INFLUENCE OF DELAYED FEEDBACK ON LEARNING, PERFORMANCE AND STRATEGY SEARCH – REVISITED

PIOTR GOLEC[1]

Abstract. The purpose of this paper is to revisit the influence of delayed feedback on learning, performance and strategy search. Two main working hypotheses have been used as a guideline of the study. First hypothesis claimed that when learning environment is *"lenient"* (errors are not heavily punished, as a consequence of which generally positive feedback is observed), postponed feedback aids learning process (to the detriment of the performance - at least in the short run). Second hypothesis asserted that when learning environment is *"exacting"* (penalties imposed for errors are severe, therefore generally negative feedback is observed), postponed feedback adversely affects learning process and performance. An experiment has been designed and conducted in order to test above stated propositions. While the data from the experiment disconfirmed both theses, it showed an interesting phenomenon, which has not been conjectured before. Subjects in the lenient learning environment, while presented with delayed feedback, tended to re-represent the task more often, therefore changing the strategy they were applying more often than the subjects receiving immediate feedback. On reverse, delaying the feedback in the exacting environment, made the subjects to be more consistent in the application of their strategy.

INTRODUCTION

The literature dealing with the timing of feedback's importance presents a rather misty picture. One can find evidence that delayed feedback can be detrimental to the learner, can help the learner or does not affect the learner at all. But what does the delayed feedback really influence?

In a typical decision making task timing of feedback can most importantly affect: stock of knowledge acquired, performance, strategy search, time of decision making and subjectively perceived mood of the decision maker. As I demonstrate in more careful literature analysis later on in this paper, the impact of delayed feedback on learning and

[1] Agradezco la ayuda económica proporcionada por el Ministerio de Ciencia y Tecnología a través del proyecto SEC2002-03403

performance is still disputable. In terms of strategy search, it is generally accepted that the positive feedback reinforces maintaining and improving current strategy and negative feedback urges to search for new strategies (Schwartz [24], Hogarth et al. [14]), but little work has been done on investigating the impact of delayed feedback on the strategy search. In terms of time consumption of decision making, Schooler and Anderson [23]), for example, report that students learning LISP receiving immediate feedback moved through the training material in 40% less time than those receiving delayed feedback, yet without detrimental effects on learning. In the second experiment, during the same study, using improved LISP editor, subjects receiving immediate feedback were still 18% faster, but they were slower on the test problems and committed twice as many errors as the students receiving delayed feedback. Finally, in terms of the influence of timing of feedback on emotional states, let me just cite after Jensen[2]: "*studies reveal that for high risk or for highly stressed students, the most beneficial form of feedback is delayed feedback. It helps mediate impulsivity.*"

The purpose of this paper is to revisit the influence of the delayed feedback on learning, performance and strategy search. The experiment conducted by Hogarth et al. [14], investigating the influence of exactingness of environments and financial incentives on learning and performance, has been adapted in order to study the influence of delayed feedback on learning and performance under environments penalizing the errors of decision makers with different severity. Two main working hypotheses have been used as a guideline of the study, namely, that when learning environment is *"lenient"* (errors are not heavily punished, as a consequence of which generally positive feedback is observed), postponed feedback aids learning process (to the detriment of the performance - at least in the short run). When learning environment is *"exacting"* (penalties imposed for errors are severe, therefore generally negative feedback is observed), postponed feedback adversely affects learning process and performance. While the data from the experiment disconfirmed both theses, it showed an interesting phenomenon, which has not been conjectured before, namely that the timing of feedback can influence subjects' quest for strategy of problem solving, depending on the exactingness of the environment.

The rest of this paper is organized as follows. First, I discuss the importance of studying the timing of feedback and the controversy aroused around this issue. Second, I report the results of the experiment investigating the influence of delayed feedback on learning, performance and strategy search. Eventually I conclude the whole with a final discussion, possible criticism and some examples of applications of my findings.

1. DELAYED FEEDBACK – WHY STUDY IT?

> *". . . any theory that depicts learning as a process of mutual influence between learners and their environments must involve feedback implicitly or explicitly because, without feedback, mutual influence is by definition impossible"*
>
> Bangert-Drowns et al. [1, p. 214]

[2]The international Association of Supervision and Curriculum Development. Downloaded on June 09; 2003 from (http://www.cascd.org/spotjensen.shtml)

The research on feedback dates back to the early days of psychology (for excellent review of feedback studies check: Mory [21] or Brinko [3]) and from the very early on, feedback has been considered a main contributing factor to knowledge acquisition. One of the most important principles in all learning theory - law of reinforcement (or law of effect) - laid a scientific base for future feedback studies. In this rule for shaping behavior by the use of rewards (reinforcers), Skinner (Hilgard and Bower [12, p. 213] after Skinner) defines primary positive (negative) reinforcer as a stimulus which, when applied (removed) following an operant response, strengthens the probability of that response. He names food, water and sexual contact as examples of positive reinforcers and very bright light, loud noise or extreme heat or cold as examples of negative reinforcers. Hilgard (Hilgard and Bower [12, p. 562] after Hilgard) mentions also learned (secondary) reinforcers, like: money, praise, social approval, attention, dominance, etc.

It is generally agreed that a learnable response followed by a reinforcing event (stimulus, state of affairs, etc.) will receive an increment in its strength or probability of occurrence. But feedback as it is defined in this paper might differ from a reinforcer. As reinforcers used for conditioning are usually binary (yes-no, reward-punishment, etc.), feedback in a decision-making task can additionally convey an extended *"evaluative"* component, making use of the whole scale between extreme values. However, the virtue of feedback in decision-making converts in its main vice. The natural ambiguity of feedback continually confounds the inferential and the evaluative role of the information being fed back to the decision maker. Let us use an example from Hogarth et al. [14], of a student writing a final paper for his academic course. After receiving her grade, the student can easily get confounded with *"how much"* of the grade should she attribute to her paper writing skills and *"how much"* to the evaluation policy of her teacher.

It is then not questionable that in almost all human learning situations, the role of feedback is to convey information about the correctness of given response. This feedback, most probably, has its roots in long history of cultural training, when being *"correct"* is associated with parental praise and approval. Whatever our opinion about the process of learning is; whether we tend to sympathize with *"learning by doing"* approach or with social learning approach (a large amount of human learning is done vicariously, through observing another person making the skilled responses and then trying to imitate), the importance of feedback is undeniable.

Over time, different investigations have been performed and a lot has been written about the appropriate timing of feedback. But still dangerous myths co-exist with proven facts. One of them is that *learning efficiency is always increased through the use of immediate feedback*. For the purpose of my analysis I will divide learning situations (as it is common in learning literature, see Char [4]) into two broad groups: *"lower order learning"* (tasks requiring knowledge and its application) *"and higher order learning"* (tasks requiring analysis, synthesis and evaluation). For the former, Skinnerian approach - treating the feedback as reinforcement- seemed later to be confirmed by other researchers (for example: Bangert-Drowns et al. [1], Berquist and Philips [2]), who reached a general conclusion that in typical classroom settings the immediate feedback has been more effective. The same has been found in on-line learning (Ogilvie[3]). Also Hogarth [13]

[3] Ogilvie, https://webct.ait.iastate.edu/ISUtools/webhtml/designer/community/ogilvie.pdf

basing on Zajonc's assumption that it is more important to learn when something new is potentially harmful or dangerous than when it is not, presents a matrix of learning structure taking as dimensions the consequence of error and the quality of feedback. He claims that: *"with speedy and accurate feedback, we can learn both easily and accurately. However, when feedback is noisy, delayed or characterized by uncertainty, the process of learning by connections and reinforcement can break down (i.e., what we learn may not be valid)"*.

While praising the search for immediate and reliable feedback is obviously correct and valuable, I would like to contrast it with some other evidence. Morrison et al. [20] found that delayed and knowledge-of-correct-response feedback may be more beneficial than answer-until-correct or no-feedback for lower level learning, but that feedback effects become weaker when higher order understanding is evaluated. Similar results have been obtained by Clariana [5] (feedback effects weaker for higher order learning). Kulhavy [15] reports studies showing that delaying the feedback for a day or more results in significant increases in students' knowledge retention on post-test scores (*"Delay - Retention Effect"*). This phenomenon was observed mostly in multiple choices testing and its explanation has been attributed to so called *"interference perseveration"* hypothesis by Kulhavy and Anderson [16]. The authors found out that if feedback is delayed, the initial errors tend to be forgotten and do not interfere with learning the correct feedback responses.

For the *"higher order learning"* situations, superiority of the immediate feedback over the delayed feedback is also not unanimously agreed. Maddox et al. [18] while investigating the delayed feedback effects on rule-based and information-integration category learning state: *"learning in rule-based tasks is dominated by an explicit system that uses working memory and executive attention. This system appears to learn through a conscious process of hypothesis generation and testing. Given this hypothesis, manipulations of the timing and placement of corrective feedback should, and do, have little effect on rule-based learning. In fact, as suggested by Ashby et al.'s (1999) results, people can learn some rule-based categories with no feedback of any kind."* Gaynor [9] and Roper [22] has indicated that low achieving students may be able to benefit more from immediate feedback, whereas high achieving students tend to make a better use from delayed feedback. This fact was contributed to the ability of high performing students to analyze cumulated information, and having time - rethink the incorrect one. Mason and Bruning [19] - after thoroughly analyzing the literature dealing with feedback's timing, state that for the instruction aiming at developing higher order skills (such as comprehension, application, or abstract reasoning), the most effective kind of feedback seem to be the delayed or end-of-session one. Finally, Farquhar [7] conducting experiments on decision making in his micro-world - LOADER - finds out that *" . . . when corrective feedback was given, the delay of feedback for non-critical errors resulted in improved performance over the use of immediate feedback for all errors"*

In terms of influence of feedback's timing on strategy search, not much work has been done. Buchwald (Hilgard and Bower [12, p. 50] after Buchwald) proposes his delayed information experiment being a response to the Thorndikean [25] explanation of the *"spread of effect"*. Thorndike set up the following experiment: a list of nouns was presented to the subjects who must have guessed a digit from 0 to 10 associated with each noun. The experimenter kept saying arbitrary *"wrong"* to all of the attempts except (let say) the fifth item. Following the first cycle, a second - identical - cycle was conducted. In the illustration (Table 1) we can see a repetition of rewarded pair *"pencil - 3"*. Interestingly, we also see an exact repetition of responses given in the first trial on the 3rd and 4th position (before the rewarded pair) and on the 6th and 7th position (after the rewarded pair). This phenomenon was called the *"spread of effect"*.

Table 1. Illustrative events on two trials of a Thorndikean and Buchwald's associative learning task

Thorndike:		**Trial 1**		**Trial2**	**Buchwald: Trial 2**
Serial Position	*Cue Word*	*S's Guess*	*E's Feedback*	*S's Guess*	*S's Guess*
1.	cup	7	wrong	2	2
2.	scissors	9	wrong	1	1
3.	plate	5	wrong	5	8
4.	sky	6	wrong	6	2
5.	pencil	3	RIGHT	3	3
6.	book	1	wrong	1	9
7.	house	8	wrong	8	4
8.	chair	2	wrong	6	6

In contrast, Buchwald run the same kind of experiment, with only difference of withholding subjects' feedback during the first trial and revealing it shortly before the second trial decisions were to be taken. In this case only the rewarded pair has been repeated. Thorndikean prediction that subjects receiving immediate feedback should show larger tendency to vary their responses happened to be false since it was Buchwald's subjects with delayed feedback who showed more variability in answers. Buchwald's explanation of the *"spread of effect"* phenomenon is the following: While Thorndikean subjects had to remember all combinations of Stimuli - Response - Outcome from the first trial (Stimuli - Response, Response - Outcome and Stimuli - Outcome), Buchwald's subjects had to remember only the Stimuli - Response association.

Recently, the emergence of information-processing theory opened new opportunities for feedback's timing studies, but mostly in the direction of the kind of feedback (the most simple differentiation: corrective vs. non-corrective, see for example in Kulhavy and Stock [17]) and timing of the feedback as a function of error's criticality (see for example in Farquhar [7]). Taking into account this feedback's delivery options increase, some researchers (Dempsey et al. [6]) call for reviewing the theory explaining when delayed feedback can help learners. The experiment reported below is supposed to shed more light on the discussed issues.

2. EXPERIMENT

2.1. Theoretical background. The model used by Hogarth et al. [14] has been adapted to the need of this study. Therefore, theoretically, the experiment consisted of two main manipulations. First of all - as in Hogarth et al. [14] - the exactingness of learning environment has been varied. Exactingness can have both positive and negative influence on learning, performance and strategy search. If the learning environment is very lenient - decision maker receiving positive feedback regardless of the action he has just taken (as a consequence of not being penalized for errors committed) - learning is practically impossible and the search for alternative strategies minimized. On the other hand, when the learning environment is exacting - decision maker receiving negative feedback most of the times (stemming from each error being severely penalized) - learning is boosted through alternative strategy search, at the same time affecting the performance negatively (at least in short term).
Second of all, the variable used by Hogarth et al. [14] – "incentives" – has been replaced with feedback's timing. Exactly half of the subjects received their feedback immediately after making each decision; the other half saw their feedback with a delay.

The task used in the experiment resembled many situations from every day life, when subjects must learn from outcome feedback. Over a series of periods, subjects were given a signaling variable (W), on the basis of which, they were supposed to predict a criterion, which was probabilistically connected to the signal. After their decision, subjects have been informed about their performance, however not via observing the correct value of the decision variable, but through observing another variable - Performance Evaluation Measure (PEM). The scale and the upper and lower bound of PEM were not known to the subjects; there were only informed that the bigger it is the better they have performed.

The experiment has been based on the following predictions:

> **Hypothesis 1**: when learning environment is *"lenient"* (errors are not heavily punished, as a consequence of which generally positive feedback is observed), postponed feedback aids learning process (to the detriment of the performance - at least in the short run).

The first hypothesis is derived from the fact that generally positive feedback reinforces maintaining of existing strategy. If so, there is a risk that no learning process occurs (one might get stuck in a local optimum). Postponing of the feedback in these circumstances could force the subject to render higher intellectual effort, use the available information to a higher extent and - gradually as the task develops - increase the ability of thinking about her answers and self-correct misconceptions prior to receiving feedback.

> **Hypothesis 2**: When learning environment is *"exacting"* (penalties imposed for errors are severe, therefore generally negative feedback is observed), postponed feedback adversely affects learning process and performance

The second hypothesis is based on the idea that both - very exacting environment and delayed feedback - would turn the learning environment into a very wicked one. Since given generally negative feedback subjects tend to seek for alternative strategies already, the delay in feedback information can only confound the inferential and evaluative role of feedback, therefore decrementing the learning process and performance.

Additionally, the following working hypotheses have been drawn:

(1) As in Hogarth et al. [14], it was conjectured that the performance is a single-peaked (inverted-U-shaped) function of exactingness.

(2) Basing on Kulhavy and Anderson [16] findings that subjects given delayed feedback spend more time studying the feedback, the same was supposed.

(3) The subjects were expected to declare worse mood with increasing levels of exactingness. At the same time, subjects were supposed to declare (on average) worse mood for delayed feedback than for immediate feedback condition.

2.2. Detailed design. ***Subjects.*** The subjects were all graduate students from the Master and Doctoral program in Economics, Finance and Business at the Univeristat Pompeu Fabra[4]. All together 36 subjects were recruited via an email message, informing all students of the department of Economics and Business about a possibility of participation in a remunerated experiment on decision-making[5]. Mean age of the subjects was 26,7 years. There were 27 males and 9 females in the sample. Subjects were remunerated with 5 euros for their participation, but not for performance.

Procedure. At the beginning subjects were instructed that the experiment consisted of 2 rounds and that in each round they had to make 25 decisions. They were asked to complete a questionnaire after each round. They were informed that there was no time constraint to finish the task but that time was going to be measured. Subjects were allowed to take notes (even though the software kept track of their decisions in each round and the subjects could browse their previous answers without any restrictions). All subjects were asked to read carefully the following instruction:

> *"It is summer . . .*
> *You are an ice cream producer - the only one in your small town in Iowa. Every day you have to make a decision about how much ice cream to produce. The capacity of your production line restricts you to choose a value of production between* 1 *and* 600 *liters a day. To help you, the only information you can observe freely is the weather forecast for the next day, especially the foreseen temperature (Fahrenheit scale).*
> *Your friend is a student of the Iowa University and for her tesina she is conducting a study of how well producers are able to predict the future demand basing on the weather forecast. In order to have a comparison between all national producers, she developed a measure for evaluating performance based on the discrepancy between the actual demand (she calculates it calling all potential clients) and the decisions taken by the producers. You are not exactly sure how this measure is calculated, but your intuition urges you to focus on maximizing it since it captures information*

[4] With exception of two students who were also post-graduate students from other universities.

[5] All together 39 students had been recruited. However, two were used as a pilot (to better calibrate models' parameters) and one had to be excluded since he clearly misunderstood the task.

that otherwise would never be available to you (i.e. the actual demand for each day). Your friend - the Iowa University student - additionally offered to help you to understand the system of the evaluation measure, through supplying you with two auxiliary variables A and B. Since her research is financed via a governmental grant, she has to maintain her neutrality and cannot favor you in any other way. She insisted however that the variables A and B convey information that will help you to understand the system.

Subjects in the *"delayed feedback"* condition additionally received the following explication:

Unfortunately, calculating the total demand of the market takes your friend a while, so you are only able to see your evaluation (and variables A and B) with a 48 *hours delay, i.e. the evaluation from the day 1 is going to be available to you after taking the decision on the production level in day* 3.

Task. The task, individually administered by microcomputer, involved a series of decisions (25 in each of the two rounds). The experiment was programmed and conducted with the software z-Tree (Fischbacher [8])[6]. Subjects were randomly assigned to one of the six groups created by crossing two variables, one with three levels (exacting vs. intermediate vs. lenient environment) and the other one with two levels (immediate vs. delayed feedback).

At the beginning of each period a signal W (weather forecast - temperature in Celsius degrees) was randomly chosen each time from a normal distribution with mean 25 and standard deviation 7. The temperature was than translated into Fahrenheit degrees and presented to the subject[7]. Immediately after observing the signal subjects had to take a production decision (values restricted to the natural interval <1,600>). Subjects following the *"immediate feedback"* condition received their PEM, and auxiliary variables immediately. Subjects following the *"delayed feedback"* condition received their performance evaluation measure with a delay of two periods (n+2) [8].

Based on *"Evaluation points"* from Hogarth et al. [14], the Performance Evaluation Measure (PEM) was calculated using the formula:

[6] Due to a specific (individual) application of the software, subjects were also informed that during the time they fill in the first questionnaire, the experimenter would manually restart the "game" for them.

[7] This manipulation has been done in order to make the task somewhat more diffcult. The conjecture was that most of the subjects were not going to know the Fahrenheit scale, so immediate "optimal scaling" of the task would be impossible. In practice only 10 out of 36 subjects knew the relationship between the Fahrenheit and Celsius scale. More strict statistical analysis was not possible, but the author had not found signs of subjects knowing the difference between the two scales, consistently outperforming the subjects who did not know it.

[8] In the first two periods subjects were informed: "No performance evaluation available yet". Starting from period 3 the following message was displayed: "Your performance (two periods ago) received a score of: XXX. The variable A had a value of: YYY. The variable B had a value of: ZZZ". In the last period of each round, subjects were given feedback from 3 periods at once ("two periods ago", "one period ago" and "current period").

$$(1)\ PEM = 600 - \alpha*(Q - D)^2$$

Where: α is the exactingness parameter (respectively 0:02, 0:4, 2 for lenient, intermediate and exacting environment)

Q is subject's response, and

D is correct value of the criterion

As can be seen, PEM function is a negative function of squared discrepancy between the subject's answer and the correct answer. The parameter α was used to manipulate the exactingness of the learning environment (increasing α increases the penalty for erroneous response).

D had a strong correlation with the weather signal W and was described as a linear function:

$$(2)\ D = 14*W - 60 + \textit{error term}$$

Where: *error term* is a small noise with mean of zero.

To compare performance (P) across different levels of exactingness the observed performance in a given trial *i* was measured by:

$$(3)\ P_i = 600 - |Q_i - D_i|,$$

with mean performance, π, (over n trials) calculated using the following formula:

$$(4)\ \pi = (1/n)\ \Sigma\ P_i$$

Since π is a mean performance of just one subject, Π - mean performance in a group - was calculated (over m subjects in each group):

$$(5)\ \Pi=(1/m)\ \Sigma\ \pi$$

Additionally, auxiliary variables A and B were calculated in the following way:

$$(6)\ \text{If } Q < D, \text{ then } A = Q \text{ and } B = 0$$
$$\text{If } Q >= D, \text{ then } A = D \text{ and } B = Q - D$$

Variables A and B have been introduced in order to provide useful piece of information for subjects with high understanding of the *"system"*.

Post-experimental procedure. After each round of the experiment subjects were asked to complete a short questionnaire. This questionnaire was designed in order to check the *"designer knowledge"* that subjects acquired during the experiment

Questions: (after first round[9])
(1) Age, Sex,
(2) Please indicate on the scale below (Likert scale from 1-7 where 1 is totally relaxed and 7 is very stressed) how you felt during this round of the experiment?

(after second round)
(1) Please indicate on the scale below (Likert scale from 1-7, where 1 is totally relaxed and 7 is very stressed) how you felt during this round of the experiment?
(2) Try to explain - in general terms - *"how does the game work?"* Focus specially on the roles played by different variables (*"Temperature"*, *"Actual demand"* and *"Production level"*) and their interrelations.
(3) What have you learned about auxiliary variables A and B?
(4) Can you think of any specific *"tips"* you would give to another person in order to help her to achieve high performance evaluation measure?
(5) Before the experiment, did you know the relation between Celsius and Fahrenheit degrees? (Yes, No)

3. RESULTS

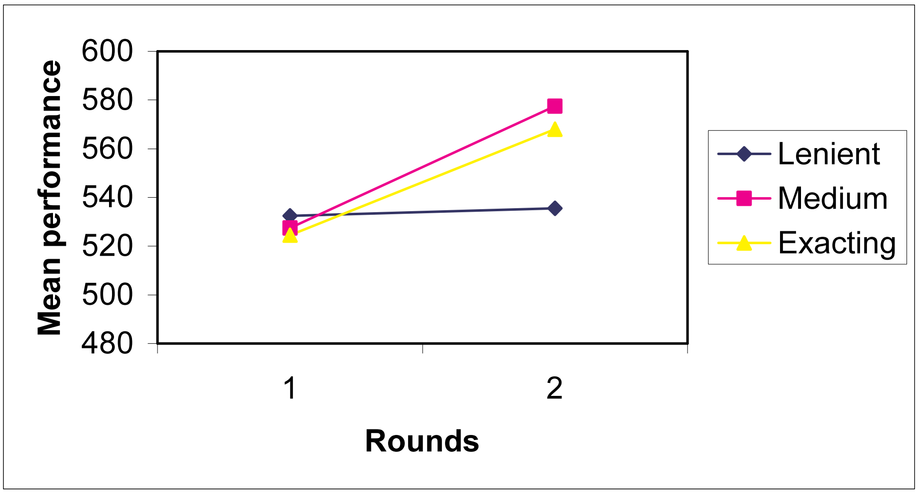
As indicated above, the manipulation of the exactingness of learning environments has been crucial to the study. As a check whether the subjects observed negative feedback to different extent, I report the percent of negative PEM (on average[10]) for each environment. That is - in lenient environment subjects observed negative PEM 12% of the times, in medium exacting environment - 34% and in exacting environment - 60%. Table 2 and Figures 1 and 2 provide an overview of the results. For all the experimental conditions, in the upper part, Table 2 reports means and standard deviations regarding performance (Π) for each round. Figure 1 shows the mean performance (Π) of subjects in each environment per round. As it can be seen, in the first round, subjects in lenient environment were on average slightly better than subjects from other groups. In the second round however, subjects from medium and exacting learning environments had outperformed "lenient" subjects. Overall, as predicted, subjects in the medium learning environment performed best.
Figure 2 shows mean performance (Π) additionally averaged over rounds. We can see that under immediate feedback, performance decreases with increasing environment's exactingness. The reverse can be observed for delayed feedback.

[9] Questionnaires (on two sheets of paper) stapled together.
[10] The percentages are given for each environment as an average between two groups of different feedback conditions. A big disproportion must be noted in the lenient environment, where subjects in the immediate feedback condition observed negative feedback only 1% of the times, but in the delayed feedback group - 23% of the times.

Table 2. Performance (Π) and inconsistency in lenient, medium and exacting environment.

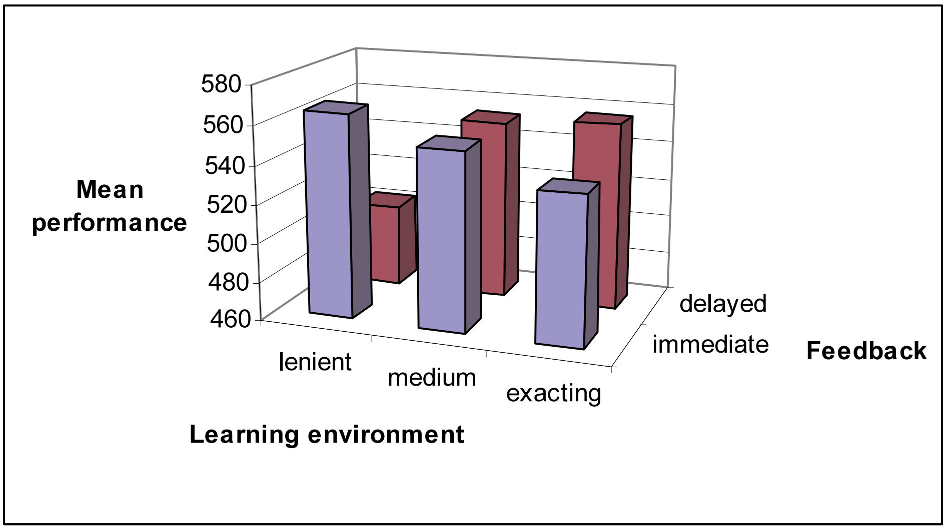
	Immediate feedback			Delayed feedback		
Round	Lenient	Medium	Exacting	Lenient	Medium	Exacting
Performance (Π)						
1						
M	559	529	511	506	526	538
SD	18	43	27	62	41	18
2						
M	572	575	561	499	580	575
SD	8	17	30	72	10	20
Average:	565.5	552	536	502.5	553	556.5
Inconsistency (1-R-square)						
1	27%	35%	62%	32%	49%	40%
2	10%	12%	32%	45%	11%	13%
Average:	18%	23%	47%	38%	30%	27%

Figure 1. Mean performance (Π) by types of environments and round



In terms of more structured statistical analysis, initial ANOVA reported that the only significant effect is the interaction between variables *"Feedback"* and *"Learning environment"*. However, the tests for homogeneity of variance for means showed statistically significant differences between the variances (both Bartlett's and Levene's tests significant at $\alpha = 0{:}01$). Therefore this result might be questionable. Since the homogeneity of variance test showed that variances differ predominantly in the lenient environment (between immediate and delayed feedback conditions and not much in other environments, I perform ANOVA with rounds as a repeated factor only for medium and exacting learning environment. For medium environment, ANOVA shows a significant difference between means of rounds: $F(1, 20) = 15.2$, $MSe = 987$, $p < 0.01$. For exacting environment, the effect for rounds is significant: $F(1, 20) = 19.1$, $MSe = 596$, $p < 0.01$. For differences between the timing of feedback, the null hypothesis of means being from the same distribution is just at the rejection region: $F(1, 20) = 4.3$, $MSe = 596$, $p = 0.05$.

Figure 2. Mean performance (Π) for rounds 1 and 2 by experimental conditions.



Since standard deviations reported in Table 2 are just indicators of groups' homogeneity, we need another measure of subjects' quest for strategy. Let us imagine a model, where all answers given by each subjects are regressed on the signal they have observed, that is:

$$(7)\ Q = \beta_0 + \beta_1 W + z$$

R-squared obtained through this regression can tell us how much of the variability of the responses can be explained by the model, i.e. how well the task is understood and how consistent the subjects are with executing their current strategies. Therefore 1-R-squared gives us a measure of inconsistency, i.e. how much subjects deviate from earlier chosen strategy[11]. I report this measure (averaged over all subjects in each experimental group) as *"Inconsistency"* in Table 2.

Under immediate feedback, as predicted, we can see that inconsistency increases with increasing exactingness. This is however not the case for delayed feedback. The difference in terms of inconsistency between delayed and immediate feedback in the medium exacting environment is not big (except of the first round – but this is intuitively

[11] R-squared from regression models has been widely used in the judgment research and lens model, see for example Stewart [25]

easy to explain). Subjects in the delayed feedback group had relatively less information in the first round, so they *"found out"* their strategy *"later"* than subjects in the immediate feedback group. By the second round, *"delayed"* subjects already outperformed the *"immediate"* group. In terms of more strict statistical analysis, two-way ANOVA with rounds as repeated factor showed that the only significant effect is the one for rounds, $F(1, 20) = 12.01$, *Mse* = 1, $p < 0.01$.

The interesting case is the one of lenient and exacting environment. In the exacting environment, subjects receiving delayed feedback from the beginning showed less inconsistency and outperformed subjects receiving immediate feedback significantly (ANOVA revealed that both effects are statistically significant, for rounds: $F(1, 20) = 8.43$, *Mse* = 1.25, $p < 0.01$ and for columns $F(1, 20) = 4.63$, *Mse* = 1.25, $p = 0.04$). In the lenient environment, subjects receiving delayed feedback, from the beginning started performing badly and by the second round, not only did not improve, but also were increasingly inconsistent in their responses (ANOVA showed significance of the feedback's timing change – $F(1, 20) = 5.99$, *Mse* = 0.86, $p = 0.02$, but not for repeated rounds).

3.1. Additional results. ***Time.*** The time of each decision has been measured. The averaged time for experimental groups per round are reported in Table 3. It can be seen that on average, subjects receiving delayed feedback took more time in order to make decisions (ANOVA with rounds as a repeated factor revealed significance for exacting environment ($F(1, 20) = 5.2$, *MSe* = 145, $p = 0.03$) and for medium learning environment ($F(1, 20) = 5.3$, *MSe* = 236, $p = 0.03$), but not for lenient environment).

Table 3. Time of decision making by round in all experimental conditions.

	Immediate feedback			Delayed feedback		
Round	Lenient	Medium	Exacting	Lenient	Medium	Exacting
1						
Time	29	29	25	32	51	39
SD	16	18	15	24	17	15
2						
Time	20	27	18	18	35	26
SD	10	11	8	10	14	7
Average	24.5	28	21.5	25	43	32.5

Stress. One of the questions asked in the questionnaires (after each round) was: *"Please indicate on the scale below (Likert scale from* 1-7*, where* 1 *is totally relaxed and* 7 *is very stressed) how you felt during this round of the experiment?"* Table 4 reports mean answers per group and the evolution over the rounds.

Table 4. Auto-evaluation of stress over rounds in all experimental conditions.

	Immediate feedback			Delayed feedback		
Round	Lenient	Medium	Exacting	Lenient	Medium	Exacting
1	2.3	2.5	2.7	2.7	4.2	2.7
2	2.8	2.8	2.5	2.7	3.3	2.5

As can be seen, there are no big differences between the groups receiving immediate and delayed feedback, except of medium exacting learning environment.

Designer's knowledge. The rest of the questions in the questionnaire was supposed to gauge subjects' *"designer knowledge"* of the task after having taken 50 decisions. The questions had been evaluated and a point had been given for each valuable piece of information reported by a subject. For question 2 – general knowledge - points might have been achieved by stating the sign and the slope of the relation between W and decision variable, being aware of two types of errors (over-estimating or under-estimating the production level), expressing the upper limit of the performance measure, etc. For question 3 - knowledge about variables A and B - points were given for stating that *"A + B = decision variable"*, indicating that variables A and B depend on the two types of error, etc. Finally, for question 4 - specific tips - points were given for any valid advice, like for example, that it is always better to overestimate the guess, and so on. The mean scores obtained by the subjects in each experimental group are reported in Table 5.

Table 5. “Designer’s knowledge” scores in all experimental conditions.

	Learning environment		
	Lenient	Medium	Exacting
Immediate feedback	3.7	3.2	3.7
Delayed feedback	2.5	4.0	3.2

Here, again data shows similar level of understanding of the task, with an exception of a clear indication of the difference between poorly performing *"delayed feedback - lenient environment"* group and best performing *"delayed feedback - medium environment"* group.

4. DISCUSSION

The two main hypotheses have not been confirmed by the experimental data. Delaying feedback in the lenient environment did not help the decision makers, whereas delaying feedback under exacting environment did. In terms of additional working hypotheses, performance was a single-peaked (inverted-U-shaped) function of exactingness. Subjects given delayed feedback spent more time on the decision-making, but they did not report significantly more stress than subjects observing immediate feedback. The mood of subjects was also robust to increases of the exactingness of learning environments.

There are three different ways of arguing if learning process occurred. The first one is just via observing the performance, indicated in Figure 1 by clearly increasing curves. Under this approach, delaying feedback in lenient environment instead of boosting learning process, impacted it negatively. The second way is to take into consideration the data collected in questionnaires. The evaluation of this data is however subjective and it does not help to reach any consistent standpoint. The third way is to search for signs of *"designer's knowledge"* in subjects' responses. Once a subject understood that whenever his production is overestimated, variable A gives him the correct answer for current temperature, and given that the temperatures tend to repeat, she should use the correct answer every time presented with the same signal. All 50 answers of each one of 36 subjects had been analyzed searching for this pattern. Unfortunately this method presented no discriminative power, since an average of two subjects per group following this rule have been found.

In terms of strategy search interesting phenomena occurred. First of all, subjects in the lenient learning environment, while presented with delayed feedback, tended to re-represent the task more often than the subjects receiving immediate feedback, therefore changing the strategy they were applying more often. On reverse, delaying the feedback in the exacting environment, made the subjects to be more consistent in the application of their strategy. I offer two ways of explaining these findings. One possible way of reasoning is that in the lenient environment, subjects searched for improving the quality of delayed feedback through excess variability. Subjects, after having taken first decisions without any knowledge about the task, had a lot of chances to receive a positive or slightly negative feedback. From the moment a representation of the task has been made, the delay of PEM required the subjects to test their strategies in a more extreme way, in a sense overreacting to the signal, in order to check their validity. Since the feedback per se was of a pretty bad quality, i.e. to a great extent every action received a positive feedback, delaying its visualization forced subjects to increase the variability of answers, therefore improving the inferential component of the information being fed back to them.

In the exacting environment, since each small deviation from the correct answer resulted in a very penalized PEM, this mechanism was not needed. Even delayed, the feedback had enough quality to guide the subjects towards a correct strategy.

Another, related, explication might take into consideration aspiration levels. Subjects were motivated to focus on maximizing their performance, but the ones in the lenient environment, through observing mainly positive PEM might have conformed with it more easily than subjects from the exacting environment[12]. The latter subjects in turn,

[12] This way of reasoning is based mainly on the interview with one of the subjects in the "lenient environment - delayed feedback" group, who was copying the signal W into the production decision consistently over a considerable number of periods. When asked whether she understood the task, the subject answered "yes" and that this heuristic strategy "seemed to work for some time".

from the beginning were much more inclined to see big negative numbers, therefore once they observed a positive PEM, they were much more motivated to focus on maintaining the prospective PEMs positive. This way of thinking is somewhat related to the theory of Higgins ([10], [11]) stating that: *"The outcome of loss and failure is congruent with the strategy of avoiding loss in prevention focus, whereas the outcome of gain and success is congruent with the strategy of approaching rewards"*. If we accept the fact that delayed feedback in the exacting environment increased the negative framing of the task, decreased variability in responses might be explained through the further loss prevention approach. If on the other hand we accept that delayed feedback in the lenient environment increased the feeling of success (through increasing the saliency of positive PEMs), increased variability of responses could be explained.

Some comments are in order here. First of all, throughout the study it is assumed that all the subjects are well motivated and their choices are in line with some long-run desire to perform well. As I have no reasons to think the contrary, marginal number of subjects reported that they have committed data introduction errors, f. ex. introducing 10 instead of 100. Second of all, it is worth stressing, that the kind of task analyzed in the experiment is different to the standard tasks used to analyze the influence of delayed feedback. The *"ice-cream"* experiment is not a dynamic decision making process in the usual sense (there is no *"carryover"* effect from one period to the next), however subjects are supposed to make decisions without knowing the outcome feedback from the previous periods (even delayed). Therefore standard terminology of *"delayed"* or *"postponed"* feedback should be used with care (I propose to merge those terms in one: *"ultra - delayed"* feedback - that means that the feedback is delayed with respect to the whole task, but postponed with respect to the given trial).

4.1. Possible criticism. There is a number of possible criticisms to the reported study. First of all, the sample size of 36 subjects makes all the findings statistically questionable. Second of all, the linear relationship between the signal W and the correct criterion D should be revisited. To a great extent, subjects - even when searching - are searching in vain, since there is no local optimum to escape from. A non-monotonic function with local and global optimum would probably increase the explanatory power of the study. Finally, a clear *"ceiling effect"* can be found, restricting options for further learning. In this sense, subjects are at risk of suffering from the *"video-game effect"*, derived from an observation that children can often obtain very high scores playing video-games, but asked how did they do it, or what one has to do in order to play well - have difficulties to answer.

4.2. Practical implementation. As expected, it has been shown that under immediate feedback, increasing the exactingness of learning environments induces strategy search. But the current study goes one step further. Another variable inducing strategy search has been identified - namely the feedback's timing. If it is true that delaying the feedback in the exacting environment could maintain subjects *"additionally"* focused on maximizing their performance at smallest possible deviations from currently used strategy, and that delaying the feedback in the lenient environment made subjects focus especially on winning strategy search, those facts can have a tremendous practical implementations.

Let us imagine a following illustrative setup: a Financial Advisor (FA) is helping several clients to manage their long-term investments in stocks. If the FA - as an expert – believes that the financial markets are passing through a temporal downturn, therefore whatever action a client takes (even a correct one) will bring penalizing effects (exacting environment), it should be justified that she (the FA) delays feedback about clients portfolio's and stock market's performance. In this manner clients bear smaller risk of *"over-reacting"*, i.e. selling good shares and buying another, which seem to be temporarily less affected by the overall slump, but are priced above their fundamental value.

On the other hand let us think of a *“creative”* advertising campaign being developed for product “Z”. Since the product is new to the market and the prior market studies showed that it would sell well, we might say that there exists a big probability that the penalty in terms of sales for not advertising the product in the *“optimal”* manner is going to be small (lenient environment). Therefore, it is in the interest of the producer of product “Z”, to delay the feedback in terms of knowledge of sales rate to the marketing agency, so to avoid being “stuck” in the sub-optimal advertising campaign.

5. END WORD

The study that has been described in this paper makes two important contributions. First of all, it adds to already existing evidence about the important role of delayed feedback. It demonstrates that some of the previous beliefs about feedback being useful only immediately after the decision should be revisited. Secondly, this paper opens a new, important area for research. Feedback's timing has so far not been taken into consideration as a variable inducing or preserving from the strategy change. The evidence collected in this study suggests that this matter should be investigated further.

References

1. R. L. Bangert-Drowns, C. C. Kulik, J. A. Kulik, and M. Morgan, *The instructional effect of feedback in test-like events*, Review of Educational Research 61(2) (1991), 213-238.
2. W. H. Berquist and S. R. Philips, *A handbook for faculty development*, vol. 1, Dansville, NY.: Council for the advancement of small colleges, 1975.
3. K. T. Brinko, *The practice of giving feedback to improve teaching: What is effective?*, The Journal of Higher Education 64 (1993), no. 5, 574-593.
4. R. O. Char, *The effect of delay of informative feedback on the retention of verbal information and higher-order learning, for college students*, Dissertation abstracts international 40,748a, Florida State University, 1978.
5. R. B. Clariana (ed.), *The effects of different feedback strategies using computer-administered multiple-choice questions as instruction*, ERIC Document Reproduction Service No. ED 347 983, Proceedings of Selected Research and Development Presentations at the Convention of the Association for Educational Communications and Technology, 1992.
6. J. V. Dempsey, M. P. Driscoll, and L. K. Swindell, *Text-based feedback*, ch. Interactive instruction and feedback, pp. 21-54, Englewood Cliffs, NJ: Educational Technology., 1993.
7. J. D. Farquhar, *A summary of research with the console-operations tutor: Loader*, [Online] Available http://www.gsu.edu/ wwwitr/research/farquhar1995/ (1995).
8. U. Fischbacher, *z-tree - zurich toolbox for readymade economic experiments - experimenter's manual*, Institute for Empirical Research in Economics, University of Zurich, working paper nr. 21 ed., 1999.
9. P. Gaynor, *The effect of feedback delay on retention of computer-based mathematical material*, Journal of Computer-Based Instruction 8(2) (1981), 28-34.
10. E. T. Higgins, *Beyond pleasure and pain*, American Psychologist (1997), no. 52, 1280-1300.
11. E. T. Higgins, *Advances in experimental social psychology*, vol. 30, ch. Promotion and prevention: Regulatory focus as a motivational principle, pp. 1-46, New York: Academic Press, 1998.
12. E. R. Hilgard and G. H. Bower, *Theories of learning.* Century Psychology Series, 1975.
13. R. M. Hogarth, *Educating intuition*, The University of Chicago Press, 2001.
14. R. M. Hogarth, B. J. Gibbs, C. R. McKenzie, and M. A. Marquis, *Learning from feedback: Exactingness and incentives*, Journal of Experimental Psychology: Learning, Memory and Cognition 17(4) (1991), 734-752.
15. R. W. Kulhavy, *Feedback in written instruction*, Review of Educational Research 47(1) (1977), 211-32.
16. R. W. Kulhavy and R. C. Anderson, *Delay-retention effect with multiple-choice tests*, Journal of Educational Psychology 63 (1972), 505-512.
17. R. W. Kulhavy and W. A. Stock, *Feedback in written instruction: The place of response certitude*, Educational Psychology Review 1(4) (1989), 279-308.

18. W. T. Maddox, F. G. Ashby, and C. J. Bohil, *Delayed feedback effects on rule-based and information-integration category learning*, [Online] Available http://love.psy.utexas.edu/ love/concepts/Maddoxetal2003.pdf (2003).
19. B. J. Mason and R. Bruning, *Providing feedback in computer-based instruction: What the research tells us*, [Online] Available http://dwb.unl.edu/Edit/MB/MasonBruning.html.
20. G. R. Morrison, S. M. Ross, M. Gopalakrishnan, and J. Casey, *The effects of feedback and incentives on achievement in computer-based instruction*, Contemporary Educational Psychology (1995), no. 20, 32-50.
21. E. H. Mory, *The handbook of research for educational communications and technology*, ch. Feedback Research, pp. 919-956, New York: Macmillan, 1996.
22. W. J. Roper, *Feedback in computer assisted instruction*, Programmed Learning and Educational Technology 14(1) (1977), 43-49.
23. L. J. Schooler and J. R. Anderson, *The disruptive potential of immediate feedback*, Proceedings of the Twelth Annual Conference of the Cognitive Science Society (1990), 702-708, Cambridge, MA.
24. B. Schwartz, *Reinforcement-induced behavioral stereotypy: How not to teach people to discover rules*, Journal of Experimental Psychology: General 111 (1982), 23-59.
25. T. R. Stewart, *Improving reliability of judgment forecast* In J. S. Armstrong (Ed.). Principles of Forecasting: A Handbook for Resaearchers and Practitioners (pp. 81 – 106). Kluwer Academic Publishers (2001)
26. E. L. Thorndike, *Educational psychology. volume 1: The original nature of man*, New York: Columbia University, Teachers College, 1913.

Universitat Pompeu Fabra
E-mail address: piotr.golec@upf.edu