) τ=100**

Fig.8 the quantities of d' and (1-p) that can replicate the data

Note: SR stands for selection rate. The blue area (replication area) is the area that replicates the data (each point in the replication area represents one successful experiment.) The line SR=0.5 shows the combination of (1-p) and d' that can result in no bias.

As we see, (1-p) is relatively high for the area of replication. This shows that people tend to underestimate false negatives in conditional feedback. Further, it shows even if a second loop learning exists, it is not effective enough as people are not able to find the correct *p* (shown by the line that represent selection rate (SR) equal to 0.5).

Furthermore, as we increase τ, the area moves upward resulting in a decline in bias. This comes from the fact that in a relatively higher τ, single noise detection will not cause a huge change in the threshold; therefore, the threshold stays in lower levels.

Considering the possibility of having different τ, we can sum up the results for τ>10 and offer Fig.9 as the possible set of d' and (1-p) that can replicate data. For each of those points there is a limited interval for τ that can replicate data. Three examples are shown in this figure as scenarios A-C. In scenario A, we are assuming an expert (d'=1.5) with a high level of confidence (p=0, and τ=30) as the decision maker in our model, and as we go toward scenarios B and C, the level of expertise and confidence decreases.

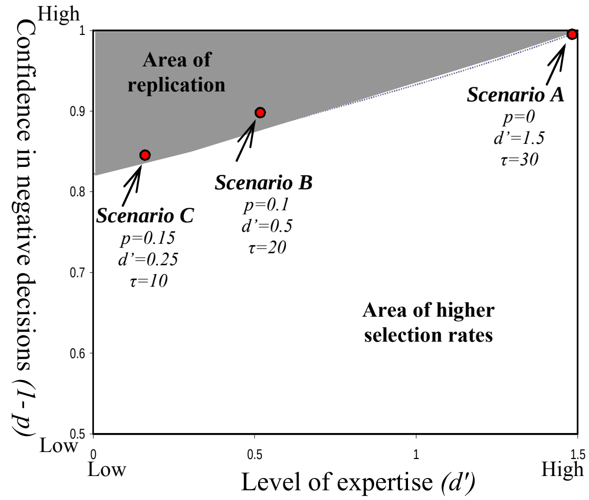


Fig. 9 The total area in which the model can replicate the results, and the three examples

Based on the represented figure, we can argue that the *x-axis* has an experience (or talent) component to it, as it is about the capability of interpreting data and judging. The *y-axis* has a personality component. So, we can say in a constant level of expertise, as confidence increases the selection rate falls. Also an increase in the level of expertise, which can be a result of learning about how to interpret cues, can result in an increase in selection rate. The interactive effect of these two parameters is very interesting for further studies.

As we see, our model is able to replicate Elwin et al.'s data for a considerable range of *d'*, *p*, and τ. However, in all of those, people underestimate *p* and are not able to learn the correct value of it.

6. Discussion and Conclusion

System dynamics as a way of analyzing nonlinear systems helped us to develop a simple model for full feedback and conditional feedback systems. The model was developed in a specific way to enable us to communicate with the decision science literature using the known framework of signal detection. Although the model was developed on individual level with disaggregated decision making processes for binary tasks, it still belongs to the family of decision making models in system dynamics. It creates insights based on activating a forgotten loop, and takes a stock flow approach in formulating variables.

The main contribution of our study is to give a new explanation for imperfectness of decision making in a series of tasks. While many scholars have intensified the negative effects of the complexity of tasks (Gonzalez 2005), misperception of delays (Sterman 1989a, Sterman 1989b, Rahmandad 2008), and feedback asymmetry (Denrell and March 2001) on learning, our work gives a different explanation for barriers to learning, that is conditionality of feedback. Our work does not reject other theories, but sheds more light from a new perspective on the problem of barriers to learning.

The simulation outcomes and the replication of data show that conditional feedback can result in bias and underestimation of the base rate. Basically, assuming people learn from their false decisions, in conditional feedback, all (or most) of negative decisions are treated as correct ones. Therefore, the dominant adjustment force comes from false positive results, not from false negative ones. Thus, forces are always toward increasing threshold (decreasing selection rate). Our experiments with different d' (level of expertise), and τ (time to adjust threshold) show that independent from these parameters, we will always face overconfidence, and bias in conditional feedback situations. This implies that in real world situations, conditionality of feedback for example for police officers, human resources management, university admission office, etc. can result in misperception of performance and overconfidence.

Our simple model of anchoring and adjustment behavior without any second loop learning fits the data from Elwin et al. Some may argue that in the existence of second loop learning, people may try new thresholds, correct their perception of false negative results, and find the optimal threshold. Although we do not have second loop learning in our model, our empirical investigation shows that people do not find the optimal threshold. The average *p* (perception about the ratio of true negatives to total negative decisions) is always overconfidently higher than the actual ratio. (Fig.10). This simply shows that even if, in the real world, second loop learning exists, it works for a limited number of people, and the average person is not able to find the optimal *p*. All of these results show that conditionality of feedback can be considered as a barrier to learning as it makes it very difficult for people to learn the optimal threshold.

Further, one of the most important implications of this study is its warning about overestimation of the relevance of full feedback assumptions in formal studies. As we see in our model, the results are very sensitive to how really people code their negative decisions' results. And, as data shows, average people underestimate their false negative results. This finding warns about the relevance of full feedback assumption in other studies.

There are some possible ways for extending this study. Discussing about how different *p* can be used to replicate the data, we find a wide range of possible *p* that can produce the data. This result comes from the fact that there is an interactive relation between the level of expertise (*d'*) and the optimal *p*. We may argue that, actually, none of *d'* or *p* are constant for an individual in the real world, but they may change dynamically through the process. Although this is more an empirical question, but intuitively we can accept that there can be some endogenous changes in these two variables. While experiencing, people learn about cue weights and it increases d'. Further, dynamics of confidence can lead to a change in *p*. Studying effects of these additional loops can be very interesting.

Also individuals can be different in how they interpret their negative decisions. This difference can be a personality trait issue. In further studies, individual level data can be gathered, and the model can be calibrated for each individual. Different parameters can then be compared. Testing a hypothesized relationship between some of the Big Five personality characteristics (like openness) and the way that people code negative decisions (*p*) is another possible and interesting way to extend this study.

Reference

Arkes, H. R., & B. A. Mellers, 2002, Do juries meet our expectations? *Law and Human Behavior, 26*(6), 625-639.

Bereby-Meyer Y. & A. E. Roth, 2006. The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation, *American Economic Review, 96*(4): 1029-1042

Denrell, J. and J. G. March, 2001. Adaptation as information restriction: The hot stove effect. *Organization Science 12*(5): 523-538.

Elwin, E., P. Juslin, H. Olsson, T. Enkvist, 2007, Constructivist Coding – Learning From Selective Feedback, *Psychological science 18* (2), 105-110

Epley, N. and T. Gilovich, 2001, Putting adjustment back in the anchoring and adjustment heuristic differential processing, *Psychological science 12* (5), 391-396

Erev I. 1998, Signal detection by Human Observers, *Psychological Review 105 (2)*: 280-298

Forrester J.W., 2007, System dynamics—the next fifty years, *System Dynamics Review 23* (2-3): 359 - 370

Gonzalez, C. , 2005, Decision support for real-time dynamic decision making tasks. *Organizational Behavior and Human Decision Processes, 96*, 142-154

Green, D. M., & J. Swets,1966, Signal detection theory and psychophysics. New York: Wiley.

Rahmandad, H. Effect of Delays on Complexity of Organizational Learning. Forthcoming, *Management Science.*

Rahmandad, H., N. P. Repenning and J. D. Sterman. 2007. Effect of Feedback Delays on Learning, Working paper. Available from http://www.mit.edu/~hazhir/papers/Learning-Rahmandad.pdf

Senge, P., 1991, The Fifth Discipline: The Art and Practice of The Learning Organization. New York: Currency Doubleday

Sterman, J., 1989a. Misperceptions of Feedback in Dynamic Decision Making. *Organizational Behavior and Human Decision Processes*. 43(3), 301-335.

Sterman, J., 1989b. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science*. 35(3), 321-339.

Sterman, J. D., 2000, Business Dynamics: Systems Thinking and Modeling for a Complex. World, McGraw-Hill: Irwin.

Swets, J. A., 1991, The science of high stakes decision making in an uncertain world (Transcript of a Science and Public Policy Seminar). Washington, D.C.: Federation of Behavioral, Psychological and cognitive sciences.

Swets, J.A., R.M. Dawes & J. Monahan, 2000, Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest 1*(1), 1-26

Stewart, T. R., 2000, Uncertainty, Judgment, and Error in Prediction. In D. Sarewitz & R. A. Pielke & R. Byerly (Eds.), Prediction: Science, Decision Making, and the Future of Nature (First ed., pp. 41-57). Washington, DC: Island Press.

Stewart, T., J. Mumpower & J. Holzworth, 2007, More on learning to make judgment and decisions in an uncertain world, *23rd Annual International Meeting of the Brunswik Society, Barcelona/Casablanca Rms, Westin Long Beach, Long Beach, CA*

Tversky, A. and D. Kahneman, 1974, Judgment under Uncertainty: Heuristics and Biases. *Science 185*, 1124-1131.

Appendix 1

Formulas (the vensim file is uploaded as a complementary document.)

I. The loops:

chng in OT=effect of gap in changing threshold*(X-optimal threshold)/time to change OT
effect of gap in changing threshold=f(gap/Normal gap)*normal effect
f([(0,0)-(10,2)],(0,0),(1,1),(1.9,1.8),(2.5,2),(10,2))
feedback availability=(1-switch to conditional feedback)+switch to conditional feedback*Positive decision
gap=perceived desired payoff-perceived payoff
Initial threshold=-2
normal effect=1
Normal gap=1
optimal threshold= INTEG (chng in OT, Initial threshold)
perceived desired payoff=Vtn+"perceived Q(X)"*(Vtp-Vtn)
perceived payoff= (1-"perceived Q(X)")*(1-Positive decision)*Vtn+"perceived Q(X)"*(1-Positive decision)*Vfn+(1-"perceived Q(X)")*Positive decision*Vfp+"perceived Q(X)"*Positive decision*Vtp
"perceived Q(X)"=feedback availability*"Q(X)"+(1-feedback availability)*signal coding ratio for CF
Positive decision=IF THEN ELSE(X>optimal threshold, 1 , 0)
"Q(X)"=IF THEN ELSE(RANDOM UNIFORM(1, 100 , NS1)>(100*(1-avrage base rate)), 1,0)
signal coding ratio for CF=0
switch to conditional feedback=0
time to change OT=50
Vfn=-1
Vfp=-1
Vtn=1
Vtp=1
X= IF THEN ELSE("Q(X)"=1, Xsignal , Xnoise)

II. The signal detection environment and additional functions

average selection rate= IF THEN ELSE(Time<T SR, total poistive decisions in last 50 decisions/(Time +1e-005) , total poistive decisions in last 50 decisions/T SR)
avrage base rate=0.5
bias in selection rate= average selection rate-avrage base rate
d prime=1
dynamic base rate=true/(true+false)
false= INTEG (fi-fo,(1-avrage base rate)*100)
fi=1-"Q(X)"
fo=false/T SR
in=Positive decision
NS1=100
NS2=1
NS3=10
out=IF THEN ELSE(Time>T SR, total poistive decisions in last 50 decisions/T SR,0)
"Q(X)"=IF THEN ELSE(RANDOM UNIFORM(1, 100 , NS1)>(100*(1-avrage base rate)), 1,0)
T SR=50
ti="Q(X)"
to=true/T SR
total poistive decisions in last 50 decisions= INTEG (in-out,0)
true= INTEG (ti-to,avrage base rate*100)
X=IF THEN ELSE("Q(X)"=1, Xsignal , Xnoise)
Xnoise=RANDOM NORMAL(-10, 10 , 0 , 1 , NS2)
Xsignal=RANDOM NORMAL(-10, 10 , d prime, 1 , NS3)