Game theory, simulated interaction, and unaided judgement for forecasting decisions in conflicts: Further evidence

Kesten C. Green[*]

Business and Economic Forecasting Unit, PO Box 11E, Monash University, VIC 3800, Australia

Abstract

When people in conflicts can accurately forecast how others will respond, they should be able to make better decisions. Contrary to expectations, earlier research found game theorists' forecasts were less accurate than forecasts from student role players. To assess whether game theorists had been disadvantaged by the selection of conflicts, I obtained forecasts for three new conflicts of types preferred by game theory experts. As before, role-players in simulated interactions were students, and other students forecast using their judgement. Game theorists did better than previously. However, when the three new and five earlier conflicts are combined, 101 forecasts by 23 game theorists were no more accurate (31%) than 354 forecasts by students who used unaided judgement (31%). Experienced game theorists were not more accurate. Neither were those who spent more time on the task. Of 105 simulated-interaction forecasts, 62% were accurate: an average error reduction of 47% over game-theorist forecasts and a halving of error relative to the current method. Forecasts can sometimes have value without being strictly accurate. Assessing the usefulness of forecasts led to the same conclusions about the relative merits of the methods. Finally, by combining simulated interaction forecasts, accurate forecasts were obtained for seven of the eight situations.
© 2005 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

Keywords: Accuracy; Methods; Role playing; Strategy; Usefulness

Conflicts are an important concern. Just how important, is illustrated by the pervasiveness of conflict stories in the news. For example, among the top 21 articles headlined on the homepage of *The New York Times* on August 4, 2003[1], more than half were concerned principally with conflicts in which two or more parties were interacting. They included stories

* Tel.: +61 3 990 52358; fax: +61 3 990 55474.
  E-mail address: kesten@kestencgreen.com.
  URLs: kestencgreen.com, conflictforecasting.com.

[1] www.nytimes.com at 12:03 AM Eastern Time. The articles considered were headlined in the central (two-column) panel of the page and included all unique articles from the top of the page and under the headings: National, Business, International, Editorials/Op-Ed, Washington, and New York Region.

with headlines such as "Food and peace just a memory in Liberian city", "Taliban are killing clerics who dispute holy war call", "Verizon contract talks extend into the night", and "Hollywood producer makes 2nd bid for German TV group".

It seems likely that improved forecast accuracy in conflict situations would lead to better decisions. For example, consider the management of *Verizon* (above) who were involved in negotiations with unions. If they could have predicted which offers would lead to agreements and which would not, and if the unions could have predicted which demands would be accepted, a speedier resolution and a better outcome would have been possible.

Green (2002) presented findings on the accuracy of forecasts from three methods for predicting decisions in conflicts. For each conflict, participants were asked to choose the most likely decision from a list. Five of the six conflicts used in the research involved direct or indirect interaction between parties. All were real situations. For those five conflicts, game theory experts' forecasts were little more accurate (28% of the predictions were accurate) than the unaided-judgement forecasts of university students (27% accurate). In contrast, forecasts from university students in simulated interactions (a type of role playing in which participants act out the interactions of parties in a conflict situation) were substantially more accurate (61%).

Erev, Roth, Slonim, and Barron (2002), and Goodwin (2002), in commenting on Green (2002), suggested the possibility that some types of conflict may be more amenable than others to forecasting by game theorists. I reasoned that if anyone knew what these conflict types were, it would be game theory experts who had done research on forecasting for real conflicts. In my search for prior evidence on forecasting accuracy in Green (2002) I found five such studies and from these I deduce the conflict types the researchers regarded as amenable to forecasting.

It is worth noting that not all game theory experts make a distinction between types of conflict suitable for forecasting with game theory, and types of conflict that are not. For example, Fraser and Hipel (1984) maintained that their game-theoretic method (conflict analysis) can usefully forecast any type of conflict.

In this article, I present new findings on the accuracy of forecasts from game theorists, unaided judgement, and simulated interaction for three conflicts that are dissimilar to the five in Green (2002). Details on the three forecasting methods are as described in that article. I have made some adjustments to the coding of responses reported in the earlier work, and these are described in Appendix A.

I first present evidence on the expectations people have regarding the performance of the three forecasting methods. Second, I describe the three new conflicts. Third, I assess whether the three new conflicts, as well as the five used in the earlier research, are of types preferred by game theorists. Finally, I present my findings.

1. Expectations of accuracy

Opinion is divided on whether it is appropriate to use game theory to make predictions about specific conflicts. For example, Robert Wilson of Stanford University was quoted as saying "Game theory does not offer any specific answers to any specific situation. It says something like 'these are the things to take into account'" (London, 2002). Other game theorists do recommend game theory for forecasting (Green, 2002). More recently, Camerer (2003) listed "predicting what is likely to happen" as one of the two important uses of game theory (p. 157), and Dixit and Skeath (2004) wrote "When looking ahead to situations where multiple decision makers will interact strategically, we can use game theory to foresee what actions they will take and what outcomes will result" (p. 37). Finally, the terms "game theory" and "forecasting" produced 60,900 hits using a Google search and 1630 hits using a Google Scholar search on December 21, 2004.

Given the controversy, Scott Armstrong and I thought it worthwhile to assess beliefs about forecasting among people with an interest in conflicts. To this end, prior to talking to various groups about forecasting for conflict situations, we obtained opinions from those who attended. Respondents were academics and students at Lancaster University (19 usable responses), Manchester Business School (18), and Melbourne Business School (6), as well as Harvard Business School alumni (8), Royal New Zealand Police College educators (4), conflict management practitioners in New Zealand (7), and

attendees at a Glasgow conference on organizational foresight (15). We asked these 77 respondents for their expectations regarding the accuracy of forecasts of decisions in conflicts from different methods. We told them that by selecting one outcome at random for each conflict a forecaster could expect 28% of forecasts to be accurate. Responses that included accuracy figures less than 28% were excluded from our analysis.

The median expectations of the percentage of accurate forecasts were 30% for novices' unaided judgements, 40% for simulated interaction using novice role-players, and 50% for game theorists. The respondents' implied rankings of the methods by forecast accuracy were broadly consistent (Kendall coefficient of concordance $W = 0.36$; $P < 0.001$, $k = 77$, and $N = 3$; Siegel & Castellan, 1988).

It seems reasonable to expect experts to provide forecasts that are more accurate than those of novices and that the body of knowledge possessed by game theorists would be particularly useful for making predictions about conflicts.

1.1. Game theorists' preferences

From studies that I found in Green (2002) to have used game theory to make predictions about real conflicts, I deduced the types of conflict that game theorists believe to be most amenable to forecasting. Brams and Togman (2000) and Organski (2000) used game theory to make conditional predictions on the outcomes of civil conflicts in Northern Ireland and the Middle East, respectively. Gruca, Kumar, and Sudharshan (1992) used game theory to predict incumbents' responses to a new competitor. Keser and Gardner (1999) discussed the use of game theory to predict the outcomes of common pool resource conflicts. Finally, Ghemawat and McGahan (1998) used game theory to predict the behaviour of a group of competing electricity generating companies. The writings of these researchers express or imply recommendations to apply game theory to particular types of conflict forecasting problems. The conflict types could be characterised as civil unrest, competitor responses to new entrants, common-pool resource allocation dilemmas and, as suggested by Ghemawat and McGahan, situations of concentrated competition (few parties), mutual familiarity, and ongoing interaction.

2. Conflict forecasting problems

I compiled material for three new conflicts: a personal grievance dispute about pay levels, a corporate takeover battle, and an international dispute with two countries edging towards war. They are referred to here as Personal Grievance, Telco Takeover, and Water Dispute, respectively. Copies of the material are available at www.conflictforecasting.com.

Names of the parties, and other details, were disguised in the conflict descriptions. Participants were asked if they recognised the conflicts and their responses were excluded from analysis if they correctly identified the actual situation. None of the participants in the research new to this paper recognised any of the conflicts.

Personal Grievance was a conflict over the importance of an employee's position and the consequent pay grade. A long-serving staff member of a New Zealand university students association believed her work was undervalued in the job evaluation that was commissioned by her new manager. The evaluation was conducted by the association's president. After some negotiation, the top of the salary band for the employee's position, set by the manager, was still below her current salary. The manager did not propose reducing the employee's actual salary, but it was clear she could not expect a pay increase in the foreseeable future. A Mediation Service mediator was appointed and a meeting between the parties was arranged. Potential decisions ranged from commissioning a new independent evaluation to accepting the salary band. To obtain information, I interviewed the three key people who were involved. They commented on drafts of my description and made suggestions as to what decisions might have been made at the meeting. I also pre-tested the material in ten simulations using 50 student role-players.

Telco Takeover was a conflict for the ownership of a regional telecommunications provider (*CenturyTel*) that occurred in the U.S.A. during 2001. *Alltel*, a telecommunications company, had been approached by managers of the somewhat smaller *CenturyTel* with an offer to sell *Alltel* their mobile telephone business. *Alltel* managers declined the offer. Shortly afterwards, *Alltel* made an offer to pay 40% more than the current share price to buy all of *CenturyTel*. *CenturyTel's*

long-standing chairman was a large shareholder of the company. Managers and staff also owned shares. The *CenturyTel* board was reluctant to sell and took measures to prevent an *Alltel* take-over. *Alltel* appealed directly to outside shareholders of *CenturyTel*. The decision to be made was whether a deal would be done and, if so, whether *CenturyTel's*, *Alltel's*, or some intermediate offer would be the basis. My description of this conflict was based on two articles in BusinessWeek Online (Haddad, 2001; Kharif, 2001) that were written before the conflict was resolved, and on an article in Wireless News-Factor (Wrolstad, 2002) that was written after a deal had been concluded. I obtained supporting information from the websites of *Alltel* (www.alltel.com) and of CenturyTel (www.centurytel.com). Eric W. Orts provided clarification on the relevant US law.[2]

Water Dispute was a conflict over access to the water of the Euphrates River that occurred in 1975 between Iraq and Syria. Syria had built a dam across the river and started to fill the reservoir, thereby reducing the flow of water into Iraq. Both are arid countries, with Iraq almost completely dependent on the Euphrates for water. The two Soviet-aligned military dictatorships were preparing for war with their troops massing on a common border. Saudi Arabia, in an attempt to mediate a peaceful outcome, called the parties together. The decision to be made was whether Iraq would declare war or go ahead with its threat to bomb the Syrian dam, or whether Syria would release more water voluntarily. The Water Dispute description is based on an account of the conflict in Keesing's Contemporary Archives (1975). Additional information came from Kliot (1994) and from internet searches.

Prior to using the descriptions in my research, I obtained reviews and predictions for the three situations from three experts.

2.1. Match with game theorists' preferences

In all, I compare forecasts for eight conflicts. Three are the new ones presented here and five are

from Green (2002). All eight conflicts align to some degree with game theorists' preferences. Artists Protest is an example of *civil unrest*. Two conflicts involved a *new entrant* to a market. They were Distribution Channel, in which an appliance manufacturer sought to sell its wares through supermarkets, and Telco Takeover, in which one corporation attempted to enter a new region by acquiring the major provider. Water Dispute was concerned with access to a *common pool resource*: the waters of the Euphrates River.

Five conflicts involved *concentrated competition*, *mutual familiarity*, *and ongoing interaction*. In Artists Protest the Dutch artists' union was, in effect, competing with other citizens, represented by the government, for public funds. The managers in Zenith Investment were competing for resources and status within the corporation. Personal Grievance, 55% Pay Plan, and Nurses Dispute were all employment relationship disputes and, as such, meet the three criteria for inclusion in this game theorist-preferred type.

As should be clear from earlier discussion, not all game theorists would support a claim that these eight conflicts are ones to which their knowledge might sensibly be applied. Other criteria could be proposed for selecting situations.

2.2. Match with decision makers' interests

To be useful, forecasting methods must address the diverse problems that decision makers face. Decision makers who operate in different arenas (industrial, commercial, civil, political, diplomatic, and military) and who interact with different parties (individuals, organisations, and governments) should be able to find familiar situations among the eight conflicts.

Three of the conflicts were industrial (employment relations) disputes. Nurses Dispute and 55% Pay Plan were public disputes over pay and conditions. Nurses Dispute was similar to many that are reported in the media while 55% Pay Plan was a struggle between sports team owners and players. Personal Grievance involved a single employee in conflict with her employer, whereas the other two industrial disputes involved many employees represented by their respective unions. Distribution Channel, Telco Take-

over, and Zenith Investment were commercial conflicts, and Water Dispute was a conflict between national governments. Finally, Artists Protest was a civil conflict.

Distribution Plan, 55% Pay Plan, Nurses Dispute, and Telco Takeover were conflicts between organisations. Nurses Dispute also involved a third party: an appointed mediator. Zenith Investment was a dispute between individuals, and Water Dispute was between governments. Personal Grievance was a dispute between an individual with a grievance and an organisation. Artists Protest was a dispute between an organisation (an artists union) and a government.

3. Findings

For each of the new conflicts, I obtained from university student participants ten simulated interaction forecasts and between nine and eleven unaided-judgement forecasts. From game theory experts, I obtained seven or eight forecasts for each conflict. Eight of the twenty-one game theory experts who provided forecasts reported in Green (2002) also provided forecasts for the new conflicts, as did two game theory experts who had not previously participated. One of the game theorists had provided commentary on Green (2002). Unlike the students, the game theorists were not paid for their time. Participants were not directly rewarded for accuracy but, with an understanding that their names would be published and professional pride at stake, the game-

theorist volunteers had reasons to be motivated to make accurate predictions.

The accuracy of the game-theorist forecasts for the three new conflicts varied from 0% for Telco Takeover to 75% for Water Dispute. This compares with 6–50% for the five conflicts in Green (2002) that involved interaction between the parties. Across the three new conflicts, the unweighted average accuracy of the forecasts was, at 39%, somewhat better than the 26% for the conflicts in the earlier research. In the case of unaided-judgement, the corresponding figures were 33% (new conflicts) and 27% (earlier research), and for simulated interaction 63% and 61% respectively.

In the following analysis and discussion, I use data from this research combined with data from Green (2002) and Green and Armstrong (2004). The data from the latter paper (185 forecasts) were forecasts from novices using their unaided judgement.

Using an unweighted average across conflicts, the forecasts of game theory experts were accurate for 31% of forecasts, which was no better than the accuracy of novices' unaided-judgement forecasts (31%). Simulated-interaction forecasts were more accurate, at 62% on average, than game-theorist forecasts, unaided-judgement forecasts, and chance for all eight conflicts (Table 1).

The simulated-interaction forecasts provided substantial error reduction relative to game-theorist forecasts for each of the conflicts. Error reduction is calculated here as the difference between error from game-theorist forecasts and error from simulated-interaction forecasts, expressed as a percentage of

Table 1
Accuracy of forecasts, combined data[a]

Percent correct[b] (number of) forecasts

	Chance	Unaided judgement by novices		Game theory experts		Simulated interaction with novices	
Artists protest	17	5	(60)	6	(17)	**29**	(14)
55% Pay plan	25	13	(38)	29	(17)	**60**	(10)
Distribution channel	33	15	(68)	23	(13)	**75**	(12)
Telco takeover	25	29	(34)	0	(7)	**40**	(10)
Zenith investment	33	35	(43)	22	(18)	**59**	(17)
Personal grievance	25	35	(31)	43	(7)	**60**	(10)
Water dispute	33	51	(35)	75	(8)	**90**	(10)
Nurses dispute	33	64	(45)	50	(14)	**82**	(22)
Averages (unweighted)	28	31	(354)	31	(101)	**62**	(105)

[a] Includes data from Green (2002) and Green and Armstrong (2004).
[b] Figures in bold indicate the most accurate forecasts for each conflict.

error from game-theorist forecasts.[3] Error reduction ranged from 24% (Artists Protest) to 68% (Distribution Channel), and averaged 47% across the conflicts.

While eight conflicts is a small sample, simulated-interaction forecasts were more accurate than game theorist-forecasts for all eight. The likelihood of this being due to chance is small ($P=0.004$, one-tailed permutation test for paired replicates; Siegel & Castellan, 1988).

Although the conflict situations were chosen in ignorance of the findings on relative accuracy, it is still possible that the sample might be biased against forecasting by game theorists because conflicts that are interesting to forecasting researchers are unusually difficult for game theorists to model. Following Armstrong (2002), I compared the average error reduction for conflicts that were easy for game theorists to forecast with the average for those that were hard. In the case of the four easiest conflicts (49% accurate) the reduction in error between game-theorist forecasts and simulated-interaction forecasts was 49%. For the four hardest conflicts (13% correct) the error reduction was 45%. From this test, then, there was no evidence of such a bias.

3.1. Effect of game theorists' experience and time spent on accuracy

For the forecasts of each of the 17 game theorists who provided three or more forecasts, I calculated the error reduction versus the game-theorist average error reduction for matching conflicts. On that basis, there was no correlation between the experience of the game theorists and the accuracy (error reduction) of their forecasts (Kendall rank-order correlation coefficient $T=-0.08$). Put another way, when averaged across the conflicts, 38% ($n=27$) of the forecasts of game theorists with fewer than five years experience were accurate compared to 29% ($n=74$) for those with more experience. Overall, game theorists had by their own reckoning as much as 30 years experience, with a mean of 11 years and median of 6.

To examine the effect of time spent on game-theorist forecast accuracy, I allocated forecasts into five groups such that each group included a similar number of forecasts. The upper limits of the groups were: 10, 20, 30, and 59 min. For each group, I calculated the error reduction versus the game-theorist average accuracy for matching conflicts. The correlation between the average self-reported time spent forecasting and the error reduction of the five groups' forecasts was negative (Kendall rank-order correlation coefficient $T=-0.40$). Moreover, 34% ($n=68$) of forecasts by game theorists who took less than 40 min were accurate when averaged across the conflicts, compared to 30% ($n=29$) for those who took longer.

4. Determining accuracy

Sometimes the set of potential decisions in a conflict situation is clear and the decision that is made is unambiguous. For example, in Zenith Investment (which was used in the original study) the choice facing the decision-makers was whether or not to invest in expensive new technology and, if so, whether to invest in one new plant or two. In other situations, such as Personal Grievance, this was not the case. Specifically, two, perhaps three, of the four decision options provided to participants could reasonably be interpreted as at least partly accurate representations of the actual outcome. Even when the actual decision maps perfectly onto one of the decision options provided to participants (as with Zenith) it is possible that forecasts that are not entirely accurate may nevertheless still be useful. For example, in the case of Zenith, given that two new plants were purchased, a forecast of "one new plant" would have been more useful to the plant-maker's planners than a forecast of none.

I recruited a convenience sample of five raters, and sent each of them a questionnaire by email. The questionnaire asked the raters to read descriptions of the actual decisions that were made in the eight conflicts used in this research. After reading the actual decisions, they rated for usefulness, on a zero-to-ten scale, the decision options that had been provided to research participants. The raters were told that a decision option that matched the actual decision should be given a rating of 10, a decision that was the opposite of the actual decision should be rated zero, and other

---

[3] Error reduction figures were calculated as $\{(100-a_c)-(100-a_x)\}/(100-a_c)*100$, where a_c is the percentage accuracy of the comparison forecasts (or chance) and a_x is the corresponding figure for the forecasts of interest.

decisions should be given some intermediate value. The instructions given to the raters are appended together with one rater's responses for one of the conflicts as an illustration (Appendix B).

The raters each rated the 31 decision options provided for the eight conflicts. The level of agreement between the raters was modest but unlikely to have arisen by chance (Kappa coefficient of agreement $K=0.249$, $z=7.89$, $P<0.001$; Siegel & Castellan, 1988). The coefficient K can take on any value between zero and one: zero where there is no more than chance agreement and one where there is complete agreement.

The median usefulness rating was used for each option. Using medians avoided the influence of outlier ratings (quirky interpretations by raters). The use of medians also had the desirable result that decision options rated 10 by a majority of raters received a usefulness rating of 10.

Interestingly, in the cases of Personal Grievance, Nurses Dispute, and Water Dispute, none of the decision options received a median rating of 10. That is, a majority of raters considered none of the options for these conflicts entirely matched the decisions that were actually made. In the case of Nurses Dispute, raters who did not give any of the options a rating of 10 argued that the option I considered a match with the actual decision was incomplete. Specifically, they wrote that the option was: "incomplete in that it does not give information about 'where the parties met'"; "literally correct, but 'wishy-washy' (vague)"; and "a compromise is a subtle thing". In the case of Water Dispute, raters argued that the option I considered matched the actual decision: "…implies that the amount will be sufficient to meet the needs—which is not stated in the actual decision"; "explains that Midistan will release water, but doesn't say how"; and "reflects part of what happened".

4.1. Usefulness assessment

In the Table 1 data, individual forecasts were either accurate (the option chosen was the one with the highest rating), or not.[4] To ascertain whether a comparison based on forecast usefulness would lead to a different conclusion about the relative worth of game-theorist and simulated-interaction forecasts, I compared average usefulness ratings for each of the conflicts.

Simulated-interaction forecasts were more useful than game-theorist forecasts for every conflict, with an average usefulness 6.6 versus 4.3 on the ten-point scale. The average expected usefulness of decisions chosen at random was 4.1.

Thus, simulated-interaction forecasts were superior to game-theorist forecasts whether one looks at usefulness or accuracy. The correlation between the two measures was modest (Kendall rank-order correlation coefficient $T=0.50$, $P=0.054$, $N=8$, one-tailed).

5. Combining forecasts

The error rates provided in earlier sections were based on individual forecasts. In practice, one should not depend on a single forecast, as the error rate can be high compared to that of a combined forecast (Armstrong, 2001b). For combining to be useful when predicting an outcome from a list of possible outcomes, however, individual forecasts must be more likely than chance to agree with the actual outcome. This was barely the case with game-theorists' forecasts and, as a consequence, the modal forecast was accurate for only three out of the eight conflicts (an error rate of 63%). In contrast, by combining simulated-interaction forecasts for each conflict, accurate forecasts were obtainable for seven of the conflicts (an error rate of 13%). This represents an error reduction of 67% from the 38% error for the accuracy of the individual simulated-interaction forecasts and an error reduction of 79% compared to combined game-theorist forecasts. Most importantly, these findings show that decisions in conflict situations can, in fact, be forecast accurately.

6. Further research

There is no evidence that the game theorists in my studies developed formal game-theoretic models and it is therefore plausible to assert that game theory, as opposed to game-theorist knowledge, has

---

[4] Distribution Channel was an exception, in that option C was "Either A or B". While there were nominally four options, option C was not counted in calculating the chance of choosing at random the option that coincided with the actual outcome. When C was chosen, the forecast was coded as half-accurate (0.5).

not been formally tested. While it is unclear why the game theory experts would not have developed models if doing so would have helped them to make accurate predictions, a study that required game theorists to provide models would resolve the issue.

Eight conflicts is a small sample when findings are contrary to many people's expectations, when the representativeness of the conflicts cannot be demonstrated, and when the type of conflict to which game theory is applicable is controversial. Perhaps the selection of conflicts that are not amenable to game-theoretic analysis was an unintended consequence of choosing situations about which information was readily available. Whatever the reason, it remains possible that there are types of conflict situations for which game-theoretic analysis would provide forecasts that were more accurate than simulated-interaction forecasts. I welcome extensions that would address this issue.

Participants in the research were asked to make a prediction by choosing from a list. This is a reasonable way to structure a forecasting problem and in many cases domain experts would be able to provide a list that was relevant to decision makers. The practicality of assembling such lists and how this should best be done merit further attention. For example, excepting 55% Pay Plan and Nurses Dispute, lists of decisions were compiled after the conflicts were over. More research using conflicts that are still unresolved at the time is warranted.

7. Conclusions

Despite the importance of accurately predicting people's decisions in conflicts, I was unable to identify research other than my own and Armstrong's (2001a) that compared the accuracy of forecasts from reasonable alternative approaches. Findings were initially obtained for only five situations that involved interaction and the findings were at odds with people's expectations. To test generalizability, I obtained new forecasts for three conflict situations that matched types of situations some game theorists have made predictions about. The new findings were consistent with the earlier ones. Specifically, knowledge of game theory was not useful, whereas modal forecasts from

the simulated interaction method were accurate for all but one conflict.

While further research is needed in order to identify whether there are situations for which game-theoretic analysis is more useful than other methods, current evidence suggests that decision makers would be wise to prefer forecasts from simulated interaction.

Appendix A. Coding refinements

In his analysis of my findings in Green (2002), Armstrong (2002) excluded the Panalba Drug Policy situation, arguing that as described it did not include interaction with other parties. I have followed that policy here, as my purpose is to compare methods for predicting decisions in conflicts that involve interaction between parties.

In Green (2002), I included an ambiguous game-theorist forecast for Artists Protest. I have excluded that forecast from analysis presented here. The change makes no difference to findings.

Distribution Channel option C is "either A or B". In Green (2002), I coded C as accurate. In subsequent research, I have coded C as half-right (0.5) and have adopted that practice for the analysis presented here. The effect of the change on previous findings is that the game-theorist accuracy for the Distribution Channel conflict is 23%, rather than 31%.

Appendix B. Usefulness questionnaire with example response for one conflict

B.1. Rating forecast usefulness

This rating task is part of a larger project to investigate the accuracy of forecasts from different methods for predicting decisions made by parties in conflict.

Following this note are sets of options that were judged by experts to be the decisions that might have been made in eight real conflicts. (Although they are all real, most of the conflicts are disguised.) Although every effort was made to ensure that the options were complete and mutually exclusive, in some cases the decision that was actually made does not match exactly any single decision option. Further, from a decision-maker's point of view, a forecast need not always be spot-on to be useful. For example, a forecast in August 2001 that attacks on the West's oil supplies by al Qaeda were imminent would not have been strictly accurate but could have been useful had governments and businesses responded by increasing security.

For each conflict, please read the brief description of the actual outcome then use your judgement to rate (on a scale of zero to 10) each of the decision options that were provided to forecasters. Using the September 11 example, a forecast that al Qaeda would never attack targets in the US should be given a score of *zero*, whereas a forecast that, within 1 month, they would use passenger planes to attack targets in New York and Washington should be given a score of *10*. The "attack on oil supplies" forecast described in the previous paragraph should be given a usefulness score *between* zero and 10. Note that your ratings of the decision options for any one conflict don't have to add to 10, or to any other number. In other words: rate each option independently.

You have about 3 min for each conflict.
55% Pay Plan
Forecasters were asked: Will there be a strike?
Actual decision: NFL players went on strike for most of the regular season.

Please rate how useful each of these decision options would have been as a forecast

---

A response from a research participant

	(Usefulness: 0–10)
(A) Yes, a long strike; (½ or more of the regular season games will be missed)	[10]
(B) Yes, a medium length strike; (less than ½ of the regular season games will be affected)	[8]
(C) Yes, a short strike; (only preseason games missed)	[4]
(D) No strike will occur	[0]

Briefly, what are your reasons for your ratings?
```
A and B are clear indicators of what
actually happened. C would allow for
some planning. D is not useful.
```

References

Armstrong, J. S. (2001a). Role playing: A method to forecast decisions. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 15–30). Norwell, MA: Kluwer Academic Publishers.

Armstrong, J. S. (2001b). Combining forecasts. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 417–439). Norwell, MA: Kluwer Academic Publishers.

Armstrong, J. S. (2002). Assessing game theory, role playing, and unaided judgement. *International Journal of Forecasting*, *18*, 345–352.

Brams, S. J., & Togman, J. M. (2000). Agreement through threats: The Northern Ireland case. In N. Miroslav, & J. Lepgold (Eds.), *Being useful: Policy relevance and international relations theory* (pp. 325–342). Ann Arbor, MI: University of Michigan Press.

Camerer, C. F. (2003). Behavioral game theory: Plausible formal models that predict accurately. *Behavioral and Brain Sciences*, *26*, 157–158.

Dixit, A. K., & Skeath, S. (2004). *Games of strategy* (2nd ed.). New York: Norton.

Erev, I., Roth, A. E., Slonim, R. L., & Barron, G. (2002). Predictive value and the usefulness of game theoretic models. *International Journal of Forecasting*, *18*, 359–368.

Fraser, N. M., & Hipel, K. W. (1984). *Conflict analysis: Models and resolutions*. New York: North-Holland.

Ghemawat, P., & McGahan, A. M. (1998). Order backlogs and strategic pricing: The case of the US large turbine generator industry. *Strategic Management Journal*, *19*(3), 255–268.

Goodwin, P. (2002). Forecasting games: Can game theory win. *International Journal of Forecasting*, *18*, 369–374.

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.

Green, K. C., Armstrong, J. S. (2004). *Value of expertise for forecasting decisions in conflicts*. Monash University Working Paper 27/04.

Gruca, T. S., Kumar, K. R., & Sudharshan, D. (1992). An equilibrium-analysis of defensive response to entry using a coupled response function model. *Marketing Science*, *11*(4), 348–358.

Haddad, C. (2001). *The telecom small fry that ate the boonies*. BusinessWeek Online, September 3. Retrieved May 3, 2002, from http://www.businessweek.com/magazine/content/01_36/b3747057.htm

Keesing's Contemporary Archives (1975, August 18–24). *Iraq: Syria–Iraq dispute* (pp. 27284–27285).

Keser, C., & Gardner, R. (1999). Strategic behavior of experienced subjects in a common pool resource game. *International Journal of Game Theory*, *28*(2), 241–252.

Kharif, O. (2001). *A small town vs. a very big deal.* BusinessWeek Online, August 21. Retrieved May 3, 2002, from http://www.businessweek.com/bwdaily/dnflash/aug2001/nf20010821_745.htm

Kliot, N. (1994). *Water resources and conflict in the Middle East*. London: Routledge.

London, S. (2002, March 26). Games or serious business. *Financial Times*, 16.

Organski, A. F. K. (2000). The outcome of the negotiations over the status of Jerusalem: A forecast. In N. Miroslav, & J. Lepgold (Eds.), *Being useful: Policy relevance and international relations theory* (pp. 343–359). Ann Arbor, MI: University of Michigan Press.

Siegel, S., & Castellan, N. J. (1988). *Non-parametric statistics for the behavioral sciences*. (2nd ed.). Singapore: McGraw-Hill.

Wrolstad, J. (2002). Alltel pays $1.65b for CenturyTel Wireless. *Wireless NewsFactor*, March 19. Retrieved May 3, 2002, from http://www.wirelessnewsfactor.com/perl/story/16833.html

In recent years, **Kesten Green** has been researching the problem of how best to predict the decisions people will make in conflict situations. He has presented papers on his work at International Symposia on Forecasting, RAND, Warwick Business School, the IMA Conference on Conflict and its Resolution, Melbourne Business School, and to the London Judgement and Decision Making Group. Dr Green created conflictforecasting.com. Prior to his academic career, he spent more than twenty years in business as a founder of four companies.