

ISSN 1440-771X



MONASH University

Australia

Department of Econometrics
and Business Statistics

<http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/>

Structured Analogies for Forecasting

Kesten Green and J. Scott Armstrong

November 2004

Working Paper 17/04

Structured Analogies for Forecasting

Kesten C. Green,* Department of Econometrics and Business Statistics,

Monash University, VIC 3800, Australia

e-mail: kestenc@kestencgreen.com

Phone +61-3-990-55438

J. Scott Armstrong, The Wharton School, University of Pennsylvania

Philadelphia, PA 19104

e-mail: Armstrong@wharton.upenn.edu

Phone 610-622-6480

Fax 215-898-2534

Version 113

November 2, 2004

* Corresponding author. Address correspondence to e-mail address or to Department of Econometrics and Business Statistics, PO Box 11E, Monash University, Victoria 3800, Australia.

Structured Analogies for Forecasting

When people forecast, they often use analogies, but in an unstructured manner. We propose a structured judgmental procedure that involves asking experts to list as many analogies as they can, rate how similar the analogies are to the target situation, and match the outcomes of the analogies with possible outcomes of the target. An administrator would then derive a forecast from the experts' information. We compared structured analogies with unaided judgments for predicting the decisions made in eight conflict situations. These were difficult forecasting problems; the 32% accuracy of the unaided experts was only slightly better than chance. In contrast, 46% of structured-analogies forecasts were accurate. Among experts who were independently able to think of two or more analogies and who had direct experience with their closest analogy, 60% of forecasts were accurate. Collaboration did not improve accuracy.

Key words: availability, comparison, decision, method.

JEL classification: D74

It seems natural to use analogies when making decisions or forecasts as, by definition, they contain information about how people have behaved in similar situations in the past. For example, Breuning (2003) found that one-third of testimony at the Senate hearing on proposals for the first U.S. program for development aid was based on analogies. Khong (1992) concluded that most of the decisions made early in the Vietnam War were based on forecasts derived from analogies. Indeed, Kokinov (2003, p. 168) asserts "...we may explain human behavior by assuming that decisions are made by analogy with previous cases...". In the belief that such information is useful, MIT professor Lincoln P. Bloomfield has developed a historical database of post-World War II conflicts (web.mit.edu/cascon) in order to help policy analysts and others identify appropriate analogies.

We agree that information about analogies should be useful for forecasting, but we suspect that without structure people will often choose inferior analogies. They will tend to choose analogies that are easy for them to recall. The approach is called the availability heuristic by Tversky and Kahneman (1973). When people use the heuristic, they judge an outcome as likely if it matches the modal outcome of readily-recalled similar situations. Where recalling analogous situations is onerous, one analogy may suffice; particularly if it confirms prior beliefs. For example, when the U.S. Environmental Protection Agency approved a new oil refinery in Eastport, Maine, decision makers relied on the analogy of Milford Haven in the U.K (Stewart and Leschine 1986). The EPA decision makers considered Milford Haven was the most comparable site, and looked no further, but Stewart and Leschine observed that Milford Haven had not been in operation long enough to provide evidence that it was safe. They were right. The supertanker Sea Empress ran aground near Milford Haven on 15 February, 1996, spilling 70,000 tonnes of crude oil (Canada Centre for Remote Sensing 1996).

Neustadt and May (1986) described how inappropriate selection and inadequate analysis of analogies led U.S. government decision makers to make poor forecasts of the decisions of other governments' leaders. Drawing on their litany of poor decisions by political leaders, they described an elaborate structured

approach to analyzing current and historical information that they suggested should lead to a more effective use of experts' knowledge and hence to improved prediction. Many areas of judgmental decision making and forecasting have shown that structured judgmental processes make more effective use of the information people possess. This occurs, for example, when people are asked explicitly to decompose a problem (MacGregor, 2001). More generally, Armstrong (1985, Chapter 6) summarizes evidence that structured methods of judgmental forecasting are more accurate than unstructured ones. A structured approach to forecasting with analogies, then, might encourage experts to consider more information on analogies, and to process it in an effective way.

Kahneman and Lovallo (1993) report an anecdote that illustrates how inducing an expert to use analogies in a structured way can affect predictions. Kahneman had worked with a small team of academics to design a new judgmental decision making curriculum for Israeli high schools. He asked each team member to predict the number of months it would take them to prepare a proposal for the Ministry of Education. Predictions ranged from 18 to 30 months. Kahneman then turned to a member of the team who had considerable experience developing new curricula and asked him to think of analogous projects. After some consideration, the man stated that, among the many analogous situations he could recall, about 40% of the teams eventually gave up. Of those that completed the task, he said, none did so in less than seven years. Furthermore, he thought that the present team was probably below average in terms of resources and potential. In the event, the project took eight years to complete.

We were unable to identify a theory relevant to our expectation that better predictions would result from a structured approach to analyzing analogous data. Reviewers and colleagues were not able to help us. In retrospect, this is not so surprising as if such a theory was in common currency, researchers would likely already have proposed or performed tests.

Procedure for forecasting with structured analogies

Because the literature provides no evidence on how to structure forecasting with analogies, we started with a simple procedure. If analogies are useful, it is because they are similar to a target. Imposing structure on experts' assessments of similarity should encourage more complete processing of information and reduction of biases. We also wanted a procedure that would be easy for experts to use. At a minimum then, a structured approach to using analogies for forecasting requires experts to identify analogies and the outcomes they imply for the target, and to assess the analogies' similarity to the target in a structured way. Our structured analogies procedure involves five steps, two of which involve experts analyzing analogies. The administrator (1) describes the target situation, and (2) selects experts; the experts each (3) identify and describe analogies, and (4) rate similarity; the administrator (5) derives forecasts.

(1) Describe the target situation

The administrator prepares an accurate, comprehensive, and brief description. To do so, the administrator should seek advice either from unbiased experts or from experts with opposing biases. When feasible, include a list of possible outcomes for the target situation to make coding easier.

(2) Select experts

The administrator recruits experts who are likely to know about situations that are similar to the target situation. The administrator should decide how many experts to recruit based on how much knowledge they have about analogous situations, the variability in responses among experts, and the importance of obtaining accurate forecasts. Drawing upon the research on the desirable number of forecasts to combine, we suggest enlisting the help of at least five experts (Armstrong, 2001).

(3) Identify and describe analogies

Ask the experts to describe as many analogies as they can without considering the extent of the similarity to the target situation.

(4) Rate similarity

Ask the experts to list similarities and differences between their analogies and the target situation, and then to rate the similarity of each analogy to the target. We suggest providing a scale against which the experts can rate the similarity of their analogies. Ask them to match their analogies' outcomes with target outcomes.

(5) Derive forecasts

To promote logical consistency and replicability, the administrator should decide on the rules to derive a forecast from experts' analogies. Many rules are reasonable to use. For example, one could select the analogy that the expert rated as most similar to the target and adopt the outcome implied by that analogy as the forecast.

Hypotheses

We examined predictive validity using conflicts. Prior research has shown that the method currently used for making predictions for conflicts, unaided judgment, produces inaccurate forecasts (see, for example, Green 2002). We hypothesized that forecasts derived from experts' structured analysis of analogies would be more accurate than forecasts by experts who used their unaided judgment.

Our structured analogies procedure is based on the assumption that while unaided experts can provide useful information, they are not good at processing complex information reliably. For that reason, we did not rely on the experts to make forecasts but instead used a rule. On the other hand, perhaps experts' understanding of their own analogies might enable them to forecast more accurately than we could by

using rules. To test this aspect of our procedure, we asked our experts to predict the decision made in the target situation after they had described and rated their analogies.

Does it help if experts collaborate and discuss analogies with others? Collaboration could help experts to produce more analogies and flesh out the details, or it could hinder them by suppressing their creativity and search. Both positions are reasonable, so we had no prior hypothesis on collaboration. We asked some experts to collaborate with others, and all experts were asked to report the number of people they discussed the forecasting problem with.

Prior evidence

We searched for evidence on methods for forecasting with analogies. Schrodt (2002) searched for empirical evidence on the accuracy of forecasts for decisions in conflicts in the foreign policy arena. He found no evidence on the accuracy of forecasts based on analogies relative to that of forecasts based on any other method.

In a marketing study, McIntyre, Achabal and Miller (1993) tested a procedure called case-based reasoning, which is a way to structure analogies, for forecasting sales during sales promotions. When tested on two products, the forecasts were no more accurate than those of an expert buyer.

We conducted a further search for evidence by using the *Social Sciences Citation Index* (SSCI) for the period 1978 to August 24, 2004 using the terms “analogies” and “forecasting,” and then “analogies” and “prediction.” We searched the Internet on August 24, 2004 using Google™ and the terms “comparative,” “forecasting,” “prediction,” “accuracy,” and “analogies”. We conducted similar searches on JSTOR. In November 2001, we sent e-mail appeals to the 278 members of the International Institute of Forecasters list server and to the 579 members of the Judgment and Decision Making mailing list. We also contacted key researchers. The only relevant study we uncovered was Buehler, Griffin, and Ross’s (1994). The

authors asked 123 participants to estimate how long it would take to complete a computer assignment. Their predictions, made using unaided judgment, were inaccurate as they were overly optimistic. Participants who had been asked to think of analogous situations were less biased, especially when they described how the analogies related to the assignment. As a consequence, proportionately twice as many of those who recalled analogies finished their assignments before their estimated completion times.

In sum, prior to the research we describe here, little evidence was available on the accuracy of forecasts based on the use of analogies relative to the accuracy of forecasts made using other methods. Furthermore, no prior evidence exists on the use of structured analogies.

Procedures used for the study

Preparing materials

We compiled descriptions of conflicts, including brief descriptions of the roles of the parties involved in the conflict. The conflict descriptions were accounts of real situations. We abstracted all but one (Personal Grievance) from mass media reports or experts' accounts. The lead author developed the Personal Grievance from information collected in interviews and from exchanges of e-mail messages with the parties involved in the dispute. In the case of Nurses Dispute, he gathered information from published sources (Langdon, 2000a; 2000b; 2000c; *Radio New Zealand*, 2000a; 2000b; 2000c) and by interviewing representatives of the two disputant parties. When we considered it necessary, we disguised the conflicts that had already occurred to reduce the chance that our participants would know the outcomes. As a precaution, we asked our experts whether they recognized the situations. In eight cases, experts correctly identified a conflict, and their responses were eliminated.

In all, we used eight conflict situations in our research. We provided between three and six possible outcome options for each of them (Table 1). Our descriptions were short, running to no more than two pages. The full descriptions are provided at conflictforecasting.com. [For reviewers, descriptions are

attached as Reviewer Appendix 1 and outcome options as Reviewer Appendix 2.] The materials, identity of the disguised conflicts, and descriptions of actual outcomes are available to researchers on request.

Table 1

Conflict Situations

Artists protest: Members of a rich nation’s artists’ union occupied a major gallery and demanded generous financial support from their government. What will be the final resolution of the artists’ sit-in? (6 options)

Distribution channel: An appliance manufacturer proposed to a supermarket chain a novel arrangement for retailing its wares. Will the management of the supermarket chain agree to the plan? (3 options)

55% Pay plan: Professional sports players demanded a 55% share of gross revenues and threatened to go on strike if the owners didn’t concede. Will there be a strike and, if so, how long will it last? (4 options)

Nurses dispute: Angry nurses increased their pay demand and threatened more strike action after specialist nurses and junior doctors received big increases. What will the outcome of their negotiations be? (3 options)

Personal grievance: An employee demanded a meeting with a mediator when her job was downgraded after her new manager re-evaluated it. What will be the outcome of the meeting? (4 options)

Telco takeover: An acquisitive telecommunications provider, after rejecting a seller’s mobile business offer, made a hostile bid for the whole corporation. How will the standoff between the companies be resolved? (4 options)

Water dispute: Troops from neighboring nations moved to their common border, and the downstream nation threatened to bomb the upstream nation’s new dam. Will the upstream neighbor agree to release additional water and, if not, how will the downstream nation’s government respond? (3 options)

Zenith investment: Under political pressure, a large manufacturer evaluated an investment in expensive new technology. How many new manufacturing plants will it decide to commission? (3 options)

Selecting experts

To select experts, we sent e-mail messages to ten public list servers, two organizations' e-mail lists, the faculty of a university political science department, and a convenience sample of 15 experts. We chose lists that were likely to include high proportions of experts on conflicts or on judgmental forecasting. We took additional steps to ensure people were suitably qualified for these tasks. In our appeals, which were personalized when possible, the lead author wrote "I am writing to you because you are an expert..." and "I am engaged in a research project on the accuracy of different methods for predicting the outcomes of conflicts..." (Appendix A). We sent only descriptions of conflicts that were likely to be relevant to the particular recipients. For example, we did not send a situation dealing with a proposed new marketing channel to experts in employment relationship disputes. Most importantly, we counted on people to recognize when they had expertise on a topic.

We sent as many as three reminders. Details of the lists and participation are provided at conflictforecasting.com. [For the purpose of review, the details are attached as Reviewer Appendix 3.]

Using the methods

In our e-mail appeal, we gave experts instructions on how to participate (Appendix A). For structured-analogies participants, our one-page questionnaires asked the experts to (1) describe each analogous situation; (2) describe their source of knowledge about it; (3) list similarities and differences compared to the target conflict; and (4) provide an overall similarity rating (where 0 = no similarity... 5 = similar... 10 = high similarity). Finally, we asked the experts to select (from a list of possible outcomes that we prepared for each target conflict) the outcome closest to the outcome of their analogy. To illustrate, a completed structured-analogies treatment questionnaire for one of the conflicts, Telco Takeover is provided as Appendix B.

Questionnaires for unaided-judgment participants first asked them to select the outcome they thought would occur. We gave them the same lists of possible outcomes that we gave to the structured-analogies participants.

We varied the order in which we attached the conflict documents to our e-mail appeals. To test our hypotheses, with our appeals we sought responses for each of the following treatments:

1. unaided judgment (no instructions on how to forecast) without collaboration,
2. unaided judgment with collaboration,
3. structured analogies without collaboration,
4. structured analogies with collaboration.

For our first appeal, we sent equal numbers of each treatment to members of the International Association of Conflict Management mailing list. The structured-analogies and collaboration treatments were more onerous for participants than unaided judgment, so we obtained relatively few responses for those treatments. As a consequence, in most of our subsequent appeals we sought responses for structured analogies with collaboration. Finally, we sought responses for combinations of conflict and treatment for which we needed more forecasts. Because we were seeking participants for their expertise, rather than as part of a representative sample of some larger group, random assignment to treatments was unnecessary. The form of collaboration was at the discretion of the participants.

Coding responses

We obtained two groups of unaided-judgment forecasts from experts. One was from the unaided-judgment treatment (62 forecasts), and the other from experts who were asked to use structured analogies but could think of no analogies (44 forecasts). We analyzed results separately for each group and the forecasts were similar; the latter group's being somewhat more accurate. We combined the two groups under the title "unaided judgment" for our analyses, reasoning that neither of these groups used structured analyses and that our action favored unaided judgment relative to the structured analogies method.

For each conflict, we derived a structured-analogies forecast from each expert's analogy information, where the information was available. It is trivial to derive a forecast from analogies information when an expert provides a single analogy. On the other hand, many mechanical schemes could be used to derive a forecast when an expert provides information on more than one analogy. To obtain a forecast, we selected the target conflict outcome implied by the analogy given the highest similarity rating by the expert. Our reasoning was that predictive validity should increase with relative similarity. Where there was a tie, we selected the outcome that had the most support from the expert's analysis of analogies. (Details on the rules for determining support are provided at conflictforecasting.com). [For the purpose of review, details of the rules are attached as Reviewer Appendix 4.] Given our uncertainties about the best procedure, we subsequently analyzed other mechanical schemes.

We asked a convenience sample of five people who knew the actual outcomes of the conflicts to rate the outcome options we provided to the research participants. The raters were told that an option that matched the actual outcome of a conflict should be given a rating of 10. Forecasts were counted as accurate if the outcome option chosen by our rule was the option that had been given the highest median rating by our raters. Outcome options were unconditional statements of decisions and did not specify timing, for example, "Expander's takeover succeeded at, or close to, their August 14 offer price of \$43-per-share."

Results

As Tetlock (1999) demonstrated, it is difficult for experts to forecast decisions made in conflicts situations. He found that forecasts by 20 experts of the outcomes of foreign-policy conflicts were no more accurate than could be expected from chance. Our results were similar. Our 66 unaided experts were correct for 32% of predictions in an unweighted average across the eight conflicts (Table 2).

As hypothesized, forecasts from structured analogies were more accurate. They were more accurate for seven of the eight conflicts. Averaging the accuracy figures across the conflicts, structured-analogies forecasts were 46% accurate ($P = 0.04$, one-tailed permutation test for paired replicates; Siegel and Castellan 1988). We used the permutation test for paired replicates to compare the differences in the percentage of correct forecasts between the two methods for each conflict (e.g., for Artists Protest, the difference between structured analogies and unaided judgment was 17%). Viewed another way, structured analogies reduced the average forecast error by 21% (where forecast error is the percentage of forecasts that were wrong)¹.

¹ We calculate average error reduction figures as $\{(100 - A_C) - (100 - A_X)\} / (100 - A_C) * 100$, where A_C is the unweighted average percentage accuracy across conflicts of the comparison forecasts (or chance) and A_X is the corresponding figure for the forecasts of interest.

Table 2

**Accuracy of structured-analogies
and unaided-judgment forecasts by experts**

Percent correct forecasts ^a (number of forecasts)

	Chance	Unaided judgment	Structured analogies	
Telco Takeover	25	0 (8)	8 (12)	
Artists Protest	17	10 (20)	27 (11)	
55% Pay Plan	25	18 (11)	57 (14)	
Personal Grievance	25	31 (13)	36 (14)	
Zenith Investment	33	36 (14)	38 (8)	
Distribution Channel	33	38 (17)	50 (12)	
Water Dispute	33	50 (8)	92 (12)	
Nurses Dispute	33	73 (15)	57 (14)	
<hr/>				
Averages (unweighted)	28	32 (106)	46 (97)	

^a Bold figures denote the most accurate forecasts for each conflict, and overall.

Value of experts' experience

We tested whether structured-analogies forecasts were more accurate when they came from experts with more experience than when from those with less. We used two measures: (1) we asked our experts how many years experience they had as “a conflict management specialist,” and (2) we asked them to rate their experience (on a scale from 0 to 10) with situations similar to the target conflict.

Structured-analogies forecasts from experts with five or more years experience as conflict management specialists were *less accurate* (average across conflicts) with 21% error reduction compared to chance, than those with less experience (26% error reduction). Furthermore, where experts gave high ratings to their experience with similar conflicts their forecasts were *less accurate* (16% error reduction) than where they gave themselves lower ratings (31%).

Effect of number of analogies

We found that forecasts based on data from experts who could think of two or more (plural) analogies were more accurate than those based on data from experts who recalled a single analogy for six of the eight conflicts. Accuracy averaged 38% for forecasts derived from single-analogy data, but 56% for those derived from plural-analogy data ($P = 0.02$, one-tailed permutation test for paired replicates).

All else being equal, conflicts with more outcome options are more difficult to forecast than those with fewer options. To control for this, we examined the reduction in error versus chance. Forecasts based on recall of a single analogy reduced error by an average of 15% compared to chance, while forecasts derived from plural analogies reduced error by 39% (Table 3). The difference in error between single-analogy forecasts and plural-analogy forecasts is $P = 0.02$ using the one-tailed permutation test for paired replicates. The error was reduced by 42% versus chance by accepting data only from experts who described three or more analogies. Thus the usefulness of an individual expert was related to the number of analogies he described.

Table 3

Accuracy of forecasts by number of analogiesPercent error reduction versus chance ^a (number of forecasts)

	None ^b		One only		Two or more	
Telco Takeover	-33	(8)	-33	(5)	-14	(7)
55% Pay Plan	-33	(2)	26	(9)	73	(5)
Distribution Channel	-19	(5)	0	(6)	50	(6)
Artists Protest	-3	(7)	-3	(7)	40	(4)
Personal Grievance	20	(5)	0	(8)	33	(6)
Water Dispute	25	(8)	100	(4)	81	(8)
Zenith Investment	25	(6)	-12	(4)	25	(4)
Nurses Dispute	100	(3)	40	(10)	25	(4)
Average error reduction	10	(44)	15	(53)	39	(44)
(unweighted)						
Average % correct	34		38		56	
(unweighted)						

^a Bold figures denote the most accurate forecasts for each conflict, and overall.

^b Forecasts from experts we asked to use the structured-analogies method, who were unable to think of analogies. We classified these forecasts as unaided-judgment forecasts in all our other analyses.

Effect of experts' familiarity with their analogies

We expected that the information experts provided would be more useful the more closely involved they had been in the analogous situations they identified, because they would be likely to know more about the situations. For example, someone who was an adult during the Vietnam War is likely to know more about that situation than someone born since, and someone who fought in the war is likely to know more again. To examine this, we identified forecasts that had been based on analogies from either experts' own experiences (45) or that of close others (5 forecasts based on the experiences of, for example, a wife or brother-in-law). In an unweighted average across the eight conflicts, these direct-experience forecasts were more accurate (49%) than the 45 forecasts based on analogies from third-party accounts (37%); $P = 0.07$, one-tailed permutation test for paired replicates. Viewed another way, the forecasts based on analogies from experiences close to experts reduced the average error across conflicts by 31% (compared to chance) while forecasts that were based on indirect experience provided only 13% error reduction.

Familiarity and plural analogies

The ideal situation when forecasting with structured analogies is to find experts who can think of many analogies with which they have had direct experience. When our experts were able to think of two or more analogies and they had direct experience with the analogy that was most similar to the target, structured-analogies forecasts were 60% accurate (23 forecasts). In other cases, 72 forecasts were 39% accurate ($P = 0.04$, one-tailed permutation test for paired replicates).

Mechanical schemes to derive forecasts

We wondered whether experts who had used the structured analogies process then provided forecasts that were more accurate than unaided experts. They did. Their predictions were on average 42% accurate (94 forecasts) compared to 32% for unaided-judgment forecasts ($P = 0.06$, one-tailed permutation test for paired replicates). As we anticipated, however, a structured mechanical process was more effective for deriving forecasts from the experts' analogies information than experts' own judgments. As we have

seen, structured-analogies forecasts were 46% accurate. Why the difference when experts derived their own forecasts? Analogies are only useful if they are used. In 22 cases, experts made forecasts that were inconsistent with the outcomes of their own analogies; of these, 25% were accurate. When the mechanical rule was used to derive forecasts from these experts' analogies, 45% were accurate.

When experts thought of more than one analogy, our mechanical scheme did not use all of the analogical information to make predictions. We tested four alternative approaches in order to determine whether we would improve accuracy further if we derived combined forecasts from all of the 210 analogies with similarity ratings and implied decisions. For example, if an expert provided information on three analogies, for the purpose of testing our four combining alternatives we effectively derived three forecasts instead of the one we would have derived using the structured analogies method.

For our first alternative, we used the outcome implied by the most analogies, and obtained an average accuracy of 40% across all conflicts, compared to 46% for the approach we had adopted. For the second, instead of assuming that the analogies were all of equal value as we did for the first alternative, for each conflict we chose the option with the highest total similarity rating as our forecast (39% accurate). For the third alternative, we allocated experts' analogies in proportion to each option's share of the expert's similarity ratings (40%). For the fourth alternative, we calculated each expert's average similarity rating for each option. We then weighted his analogies by that average as a proportion of his total average similarity ratings (39%). In sum, all of these alternatives provided forecasts that were less accurate than those derived by applying the mechanical scheme that we had specified prior to testing the accuracy of structured analogies.

Effect of collaboration

While we had no directional hypothesis about collaboration, we analyzed the data to see whether collaboration among experts was useful. When experts using structured analogies collaborated with others, their median working time was 45 minutes compared to 30 minutes for those who worked alone. (We do not know how much time the collaborators spent on the task, nor do we know the nature of their collaboration.) As it happened, those who collaborated claimed to have had much more experience with conflict-management (median of 14 years versus 5 years) and experience with similar conflicts (a median self-rating of 4.0 out of 10, versus 2.8). Despite the greater investment of resources by more knowledgeable experts, collaboration produced no gain in accuracy: forecasts from solo experts were on average 44% accurate across conflicts (75 forecasts), compared to 42% for forecasts by collaborating experts (22 forecasts).

Given our findings, we saw no need to distinguish between solo and collaborative forecasts in our analysis. In view of the time savings, we recommend that structured analogies be done by individuals.

Limitations

The structured analogies method is useful only in cases in which experts can think of analogies. This limitation can be overcome in many situations by identifying people with relevant expertise. While this may be difficult to know in advance, one can gauge people's expertise from their responses – that is, did they provide analogies; if so, how many; and did they have direct experience?

Using structured analogies is more costly than using unaided judgment. However, relative to the costs of making bad decisions in many conflict situations, such as selecting strategies to achieve peace in the Middle East or to deal with threatening behavior by the North Korean government, the costs are negligible.

Further research

Research on additional situations would help to better assess the improvements that might be expected, and the conditions under which structured analogies is most effective. Our conclusions are based on a sample of only eight situations.

This is the first published study on the use of structured analogies. More research needs to be done to develop the operational procedures for the method. For example, what is the best way to frame the issues for the experts so that they provide more and better analogies? Would a more structured approach to rating analogies' similarity to a target help administrators derive more accurate forecasts? To what extent might improvements in accuracy be obtained, in the case of well-documented analogies, by checking the facts of the situation and correcting any errors in experts' matching of analogy outcomes with potential target outcomes?

It seems plausible that the Delphi technique could be used to improve assessments of analogies' similarity to a target, potentially increasing accuracy further at a low cost. Rowe and Wright (2001) provide evidence on the value of Delphi, and software for implementing of Delphi is provided at forecastingprinciples.com. Experts' confidence ratings may be useful for weighting structured-analogies forecasts in a combination (Arkes, 2001).

We have examined conflict situations because of their importance and the difficulty of obtaining useful forecasts. Structured analogies might also improve forecasting for situations other than conflicts. We expect that it would be most useful where situations are complex and where there are plural analogies.

Research is needed on how to encourage adoption of structured analogies. Currently, people use unaided judgment, a method that is little better than chance, to decide whether to go to war, get a divorce, make a

hostile takeover bid, go on strike, or mount a competitive pricing campaign. Better forecasts would aid decision making in such situations.

Conclusions

It is difficult to forecast decisions made in conflict situations. On average, unaided experts were correct for only 32% of their predictions. This was little better than chance at 28%.

For our structured analogies method, the two key criteria for identifying an expert were the number of analogies generated, and the presence of direct knowledge about those analogies. When experts produced two or more analogies from experience, forecasts from structured analogies were correct for 60% of the predictions. Given the importance of forecasts in conflict situations and in other arenas, such improvement could have considerable benefits.

Appendix A

E-mail message appeal and instructions: Structured analogies / collaboration treatment

Subject: Using analogies to predict the outcomes of conflicts

Dear Dr _____

I am writing to you because you are an expert on _____. I am engaged on a research project on the accuracy of different methods for predicting decisions made in conflicts. At this stage, I'm investigating the formal use of "analogies" for forecasting. That is, forecasting on the basis of the outcomes of similar conflicts that are known to the forecaster.

What I would like you to do is to read the attached descriptions of some real (but disguised) conflict situations and to predict the outcome of each conflict. If you can't read the attachments, please let me know and I'll send the material in your preferred format if I'm able.

Each attached file contains a conflict description and a short questionnaire. Please follow these steps for each conflict:

- 1/ Read the description and
- 2/ try to think of several analogous situations and
- 3/ about how similar your analogies are to the conflict.
- 4/ Fill-in the questionnaire (electronically if you can)
 - a) describe your analogies
 - b) rate your analogies
 - c) make your prediction (either pick an outcome or assign probabilities)
 - d) record the total time you spent on all tasks
 - e) return the questionnaire.

One of the objectives of this research is to assess the effect of collaboration on forecast accuracy. You have been allocated to the collaboration treatment, so please *do* discuss these forecasting problems with colleagues. Do *not*, however, discuss them with other people who have received this material as I want independent responses from participants.

Although I intend to acknowledge the help of all of the people who assist with this research, my report will not associate any prediction with any individual.

Your prompt response is very important to the successful completion of my project. Please help me to prove the sceptics wrong about the level of cooperation I get!

Best regards,

...

Appendix B

Telco Takeover Bid

- 1) (A) In the table below, please briefly describe
 (i) your analogies,
 (ii) their source (e.g. your own experience, media reports, history, literature, etc.), and
 (iii) the main similarities and differences between your analogies and this situation.
 (B) Rate analogies out of 10 (0 = no similarity... 5 = similar... 10 = high similarity).
 (C) Enter the responses from question 2 (below) closest to the outcomes of your analogies.

(A) (i) description,	(ii) source,	(iii) similarities & differences	(B) Rate	(C) Q2
a. Bank takeover	Personal	Issue same, industry different	8	C
b. Govt Agency merger	Personal	Takeover same, government, but ordered takeover	4	D
c. Facility Merger	Personal/family	Combine similar operations	3	B
d.				
e.				

- 2) **How was the standoff between Localville and Expander resolved?** *(check one ✓, or %)*
- a. Expander's takeover bid failed completely
- b. Expander purchased Localville's mobile operation only
- c. Expander's takeover succeeded at, or close to, their August 14 offer price of \$43-per-share
- d. Expander's takeover succeeded at a substantial premium over the August 14 offer price
- 3) **If you have *not* given a prediction, please state your reasons:**
- 4) **Roughly, how long did you spend on this task?** [1] hours
{include the time you spent reading the description and instructions}
- 5) **How likely is it that taking more time would change your forecast?** [0] 0-10
{ 0 = almost no chance (1/100) ... 10 = practically certain (99/100) }
- 6) **Do you recognise the actual conflict described in this file?** Yes No
 If so, please identify it: _____
- 7) **How many people did you discuss this forecasting problem with?** [2] people
- 8) **Roughly, how many years experience do you have as a conflict management specialist?** [20+] years
- 9) **Please rate your experience (out of 10) with conflicts similar to this one** [6] 0-10

*When you have completed this questionnaire, please return either this document as an email attachment to...
 or this questionnaire (with your initials at right) by fax to...*

Your initials: [XYZ]

References

- Arkes, H. R. 2001. Overconfidence in judgmental forecasting. J. S. Armstrong, ed. *Principles of Forecasting*. Boston: Kluwer Academic Publishers.
- Armstrong, J. S. 2001. Combining forecasts. J. S. Armstrong, ed. *Principles of Forecasting*. Kluwer Academic Publishers, Boston.
- Armstrong, J. S. 1985. *Long-Range Forecasting*. John Wiley, New York. Full text at hops.wharton.upenn.edu/forecast
- Breuning, M. 2003. The role of analogies and abstract reasoning in decision-making: Evidence from the debate over Truman's proposal for development assistance. *International Studies Quarterly*, **47** 229-245.
- Buehler, R., D. Griffin, M. Ross. 1994. Exploring the 'planning fallacy': Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, **67** 366-381.
- Canada Centre for Remote Sensing. 1996. 'Sea Empress' oil spill monitoring. *RADARSAT Image*, Milford Haven, Wales, United Kingdom, February 22, 1996. Retrieved July 1, 2003, from http://www.ccrs.nrcan.gc.ca/ccrs/data/satsens/radarsat/images/uk/ruk01_e.html.
- Green, K. C. 2002. Forecasting decisions in conflict situations: A comparison of game theory, role-playing, and unaided judgement. *International Journal of Forecasting*, **18** 321-344. Full text at forecastingprinciples.com.
- Kahneman, D., D. Lovallo. 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, **39** 17-31.
- Khong, Y. F. 1992. *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton University Press, Princeton, NJ.
- Kokinov, B. 2003. Analogy in decision-making, social interaction, and emergent rationality. *Behavioral and Brain Sciences*, **26** 167-168. Full text at <http://www.nbu.bg/cogs/personal/kokinov/bbskokinov.pdf>
- Langdon, C. 2000a. Nurses vote today on strike. *The Dominion*, **Edition 2 - 20 September 3**.

- Langdon, C. 2000b. Nurses support call for strike. *The Dominion*, **Edition 2 - 21 September 3**.
- Langdon, C. 2000c. Nurses' pay boosted, strike off. *The Dominion*, **Edition 2 - 6 December 3**.
- MacGregor, D. G. 2001. Decomposition for judgmental forecasting and estimation. J. S. Armstrong ed. *Principles of Forecasting*, Boston.
- McIntyre, S. H., D. D. Achabal, C. M. Miller. 1993. Applying case-based reasoning to forecasting retail sales. *Journal of Retailing*, **69** 372-398.
- Neustadt, R. E., E. R. May. 1986. *Thinking in Time: The Uses of History for Decision Makers*. Free Press, New York.
- Radio New Zealand Limited*. 2000a, September 20. Brenda Wilson (Chief Executive, New Zealand Nurses Organisation) interviewed by Geoff Robinson. Morning Report, Transcript: *Newztel News Agency Ltd*.
- Radio New Zealand Limited*. 2000b, September 20. Rae Lamb (Health Correspondent, *Radio New Zealand*) interviewed by Mary Wilson with excerpted material from interviews with Annette King (Minister of Health), Susan Rolls (Emergency Nurse at Wellington Hospital), and Russell Taylor (Wellington Nurses Union Organiser). Checkpoint, Transcript: *Newztel News Agency Ltd*.
- Radio New Zealand Limited*. 2000c, September 22. Margot Mains (Chief Executive Officer, Capital Coast Health) interviewed by Geoff Robinson. Morning Report, Transcript: *Newztel News Agency Ltd*.
- Rowe, G., G. Wright. 2001. Expert opinions in Forecasting: The role of the Delphi technique. J. S. Armstrong, ed. *Principles of Forecasting*. Kluwer Academic Publishers, Boston.
- Schrodt, P. A. 2002. Forecasts and contingencies: from methodology to policy. Paper presented at the American Political Science Association meetings, Boston, 29 August – 1 September. Retrieved May 7, 2004, from <http://www.ukans.edu/~keds/papers.dir/Schrodt.APSA02.pdf>.
- Siegel, S., N. J. Castellan Jr. 1988. *Non-parametric Statistics for the Behavioral Sciences*, 2nd ed. McGraw-Hill, Singapore.

- Speke, C., S. Reuter. 2003. German military historians predict Anglo-American defeat in Iraq. *Online Journal* **March 29**. Available from <http://morris.wharton.upenn.edu/forecast/Conflicts/PDF%20files/Iraqi%20conflict%20forecast.pdf>.
- Stewart, T. R., T. M. Leschine. 1986. Judgment and analysis in oil spill risk assessment. *Risk Analysis*, **6** 305-315.
- Tetlock, P. E. 1999. Theory driven reasoning about possible pasts and probable futures: Are we prisoners of our perceptions? *American Journal of Political Science*, **43** 335-366.
- Tversky, A., D. Kahneman. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, **5** 207-232.

Acknowledgements

We thank the experts who participated in the research reported here. They included Roderic Alley, Barry Anderson, Don Baker, Corrine Bendersky, Constant Beugre, Doug Bond, Michelle Brackin, José Ramón Cancelo, Nihan Cini, David Cohen, Ike Damayanti, Serghei Dascalu, Nikolay Dentchev, Ulas Doga Eralp, Miguel Dorado, Erkan Erdil, Jason Fello, Paul Gaskin, Andrew Gawith, Kristian Skrede Gleditsch, Joshua Goldstein, David Grimmond, George Haines, Claudia Hale, Ragnar Ingibergsson, Patrick James, Michael Kanner, John Keltner, Daniel Kennedy, Susan Kennedy, Oliver Koll, Rita Koryan, Talha Köse, Tony Lewis, Zsuzsanna Lonti, Dina Beach Lynch, David Matz, Bill McLaughlan, Kevin Mole, Ben Molloy, Robert Myrtle, W. Bruce Newman, Randall Newnham, Konstantinos Nikolopoulos, Glenn Palmer, Dean G. Pruitt, Perry Sadorsky, Greg Saltzman, Amardeep Sandhu, Marlies Scott-Wenzel, Deborah Shmueli, Mike Smith, Marta Somogyvári, Harris Sondak, Dana Tait, Scott Takacs, Dimitrios Thomakos, William Thompson, Ailsa Turrell, Bryan Wadsworth, James Wall, Daniel Williams, Christine Wright, Becky Zaino. We received useful comments from delegates at the 2003 and 2004 International Symposia on Forecasting and at the Institute of Mathematics and Its Applications' Conference on Conflict and Its Resolution, and from attendees at talks at RAND Organization, Warwick Business School, University College London, Monash University, and Melbourne Business School to

whom we presented elements of the work reported here. We also thank Lisa Bolton, Nikolay Dentchev, Don Esslemont, Stanley Feder, Paul Goodwin, Clare Harries, Rob Hyndman, Oliver Koll, and Tom Yokum for providing pre-submission peer review. Editorial assistance was provided by Mary Haight, Marian Lee, and Catherine Morgan.