An analysis of calibration; the hard-easy effect and the emotional disappointment of overconfident behavior: Some experimental evidences

Preliminary versions; prepared for the IAREP/SABE – Halifax, 2009

Pablo Fajfar; Economics and Mathematics Department of the Economic Science Faculty at Buenos Aires University – Argentina.
Correspondence to: pffajfar@yahoo.com.ar

Natacha Gurman; Faculty of Psychology at Buenos Aires University – Argentina.
Correspondence to: natachagurman@hotmail.com

Abstract

The over/underconfidence behavior, like a bias, has been explained as a cognitive process involved on the decision maker. Such cognitive process has been principally justified on the difficulty of the task or the problem to decide, the so called "hard easy effect" (Lichtenstein and Fischhoff; 1977). In this paper we demonstrate that underconfidence bias dramatically increases whenever one's decision consequences are immediately known. In this way, we propose that a closer relationship between the subjective performance expectations and the effective or real performance makes an alert signal on subject. Such signal activates the psychical principle of loss aversion and makes most subjects turn into underconfident.

1. Introduction

Calibration is defined as the difference between the subjective performance expectations and the objective or real performance in a decision-making process (Oskamp, 1965; Lichtenstein, Fischhoff & Phillips, 1982; Yates, 1990). When the subjective performance expectations for any decision maker are greater than their real performance, they are defined as overconfident subjects. On the other hand, when the subjective expectations are lower than the real performance they are defined as underconfident subjects. The over/underconfident behavior, like a bias, has been explained as a cognitive process involved on the decision maker (Björkman, 1994; Gigerenzer, Hoffrage & Kleinbölting, 1991; Juslin, 1993a, 1993b, 1994). Such cognitive process has been principally justified on the difficulty of the task or the problem to decide. The main previous in this way were Lichtenstein and Fischhoff (1977) who found that the overconfidence scenario increases/decreases (in terms of the questionnaire item sample) when the difficulty of the item-task increases/decreases. That's for lower effective performance (lower level of correct answers) they find overconfidence, and for greater they find underconfidence. This funding was the so called "*hard- easy effect"*.

The most contemporary studies about over/underconfidence keep focusing their attention in terms of the *hard- easy effect,* or well, the difficultly of the task (Griffin & Tversky, 1992; Keren, 1991, 1997; McClelland & Bolger, 1994; Pulford & Colman., 1997; Juslin et al, 2000; Merkle 2008). Nevertheless, none of these studies (including others not quoted here) pay attention to the temporary relationships between the subjective expectations –i.e., the subjective expected performance- and the effective or actual performance. In this way, we must begin with the fact that nobody wants to see

him/herself as an overconfident subject. Those people tend to avoid the narcissist wound that overconfidence generates (Sigmund Freud “Introduction to Narcissism” – 1914). According to this; the psychical principle off loss aversion can be intensified whenever the subject can’t transfer or process his own disappointments along time. Formally; suppose that you have just finished a grade examination. Now, you must declare your expected performance! In which of these scenarios would you be less overconfident?

Scenario I: Your real performance or result would be given to check-up in 10 minutes.

Scenario II: Your real performance or result would be given to check-up next week.

It’s natural that you will be less overconfident in the first one. Now why? Your reality principle of loss aversion or overconfidence avoidance is activated now.You ovoid the overconfidence state and you know that your real performance is given up. You prefer to see yourself as an underconfident subject than overconfident. The underconfidence sate makes you happier!

In the following experiment we demonstrate that underconfidence bias dramatically increases whenever one's decision consequences are immediately known. In other words, we present an alternative explanation for the over/underconfidence phenomena.

2. Our Experiment

One hundred and fifty five voluntary students (83 males and 72 females) of Economics and Social Sciences at three different Universities were requested to answer a 34 multiple-choice questionnaire from the Baires´s test of verbal performance (Cortada de Kohan; 1999, 2003).[1] Students were randomly recruited by a public electronic announces – web pages - made at the three Universities. For every correct answer students were paid $1, without discounting any monetary value for the incorrect ones. At the end of the questionnaire we asked the participants to report the number of questions they believe they answered correctly – i.e, the money they expect to have accumulated.

2.1. Experiment Treatments

Two treatments have been put into practice. In the first one (TR1- 79 subjects) the participants knew they would be informed of their effective or actual performance-i.e, the money they have accumulated- once the task was finished. In the second one (TR2- 76 subjects) the participants knew they would be informed of their effective performance after a week.

In both treatments participants receive their payments-i.e, the money they have accumulated for the correct answers- after a week and at the same day.

[1] The Baires´s test contains a 34 multiple choice items, with four options per item, where there's only one possible correct answer. The first 17 items present a noun and four possible definitions. The 17 remaining items, present a noun and four possible synonymous.

3. Hypothesis

Starting with the fact that a subject is defined over / underconfident if the expected numbers of correct answers are greater / smaller than the actual correct answers (Lichtenstein, Fischhoff & Phillips, 1982), we propose that:

The over/underconfident behaviour, like a bias, temporarily depends on the relationship between the subjective performance expectations and the actual or effective performance. Nevertheless, the difficulty of the task or the problem to decide is the same in both treatments; the underconfidence bias dramatically increases whenever one's decision consequences are immediately known! According to this, we expect to find a greater number of underconfident subjects in TR1.

4. Experiment Output's:

Table 1 presents the main descriptive statistics from the TR1 & TR2 experiment treatments.

Table 1

Experimental results. Data:	**Treatment 1 ;TR1 (N=79)**	**Treatment 2 ;TR2 (N=76)**	**Diferences (Wilcoxon/Mann-Whitney)**
Overall data:			
Age	22	25	
	(4.28)	(4.9)	
Gender	44 % males	64 % males	
Calibration (Mean)	-3.127	-0.0921	z = 4.716; p-value = 0.0000
	(3.82)	(4.00)	
Overconfidence Cases:	***11 Subjects; 14 %***	***35 Subjects; 46%***	
Calibration (Mean)	2.36	3.26	z = 1.818; p-value = 0.0691
	(2.3)	(2.33)	
Number of Correct Answers (Mean)	20.9090/34; 61.5%	19.628/34; 57.73%	z = -0.854; p-value = 0.3930
	(3.59)	(4.34)	
Expected Number of Correct Answers (Mean)	23.2727/34; 68.45%	22.886/34; 67.31%	z = -0.417; p-value = 0.6765
	(3.52)	(4.35)	
Correlation Expected Nº of Correct Ans. - Nº of Correct Ans.	0.6177; p- value = 0.0428	0.8727; p- value = 0.000	
Underconfidence Cases:	***60 Subjects; 76 %***	***34 Subjects; 45%***	
Calibration (Mean)	-4.55	-3.559	z = 1.640; p-value = 0.101
	(2.9711)	(2.57)	
Number of Correct Answers (Mean)	22.9833/34; 67.6%	21.4117/34; 63%	z = -1.856; p-value = 0.0635
	(3.44)	(4.356)	
Expected Number of Correct Answers (Mean)	18.433/34; 54.21%	17.8529/34; 52.5%	z = -0.726; p-value = 0.4678
	(3.7)	(4.887)	
Correlation Expected Nº of Correct Ans. - Nº of Correct Ans.	0.6512; p-value = 0.0000	0.8688; p-value = 0.0000	
Calibres Cases	***8 Subjets; 10%***	***7 Subjets; 9%***	
Number of Correct Answers (Mean) = Expected Number	20.75/34	21.2857/34	z = 0.661; p-value = 0.5089
	(3.5757)	(4.19)	

(Standard Deviation between brackets)

As shown in Table 1, the mean calibrations coefficient measured as the difference between the expected number of correct answers –i.e.,subjective expected performance- and actual correct answers –i.e., objective performance- was smaller in TR1 than TR2 (-3.127 TR1 vs. -0.0921 TR2). That's, the TR1 presented us a greater underconfident bias behavior. According whit the Wilcoxon/Mann-Whitney´s statistical test, we can reject the hypothesis that the TR1 expresses the same underconfident biases than TR2.

When we take into account the number of over/underconfidence cases, the TRI show us a 14 % overconfidence / 76 % underconficence relationship case, against a 46 % / 45% for TR2. It's clear that the numbers of underconfident subjects are significantly greater in TR1 than in TR2. That means, the greater TR1 underconfidence bias is explained as the difference between the number of underconfident subjects, but not by the over/underconfidence calibrations magnitude.

About the last concepts, the mean calibrations coefficient for overconfident subjects was statistically the same between TR1 and TR2: 2.36 for TR1 and 3.26 for TR2, p-value = 0.0691. The same phenomenon occurs for the underconfident subjects, where the mean calibrations coefficient was: -4.55 for TR1 and -3.559 for TR2, p-value = 0.101. For additional explanations, notice that the Expected Number of Correct Answers / Number of Correct Answers relationships, were not statistically different between the treatments.[2]

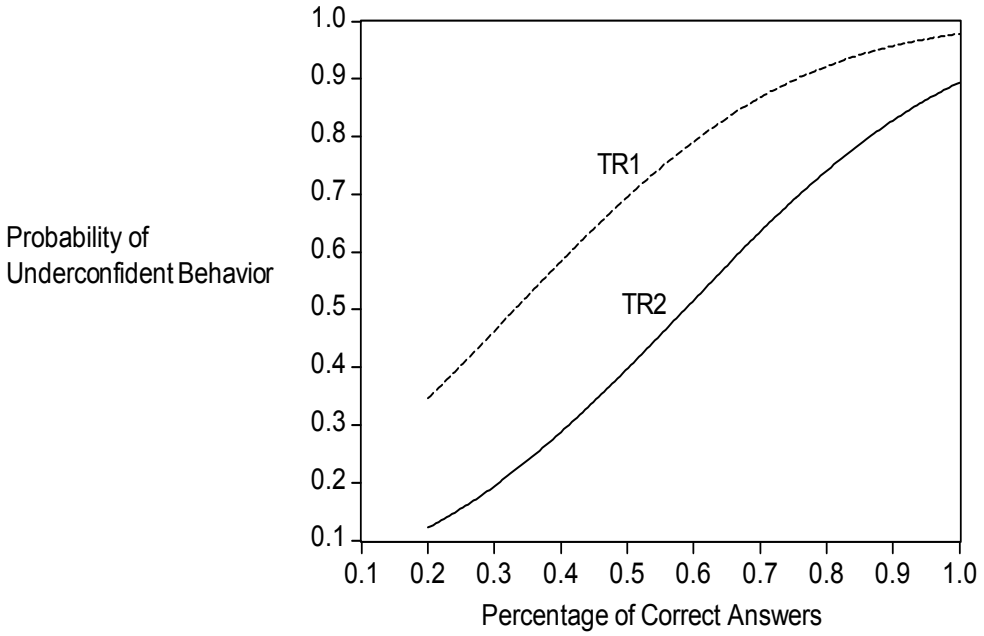
In terms of the correlation between the subjective expected performance (Expected Number of Correct Answers) and the objective or actual performance (Number of Correct Answers) they are smaller for overconfident and underconfident subjects in

[2] For the overconfident subjects, the mean of Expected Number of Correct Answers / Number of Correct Answers outputs were: 23.2727 -TR1 and 22.886 -TR2 (p-value 0.6765) / 20.9090 -TR1 and 19.628 -TR2 (p-value 0.3930). For the underconfident subjects, the mean of Expected Number of Correct Answers / Number of Correct Answers outputs were: 18.433 -TR1 and 17.8529 -TR2 (p-value 0.4678) / 22.9833 -TR1 and 21.4117 -TR2 (p-value 0.0635).

TR1. That means that the TR1 subjects show a lower confidence transference for all levels of objective or real performance.

In addition; according whit our proposal hypothesis, the number of underconfident subjects are significantly greater in TR1 than TR2. That means "the underconfidence scenario increases whenever the own consequences for a decision making process -i.e, the subjective performance expectations - are immediately know". For a better explanation, graph 1 presents us the *probability response curves of underconfident behavior* between TR1 & TR2 (probit model estimation). The probability of underconfident behaviour, like a sample phenomenon, is greater on a 30% in TR1 than in TR2. That's the probability of finding an underconfident subject when he/she knows that the objective or actual performance feedback is given ones the task is finished , was 30 % greater than when he/she knows that the real performance results are given after a week.

Graph I; Probability Response Curves
of Underconfident Behavior between TR1 -TR2 (Probit Estimation)



The vertical axis expresses the probability of underconfident behaviour as a result of the binary probit model estimation (TR1 + TR2). The horizontal axis expresses the percentage of correct answers as a difficult task indicator (Hard-Easy Effect). For all correct answers percentage, the probability of underconfident behaviour is statistically greater in TR1.

5. Some Conclusions and Discussion Proposals

Conclusion:

The over/underconfident behaviour, like a bias, temporarily depends on the relationship between the subjective performance expectations and the actual performance. In this way, we find that when subjects know that their actual performances are given immediately, the underconfidence scenario predominates above the overconfidence one. The TR1 shows us a 76 % of underconfidence subjects, against 45 % in the TR2. Note that the probability of underconfident behaviour, like a sample phenomena (probit model estimation), is greater on a 30% in TR1 than in TR2. That's the probability of finding an underconfident subject when he/she knows that the actual performance are

given once the task is finished, was 30 % greater than when he/she knows that the real performance are given after a week. Nevertheless, the over/underconfidence behaviour as a calibration measure was not significantly different between over/underconfident subject in TR1 and TR2. That means; the difference existed in terms of number of over/underconfident subjects, but not in a calibration magnitude.

In terms of the confidence transference, the correlation between the subjective expected performance (Expected Number of Correct Answers) and the objective performance (Number of Correct Answers) was smaller for overconfident and underconfident subjects in TR1. That means that TR1 subjects show a lower confidence for all levels of objective or real performance.

Discussion Proposals:

What kind of explanations can be found for this experiment?

It's reasonable that nobody wants to see him/herself as a loser. The overconfidence knowledge state, like a loser state, is a narcissus wound for anyone. The temporary relationships between the subjective performance expectations and the effective or real performance make an alert signal on subjects. Such signals say: "Remember that your actual performance is given in a few minutes, and remember also that you will avoid seeing yourself as an overconfident subject". The psychical principle of reality about overconfidence avoidance is activated now, and subjects prefer to see themselves underconfident. Underconfidence makes them happier!

Curiously, a closer nearness between expectations and results doesn't make a better calibration performance. In fact, they generate underconfident behaviours. Imagine now

a world in which every economic and political announcement is made with a short temporally relationship whit the effective or actual results. Can they be as overconfidence like they normally are?

References:

Björkman, M. (1994). Internal cue theory: Calibration and resolution of confidence in general knowledge. *Organizational Behavior and Human Decision Processes, 58*, 386-405.

Freud, Sigmund (1911). Formulations on the two principles of psychic happen. Vol. 12 Editorial Amorrortu.

Freud, Sigmund (1914). Introduction to Narcissism. Vol. 14 Editorial Amorrortu.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98,* 506-528.

Griffin, D., & Tversky, A. (1992). The weighting of evidence and the determinants of confidence. *Cognitive Psychology, 24,* 411-435.

Juslin, P. (1993a). *An ecological model of realism of confidence in one's general knowledge* (Acta Universitatis Upsaliensis : Studia Psychologica Uspaliensia, 14). Stockholm: Almqvist & Wiksell.

Juslin, P. (1993b). An explanation of the "hard – easy effect" in studies of realism of confidence in one's general knowledge. *European Journal of Cognitive Psychology, 5,* 55-71.

Juslin, P. (1994). The overconfidence phenomenon as a consequence of infor experimenter – guided selection of almanac items. *Organizational Behavior and Human Decision Processes, 57,* 226-246.

Juslin, P., Winnan, A., & Olsson, H. (2000). Naïve Empiricism and Dogmatism in Confidence Research: A Critical Examination of the Hard – Easy Effect. *Physiological Review, 2*, 384 -396.

Keren, G. (1991). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica, 77,* 217-273.

Keren, G. (1997). On the calibration of probability judgments: Some critical comments and alternative perspectives. *Journal of Behavioral Decision Making, 10,* 269-278.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance, 20,* 159-183.

Lichtenstein, S., Fischhoff, B. & Phillips, L.D. (1982). Calibration of subjective probabilities: The state of the art up to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.). *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). New York: Cambridge University Press.

McClelland, A. G., & Bolger , F. (1994). The calibration of subjective probabilities: Theories and models 1980-1993. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 453-482). Chichester, England: Wiley.

Merkle, E., C. (2008). The disutility of the hard – easy effect in choice confidence. *Psychonomic Bulletin & Review.*

Oskamp, S. (1965). Overconfidence in case-study judgments. *The journal of Consulting Psychology, 29,* 261-265.

Puldford, B. D., & Colman, A, M. (1997). Overconfidence: Feedback and item difficulty effects. *Personality and Individual Differences, 23*, 125-133.

Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice Hall.