STRATEGIC COST ADVICE AND THE EFFECT OF COGNITIVE ERROR

Chris Fortune[1] and Mel Lees[2]

[1]*School of the Built Environment, Liverpool John Moores University, Clarence Street, Liverpool L3 5UG, UK*
[2]*School of Construction and Property Management, University of Salford, Bridgewater Building, Salford, M7 9NU, UK*

The formulation of early stage building project cost advice for clients requires the professionals concerned to exercise judgement. The exercise of judgement is a human cognitive process that can be subject to errors, bias and heuristics. One of the groups of error that affects judgement is "cognitive error". This group includes sample size error, base rate error and logic error. This study identifies that construction professionals make systematic errors of judgement due to cognitive error.
The paper reports the development of an appropriate measuring instrument and the results of its application to a group of thirty-six practitioners. Subjects were tested on their propensity to make errors in judgement via text problems that were set in their own subject specific domain. The results of the work revealed that the subjects displayed the same level of error in response to the context-based word problems as had been displayed in previous studies on other biases. The paper goes on to report the findings of a follow up study to determine the significance of these errors in practice. The paper concludes by setting out the case for the development of training a package to help practitioners to deal with the propensity to make errors.

Keywords: cognitive, early cost advice, error, judgement

INTRODUCTION

Fortune and Lees (1996) reported an empirical study that obtained a practitioner assessment of the relative performance of the building project cost modelling techniques actually used in practice. The extent to which a particular cost model relied upon judgement was found to be a factor that influenced a model's incidence-of-use. The results of that study showed that judgement could be considered as being either a positive or a negative influence on the quality of advice provided as perceived by the practitioner. The positive component is the need of the practitioner to adjudicate on the raw outcome of a process or a technique. The negative side is the nature of the human judgement and the propensity of individuals to make errors. Thus it can be seen that research designed to improve the quality of early stage strategic cost advice must also address the development of a better understanding of the role of human judgement in its formulation. Previous work has established that practitioners make errors due to anchoring and adjustment bias in arriving at judgements when confronted with problems unrelated to their industry context - see Fortune and Lees (1998). In addition the potential for humans with particular learning characteristics to make similar cognitive errors has also been investigated - see Fortune and Lees (1997). This study seeks to make a further contribution to the research in the field of the formulation of strategic cost advice by ascertaining whether construction

Fortune, C and Lees, M (2000) Strategic cost advice and the effect of cognitive error. *In:* Akintoye, A (Ed.), *16th Annual ARCOM Conference*, 6-8 September 2000, Glasgow Caledonian University. Association of Researchers in Construction Management, Vol. 2, 615-22.

professionals have a propensity to make errors of judgement, due to anchoring and adjustment bias, when the problems are set in an industry context.

The paper firstly sets out the wider context for the study and then reports on the development and application of an appropriate measuring instrument to thirty-six subjects drawn from practice in quantity surveying. The results of the investigation are then analysed using *Minitab for Windows (v12*) and the paper concludes by setting out the case for the development of training a package to help practitioners to deal with the propensity to make errors.

CONTEXT

Following a review of literature related to early stage building project price forecasting and judgement Raftery (1995) asserted that reliable strategic cost advice required the input of human judgement. However, both Raftery (1995) and Birnie (1995) pointed out that humans make mistakes when making judgements and they stated that more work was needed to understand the behavioral processes involved as this may be a source of significant error.

Tversky and Kahneman (1974) and Kahneman, Slovic and Tversky (1985) provided a comprehensive review of the literature related to the existence of systematic biases that affect judgement. They asserted that in making judgements under uncertainty people in general do not appear to follow the calculus of chance or the statistical theory of prediction, instead they rely upon a number of simplifying strategies or heuristics that direct their judgements. A heuristic is a non-rational method of determining a judgement, which may or may not lead to an error. It is a 'rule of thumb' that supplants a rational decision-making process. Such heuristics can sometimes lead to reasonable judgements and sometimes lead to severe and systematic errors. The heuristics that were acknowledged as being generalisable across the population were (1) the availability heuristic, (2) the representative heuristic and (3) the anchoring and adjustment heuristic. The potential biases attributed to each of these heuristics were listed by Bazerman (1993) as being: ease of recall, retrievability, presumed associations (the availability heuristic), insensitivity to base rates, insensitivity to sample size, misconceptions of chance, regression to the mean, the conjunction fallacy (the representativeness heuristic), insufficient anchor adjustment, conjunctive and disjunctive events, overconfidence (the anchoring and adjustment heuristic). Mak and Raftery (1992) acknowledged that the majority of the research in cognitive psychology has lead to a common understanding and acceptance of the existence of the above listed heuristics and biases in lay people thinking intuitively and making judgements about problems. However, they pointed out that there was as yet no consensus in the literature on bias. In particular they noted that there was little empirical evidence of the propensity for bias in judgements made by experts considering context related problems.

One such study was carried out by Mak and Raftery (1992) in an experiment with quantity surveying students in a simulated price-forecasting situation. Their conclusions indicated that there was little support for the existence of severe and systematic bias and that the previous research findings on the existence of generalised bias may have been too pessimistic when practitioners were asked to make judgements on matters within their own field. However, Mak and Raftery (1992) went on to point out that their work had methodological limitations and they

suggested that further work with experienced practitioners and subjects from different institutions would be needed to validate their work.

This study seeks to add to the empirical evidence so far collected on practitioners' judgements on word problems set in their own subject- elated domain. The study has centred on the ascertainment of practitioners' propensity to make cognitive errors when making judgements on work related word problems

METHOD

General

The central problem facing this study was the re-working of the measuring instrument used in the previous study (Fortune and Lees 1997) to introduce an industry context. The instrument takes the form of a series of problems that the subject is required to attempt. In the 1997 study these were taken from Bazerman (1993) and were, therefore, not set in a construction context. Beach *et* al (1987) criticised the approach of asking practitioners to solve problems that were not set in the context appropriate for their expertise. They argued that this would inevitably lead to evidence of error, as the subjects did not apply themselves to the task in hand. This paper focuses on cognitive error but previous papers have covered the four main sources of error - cognitive, availability, representative and anchoring. The test used in that study had contained twelve questions - three for each of the main types of error. Each question dealt with a particular sub-type of the main error. Originally, the test included 36 questions with each sub-type having three questions randomly spread throughout the test, but piloting suggested that this made the test too long as the subjects became disinterested and the results less valid. The three sub-types of cognitive error are sample-size errors, base rate errors and logic errors and these are shown with the question numbers used in the test in Table 1.

For the 'sample size error' question the issue was related to site weather conditions and the carrying out of a research study to investigate them. The respondent is presented with a description of data collection by two groups. One of the groups – Group A – clearly makes a larger number of observations than the other group – Group B. The mean level of rainfall for the entire study is given. The respondent is asked to indicate which group will have observed more months where the rainfall exceeded the mean by 10%. The answer is Group B because since it contains fewer observations it is likely to have greater variability and a higher standard deviation. A respondent who indicates Group A is misunderstanding the significance of the larger number of observations and underestimating the variability of small data sets.

Table 1: Test questions and sub-types of errors

Cognitive error	**Test question number**
Sub-type of error	
Sample size error	4
Base rate error	10
Logic error	12

'Base rate error' is error deriving from a failure to recognise a basic underlying principle and results in decisions being made that ignore a fundamental truth. The question sets out certain information about a new client to whom the respondent has just been introduced. The information about the client relates to behaviour and physical attributes. The respondent is asked to indicate which of 'Chinese studies' or 'psychology' is the more likely first degree of the client. The information about the

client is not conclusive and any attempt to decide on the basis of the information alone would be a mistake. Since, in reality, there are many more students studying psychology than Chinese studies, a correct response is psychology. A respondent who indicates Chinese studies as the more likely first degree is ascribing too great a significance to the information about the client. This demonstrates a lack of objectivity and a tendency to rely on small data sets.

The final sub-type error is 'logic error'. Here the respondent is presented with straightforward logic test. The respondent is given the proposition that a particular estimator always wins a tender if he/she predicts that he/she will win. The respondent is then presented with four situations in the form of cards that have predictions on one side and outcomes on the other. The respondent is only shown one side of the cards. The respondent is asked to indicate which cards need to be turned over (to see what is written on the other side) in order to test the validity of the proposition. The first card shows a prediction that the tender will be won, the second card shows a prediction that the tender will be lost, the third shows that a tender was won and the fourth indicates that a tender was lost. The answer is the first and the fourth. The first card, because the prediction of a win must have resulted in a win, and the fourth card, as a lost tender must not have been a predicted win tender, if the proposition is true. The second card predicts a lost tender, which is not relevant to the proposition. The third card indicates a won tender, which, on the face of it, appears relevant, but in fact the proposition does not preclude a won tender when a lost one was predicted. Respondents who demonstrate logic errors are prone to misunderstanding the information at their disposal.

Piloting

The new test was piloted on a small group of practitioners to establish whether it could be understood and whether it was of an appropriate length. The results of the pilot indicated that the text of the test was appropriate and that the time-required undertaking the test, 20 minutes, was short enough to retain the interest of the subjects.

Given the limited resources available for the study it was decided to establish a sample of practitioners drawn from a variety of organisations that provide strategic cost advice to clients. The sample was determined by randomly selecting 153 organisations drawn from a database of over 2600 firms actively involved in strategic cost advice in England. Each of the 153 organisations was issued with a questionnaire and 36 completed forms were returned. The measuring instrument was applied in the autumn of 1999.

RESULTS

The propensity for error test used questions developed in context as described above. In addition to responding to the questions the subjects were asked to indicate on a scale of 1 to 4 how confident they were that their response was correct (1 - not at all sure, 4 - very sure the answer is correct). Therefore, for each question not only was it possible to identify whether an error had been made (i.e. an incorrect answer), but it was also possible to express the degree of error by using the confidence response.

For example, if a question required the respondent to indicate between two alternatives, A and B, one of the alternatives would be the correct answer. Answering incorrectly would indicate an error. But if the respondent was not sure about their response they could indicate a level of confidence of 1, if they were very confident

then they could indicate a higher confidence score. An incorrect answer with low confidence is arguably not an error at all.

The scoring system took this into account and was based on the confidence of the response minus one (i.e. a confidence level of 1 became 0, 2 became 1 and so on), which if the answer was incorrect, was expressed as a negative number. Therefore, the available scores for any question were +3, +2 and +1 for correct answers, i.e. no error; 0 were the confidence was low and it could not be assumed that an incorrect answer was indicative of an error; and -1, -2 and -3 for incorrect answers. For each sub-type of error there was a score for each of the three questions and these scores were averaged to produce an overall result for the sub-type for each subject. The results are set out in Table 2.

ANALYSIS

The analysis of the results was carried out using *Minitab for Windows (v12)*. The data shown in Table 2 was used to construct a distribution for each sub-type of error. The distribution shows how the sample as a whole responded in terms of errors of judgement. The scale of +3 to -3 is a continuous scale moving from high certainty of correctness to high incidence of error. There are two possible 'null' scenarios for the expected distribution. The first is that the subjects do not take the test seriously and that answers and confidences are generated at random. This would mean that each confidence level had the same chance of being indicated in any given response and would result in a horizontal distribution curve. The second is that the subjects do not commit errors and would indicate a level of confidence of 1 to all questions they suspected they may have got wrong. Since all confidence levels of 1 were re-graded to 0 this would result in a distribution that occurred on the positive side of the y-axis only. The actual results are shown in Figs 1, 2 and 3.

The distribution is clearly centred around negative values. The mean is –1.794 and the standard deviation 1.321. The distribution indicates that, as a group, the subjects make systematic errors of a sample size type. This means that surveyors are likely to misunderstand the variability of data and be over influenced by small data sets that are provided or used in the formulation of advice. This makes them susceptible to a lack of objectivity and likely to provide advice that is based upon errors.

For the 'base rate error' the results are far more positive. The distribution is fairly normal and centred upon a mean -0.226 with a standard deviation of 1.477.

The sample shows that a significant number of practitioners are not making base rate errors. However, a significant number are and this group will be likely to develop advice that ignores or is at odds with a fundamental truth.

On the question of 'logic errors' the sample divides itself into two camps - a small group that arrived at a correct conclusion, and a much larger one that demonstrated an error. The mean is strongly negative at –1.758 and the standard deviation is 1.838. The results show that many practitioners are susceptible to logic errors. This could result in flawed advice being given.

Table 2: Results of error test (blanks indicate incomplete responses)

Test Subject	Sample size error	**Base rate error**	**Logic error**
1	-3	0	-1
2	-3	1	-2
3	-2	0	-3
4	-2	0	-3
5	-1	-3	3
6	-3		-3
7	1	0	
8	-2	0	-2
9	-2	1	-2
10	-2	2	-3
11	-3	-1	-3
12	-1	-1	-2
13	-3	-3	-3
14	1	3	-2
15	-2	1	-1
16	-2	-1	-3
17	-1	0	-3
18	-2	-2	-3
19	-3	-2	-3
20	1	0	-2
21	-2	0	-2
22	-1	0	-3
23			
24	-2	3	-3
25	-2	1	-2
26	-3		-1
27			
28	-2		-3
29	-2	0	-2
30	-3	-2	3
31	-2	0	-3
32	-3	0	-3
33	-3	0	2
34	2	-2	-2
35	-3	-2	3
36	-1	0	-1

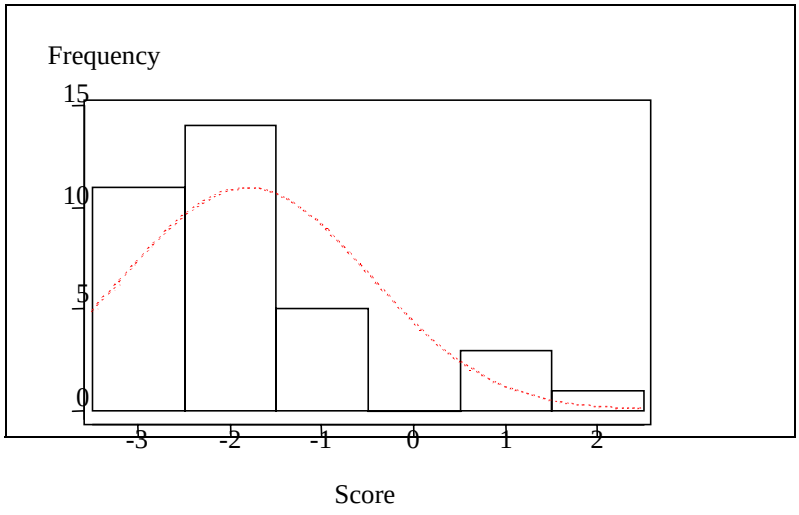

Figure 1: Distribution of responses for ‘sample size error’

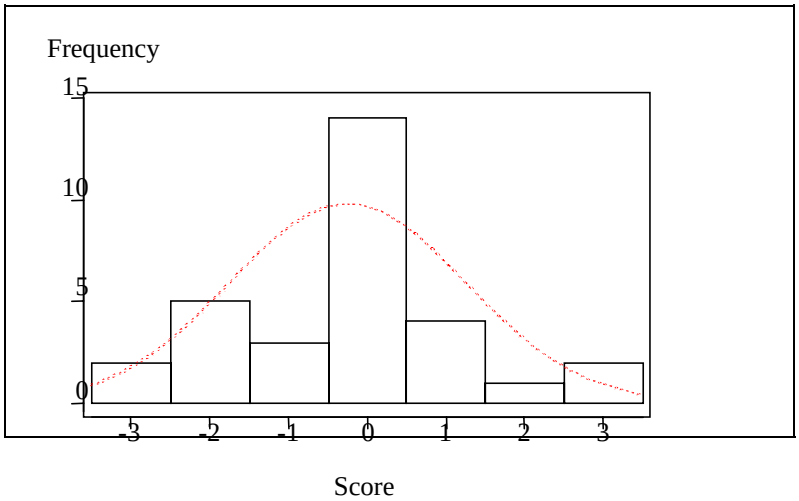

Figure 2 : Distribution of responses for 'base rate error

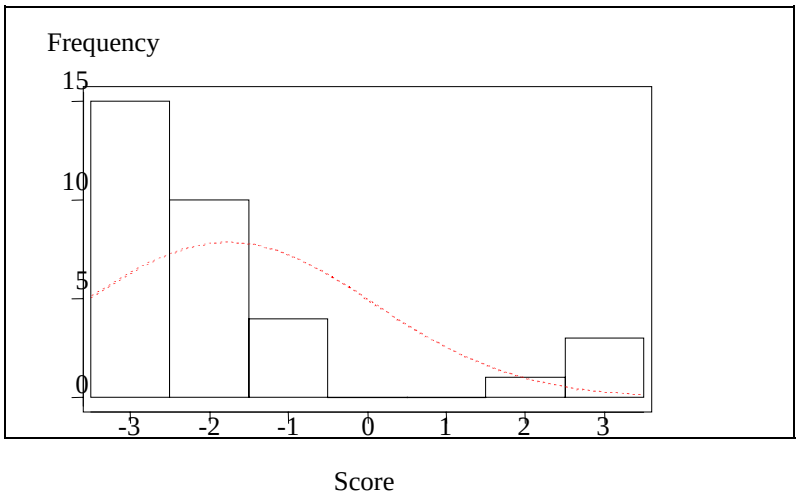

Figure 3 : Distribution of responses for 'logic error'

CONCLUSIONS

The previous discussion identified two null hypothesis scenarios - random distribution and positive skewing. The analysis of the results shows that neither of these holds true and therefore, they can be disregarded. The results show clear evidence of systematic errors of an anchoring and adjustment type. These errors have implications for the quality of strategic cost advice given to clients and, therefore, the quality of decisions made by clients when considering construction projects. The main limitation of the research is the small sample. A larger study was not possible due to resources. The recommendation for future research is that the significance of these errors in the formulation of strategic cost advice should be tested.

REFERENCES

Bazerman, H., (1993) *Judgement in Managerial Decision Making*, 3rd ed., Wiley, New York

Beach, L.R., Christensen-Szalanski, J., and Barnes, V., (1987) Assessing human judgement: has it been done, can it be done, should it be done? In Wright, G and Ayton P. (eds) *Judgemental Forecasting*, Wiley, 49-62

Birnie, J., (1995), The Possible Effects of Human Bias in Construction Cost Prediction. *Proceedings of 11th Annual ARCOM Conference*, University of York, September, 359-366

Fortune, C.J., and Lees, M.A., (1996) The Relative Performance of New and Traditional Cost Models in Strategic Advice for Clients, *RICS Research Paper Series*, **2**(2)

Fortune, C.J. and Lees, M.A. (1997) Cognitive Errors of Judgement, Learning Style Characteristics and Clients' Early Cost Advisors. *Proceedings 13th AnnualARCOM Conference*, Kings College Cambridge, September, 556-566

Fortune, C.J., and Lees, M.A., (1998) Strategic Cost Modelling for Construction Projects - The Role of Judgement. *Engineering, Architectural and Construction Management Journal*, (awaiting publication)

Kahneman, N., Slovic,P., and Tversky, A., (1985) *Judgement under Uncertainty: Heuristics and Biases.* Cambridge University Press, London

Mak, S., and Raftery, J.J. (1992) Risk Attitude and Systematic Bias in Estimating and Forecasting. *Construction Management and Economics*, **10**: 303-320

Raftery, J.J., (1995) Forecasting and Decision Making in the Presence of Risk and Uncertainty. Keynote Address to the *CIB W55 International Symposium on Economic Evaluation and the Built Environment*, Lisbon, **6**: 1-11

Tversky, A., and Kahneman, D., (1974) Judgement Under Uncertainty, Heuristics and Biases. *Science*, **185**: 1124-1131