Decisions and Delphi: The Dynamics of Group Estimation

Shari Lawrence Pfleeger
Systems/Software, Inc.
Washington, DC, USA

Martin Shepperd
Bournemouth University
Poole, Dorset, UK

Roseanne Tesoriero
The Catholic University of America
Washington, DC, USA

Note: A version of this paper has been submitted to *IEEE Software*.

Abstract

Cost estimation is often performed by a group. We examined what decision theorists have to say about group decision-making, and observed seven groups as they discussed their estimates for a small project. We found that personalities had more to do with confidence and consensus on an estimate than did expertise or years of experience.

Keywords

estimation, decision theory, bias

Lots of decisions

Sometimes it seems as if software engineering is simply a string of pressured activities connected by decision-making and estimating. We plan our projects by estimating resources and risk. We assess projects by deciding if our processes were effective, our resources appropriate, and our products satisfactory. To test our products, we weigh alternatives when we cannot test everything. And change requests and maintenance require evaluating alternative actions, estimating needed resources, and analyzing risk.

We need not make our decisions in a vacuum. There are theories that support our decision-making from two points of view: descriptive and prescriptive. Descriptive theories provide evidence about how decisions are actually made, while prescriptive theories provide frameworks and methods to help decision-makers improve their performance in problem-finding and problem-solving, given real constraints. Figure 1 illustrates how many other disciplines contribute information to our decision-making.

	Descriptive theories	*Prescriptive theories*
Individual	Psychology Marketing Psychiatry Literature	Decision theory Economics Operations research Philosophy and logic
Group	Social psychology Organizational behavior Anthropology Sociology	Game theory Organizational behavior Clinical psychology and therapy Finance and economics
Organization	Organization theory Sociology Industrial organizations Political science	Planning and strategy Control theory and cybernetics Organizational design Team theory and economics
Society	Sociology Anthropology Macroeconomics	Legal philosophy Political science Social choice

Figure 1: Roots of decision sciences (Kleindorfer *et al.* 1993)

Often, our decision-making involves at least two distinct steps. First, we make our choices individually. We predict or infer, we assign value to the different alternatives, and we assess different approaches as we make up our minds. Second, we contribute our findings to a group decision process. For example, to estimate the effort required for building a particular kind of software, each of us may make our own prediction before combining estimates to get an idea of what the group predicts. Moreover, this "group" may in fact be our projects, our organizations, or even our societies; each such decision has impact relative to the group it will reflect or affect.

Figure 2 shows us that many elements affect how we make up our minds. The context of the situation constrains both our understanding and our options. Within that context, we must understand and represent the problem before we try to solve it. Each option must be screened in several ways, to determine its likely effect on stakeholders, and its degree of rationality and realism. "Legitimation" is likely to be both highly significant and somewhat overlooked in the context of software engineering. It is conceivable that estimators and decision makers may have a preference for estimates and decisions that can be more easily justified. This preference suggests a bias in favor of a particular approach, such as the use of algorithmic models over expert judgment. And the possible solutions must be viewed through a filter of our values and beliefs.

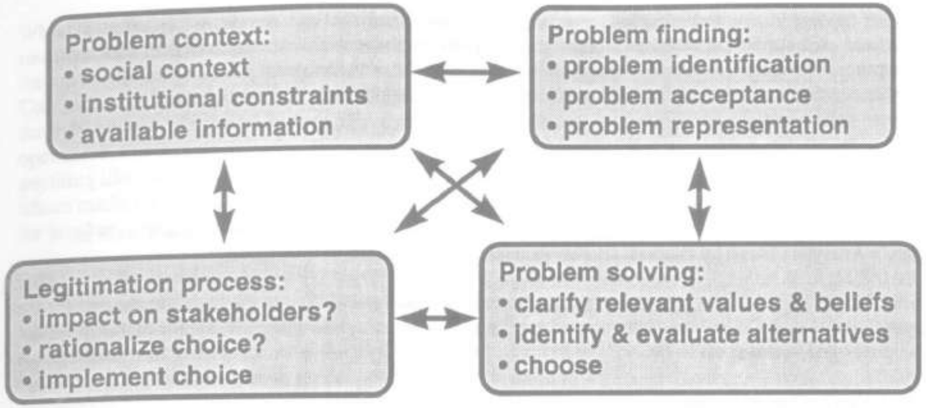

Figure 2: Aspects of decision-making (Kleindorfer *et al.* 1993)

To see how we use these elements in our decision-making, consider the problem of choosing new office space. Table 1 represents five options. Each alternative is characterized by the rent in dollars per month, the distance in kilometers from home, the size in square meters, and a general, subjective rating of the quality of space. (For instance, a "high" quality space may have more light or higher ceilings than a lower-quality one.)

Table 1: Office space options

Office option	Rent per month	Distance from home	Size	Quality
1	450	10	4000	Medium
2	475	15	2500	High
3	460	14	1500	Average
4	500	5	1750	High
5	510	7	2500	High

There are many rules we can use to select the best option. For example, we can choose the office with the lowest rent. Or we can choose the office closest to home. These rules reflect our values; someone who uses the first rule instead of the second may value money over time. Alternatively, we can use more complex rules. For instance, we can define "office value" to be a combination of rent and size, and we can balance it with the shortest travel time. Or, we can use a multi-step approach, where first we set cut-off levels for rent and distance (such as, no more than $500 for rent, and no more than 10 kilometers from home), and then balance the remaining attributes.

Of course, our selection process can be still more sophisticated. For example, we can use domination procedures, where we eliminate alternatives that are "dominated" by better choices. However, this type of rule can lead to suboptimization; we may eliminate a pretty good choice if our cut-offs are arbitrary or not carefully considered. Or we can use conjunction (every dimension meets a defined standard) or disjunction (every dimension is sufficiently high). In these situations, there is no slack when the characteristic values are close to the threshold; we may discard a choice because the rent is over $500, but in fact the other characteristics of the $501 choice are far superior to those in other options.

Another strategy is to use elimination by aspects. Here, each attribute has a preassigned criterion value, and each attribute is assigned a priority. Attributes are then reviewed in terms of their relative importance. We can formalize this approach by using an additive value model, where we assign weights or priorities (w_j) to each attribute (x_i), and then sum the products of the weights and the attribute values ($v(x_{ij})$):

$$V_i = \underset{j}{} w_j v(x_{ij})$$

Sometimes weights and comparisons are more easily made by adopting a pairwise approach, such as Saaty's Analytic Hierarchy Process, another example of a multi-criteria decision aid.

Each of these approaches suggests the "right" choice, but it may not always be the optimum choice. Or it may involve many calculations or comparisons. In reality, we may use a heuristic approach that gives us a pretty good answer.

The more the merrier

So far, we have discussed characteristics related to the problem itself. Group decision-making is in some sense more difficult because aspects of group behavior influence how the decisions are made. Figure 3 illustrates some of the issues that must be considered when several people try to choose among alternatives. For example, trust, communication and cooperation can affect the result; none of these is a factor in individual choice.

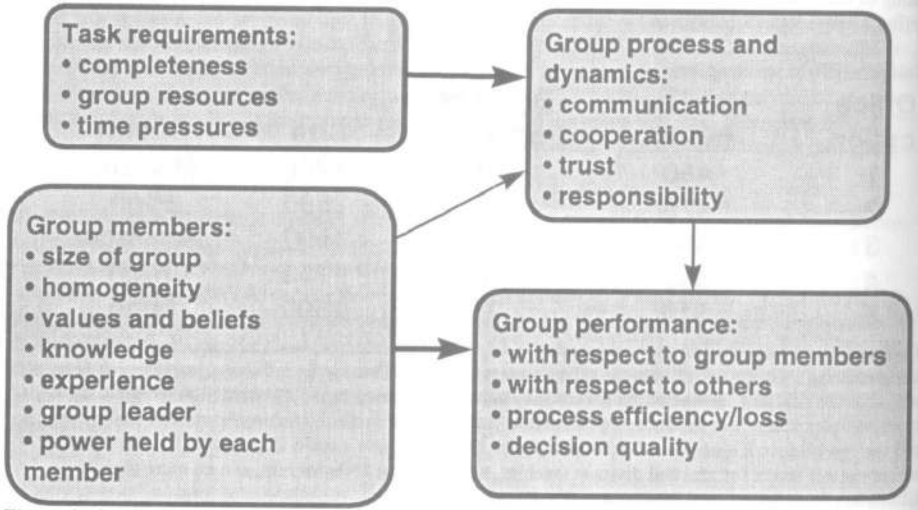

Figure 3: Issues in group decision-making

However, there are several group decision strategies to address these concerns. For example, dialectical strategies may allow one side to advance an argument, then the other side to speak. A third party may be employed to reconcile the differing viewpoints. Alternatively, brainstorming can be used to identify a full list of possibilities, including opportunities or threats. Nominal group techniques involve silent generation of ideas followed by a round robin, where ideas are shared one at a time and then evaluated spontaneously using silent voting. Or, social judgment approaches may be used to separate facts from judgments, or to distinguish science from values from social judgment.

When the group is an organization, the decision-makers must distinguish strategic from tactical and routine decisions. Strategic decisions affect the well-being and nature of the organization; typically, they involve new products, services or markets, and senior management may play a significant role. Cost estimates can be part of strategic decisions, especially when they are used to position a product in the marketplace. Tactical decisions affect pricing, employee assignments, customer interaction or operations, but they do not affect the organization's financial bottom line or commercial direction to anything like the same degree. A tactical cost estimate can be used to set the price for a new product where market share may not be an issue; for example, when one company division develops a product for another division, competition and pricing may not be of strategic importance.

Routine decision-making is usually more mundane: repetitive in nature, local in scope, and guided by organizational rules or policies. For instance, suppose a company supports its own reuse repository, where software engineers are rewarded for "depositing" a reusable component and "charged" for using a component that already exists in the repository. Determining the "price" of the component may be a routine task based on corporate guidelines.

How we really decide

The decision science and operations research literature is replete with examples of decision-making techniques. But which of those techniques do we really use when we make decisions? A survey reported by Forgionne (1986) indicates that we tend to use statistics and simulation, but very few of the more complex processes suggested in the textbooks.

There are many reasons why we often shun the more technically-sophisticated approaches. The biggest impediments to their use are the difficulty of setting up the calculations and the combinatorial explosion of possibilities. Rather than hypothesize about the best way to make decisions, Gary Klein (1998) has observed decision-makers at work, under pressure. In one study of 156 observations, he found that no one made use of preselected options (where someone else lays out what you may do, and then you choose from among those possibilities). Eighteen decision-makers did a comparative evaluation, where an initial option was chosen, and then all other options were compared to it to see if it was the best one; here, the decision-makers were optimizing their action. Eleven decision-makers created a new option. But the rest used what Herbert Simon calls a satisficing strategy: they evaluated each option on its own merits until they found the first one that met their criteria for an acceptable solution.

After watching and interviewing firefighters, emergency medical technicians, soldiers, and others who make important decisions under pressure, Klein has suggested a "recognition-primed decision model," as shown in Figure 4, to describe how we really make decisions. He points out that we tend to keep a repository of examples in our heads, and we draw on these situations when we are confronted with a choice. We mentally refer to a past situation and compare the current one to see if it is similar. Rather than comparing all the options, we grab one that is "close," and we go through a process in our heads to justify to ourselves that it is "close enough" to be used as a basis for the current decision. At the same time, we use mental simulation to determine if the actions we propose to take will really work in this situation; when we don't think they will, then we back up, choose a different scenario, and perform the mental simulation again. Eventually, when we are satisfied that we have made the right choice, we act.

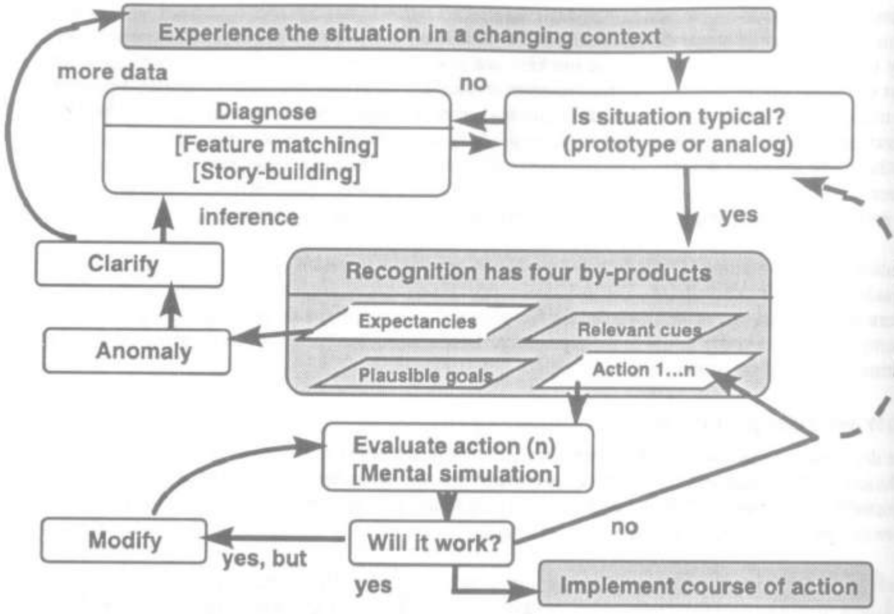

Figure 4: Recognition-primed decision model (Klein 1998)

But decision-making and estimating are not as simple as the model suggests. Hershey, Kunreuther and Schoemaker (1982) have demonstrated that the decision's context can bias the choice. To see how, consider these two questions:

> *Question 1*: You have a 50% chance of losing $200 and a 50% chance of losing nothing. Would you be willing to pay $100 to avoid this situation?
>
> *Question 2*: You can pay $100 to avoid a situation where you may lose $200 or nothing. Would you pay if there were a 50% chance of losing?

In their study, they asked each of two equivalent groups to answer the question with either "yes," "no" or "indifferent." Even though the two questions are equivalent, 6% of the group answering question 1 were willing to pay, but 32% of the group answering question 2 was willing to pay.

Tversky and Kahneman (1981) illustrated the same kind of bias in risk-analysis decision-making. They described a situation where a rare disease endangers 600 people in a village. Public health officials have enough vaccine for 200 people, and they consider two different possibilities. They can either administer the full vaccine to 200 people (program A), or they can water down the vaccine and hope that it will protect all 600 villagers (program B). The researchers asked a group to choose between the two programs, where

> *Program A*: Exactly 200 lives will be saved.
> *Program B*: 1/3 chance of saving all 600, and 2/3 chance of saving none.

In an alternative framing, the researchers asked an equivalent group to make the same choice, but this time framed the programs in terms of lives lost:

> *Program C*: Exactly 400 lives will be lost.
> *Program D*: 1/3 chance that no one will die, and 2/3 chance that 600 will die.

Even though the problems are mathematically identical, there was a dramatic difference in the responses. 72% of the subjects shown the first framing chose A; only 22% of the subjects shown the second framing chose C.

In a similar study, people linked past actions to current decisions. To see how, consider the situation where you are going to the theater to see a play. When you get to the box office, you find that you have lost a $10 bill. Would you then pay $10 for the ticket to the play? Eighty-eight percent of the respondents said yes. But if, alternatively, when you get to the box office, you find that you have lost your $10 ticket, would you pay $10 for a new ticket? In this situation, only 46% would pay, even though the situations are mathematically identical.

Thus, it is important to consider contextual bias, especially when estimating effort or risk. The way in which the problem is framed can produce radically different results.

Claude Steele's work (1998) has described a related phenomenon. He notes that expectation of performance or results can lead to a "stereotype threat." For example, if a particular group of people is told that it usually does not do well in a given situation, then the stereotype can actually result in poorer performance.

Other sources of bias can creep into decision-making. For example, people usually overvalue what they already own (called "status quo bias"). This phenomenon can influence an estimator to be overly optimistic about productivity in a familiar development situation. Similarly, probability and payoff interact: if probabilities are small, people look at payoff first; if probabilities are large, people look at probability and then payoff.

Individuals exhibit a marked preference for case-specific, or singular, information, as opposed to general statistical, or distributional, information. Busby and Barton (1996) focus on this preference when describing estimators who employed a top-down or work breakdown approach to prediction. Unfortunately, this approach failed to accommodate unplanned activity, so that predictions were consistently under-estimating by 20%. By definition, the case-specific evidence for each project will fail to account for unplanned activities, yet the statistical evidence across many projects suggests that unplanned activities are very likely to happen. Nevertheless, managers favored the singular evidence and would not include a factor for unplanned activities in their estimation process.

In addition, we must remember that recall is affected by both the recency and vividness of an experience. The further into the past a factor occurred, the greater the tendency of the recaller to discount its significance. In one sense, this diminishing significance may be sensible, given that the way in which we develop software has changed considerably over the years. On the other hand, many risks, such as requirements being modified or misunderstood, have changed little.

Anchoring-and-adjustment is another common technique employed by estimators. Here the estimator selects an analogous situation and then adjusts it to suit the new circumstances. However, there is considerable evidence to suggest that estimators are unduly cautious when making the adjustment. In other words, the anchoring dominates and then insufficient adaptation is made. This approach may be influenced by recall, in that the most suitable analogies may be overlooked because they are not recent.

A reluctance to appear negative can also affect expert judgment. Tom DeMarco (1982) reminds us that "realism can be mistaken for disloyalty," leading to undue optimism in making predictions, both individually and in groups.

How groups really make decisions

We often evaluate our estimation accuracy by comparing estimates to actuals for individual estimators. But in fact, many organizations use group decision-making or estimating techniques to derive important estimates. For example, the Delphi technique (see Sidebar) enables several estimators to combine their disparate estimates into one with which all can feel comfortable.

But the group dynamics can affect the quality of a decision or estimate. For instance, Foushee *et al.* (1986) found that it takes time for team members to learn to be productive together. The teams performed better at the end of their assignment than at the beginning, because they learned over time to work together effectively.

Group dynamics can also have negative effects. Solomon Asch tested the effect of colleagues on an individual's decision by presenting a subject with the lines shown in Figure 5. When individuals were asked which of the three lines on the right was the same length as the test line, almost all of the respondents answered correctly: line B. But when others in the room were present and gave the wrong answer, the number of errors rose dramatically, as shown in Table 2.

Figure 5: Example (Asch 1956)

Table 2: Example (Asch 1956)

Condition	*Error rate*
Subject is alone	1%
With 1 person who says A	3%
With 2 people who say A	13%
With 3 people who say A	33%
With 6 who say A and 1 who says B	6%

A modest observational study

In 1999, we explored the effects of group decision-making on effort estimation. We began with a pilot investigation at Bournemouth University. Twelve postgraduate students were organized into four teams having between two and four members. As part of coursework, each team was required to capture requirements and develop a prototype for a simple information system. We then asked them to predict the size of the prototype in lines of code. (To minimize counting problems, we defined a line of code as a statement delimiter. We chose lines of code because they could easily be verified. not because of any other special significance.) The subjects were not constrained to any particular technique, although in practice they tended to use subjective judgment. Immediately afterwards, the subjects participated in Delphi meetings leading to two additional estimates.

Table 3: Estimation Errors from Three Rounds of Predicting Size

Estimate	Median Error	Min Error	Max Error
Initial	160.5	23	2249
Round 1	40	23	749
Round 2	40	3	949

Table 3 shows that both the median error and the spread or range of errors are greatest for the initial estimate. In other words, error was greatest prior to using the Delphi technique; the subsequent Delphi rounds led to a reduction in the differences between predicted and actual prototype sizes. As shown in Figures 6 and 7, three out of the four groups exhibited clear improvement, but the fourth group diverged from the true value as the process continued. This divergence was due in part to the dominance of one member of the group. Although the Delphi technique allows anonymity for individual estimates, it appears to be vulnerable to forceful individuals. Interestingly, this result is consistent with the behavioral theories described above. As we have seen, the group performance literature (such as Sauer *et al.* 2000) suggests that a major determinant of the outcome is the choice of decision scheme, as well as how the group makes use of its expertise.

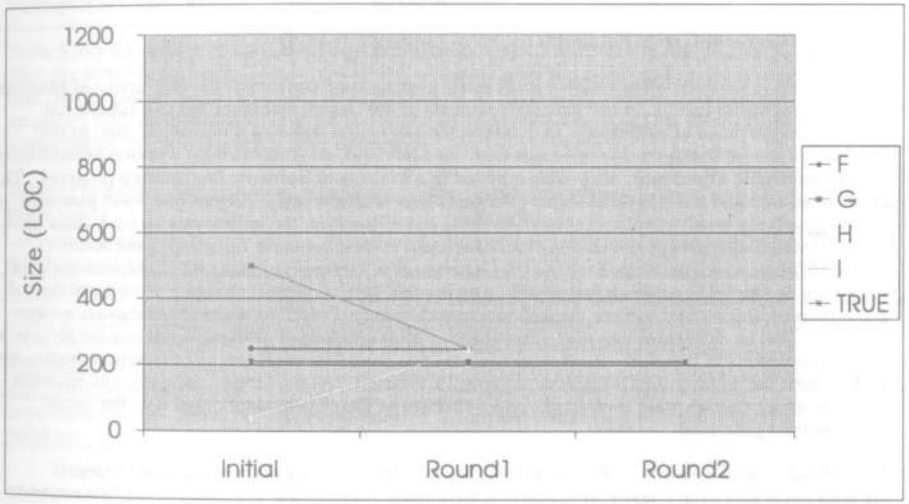

Figure 6: Convergent Group Estimates

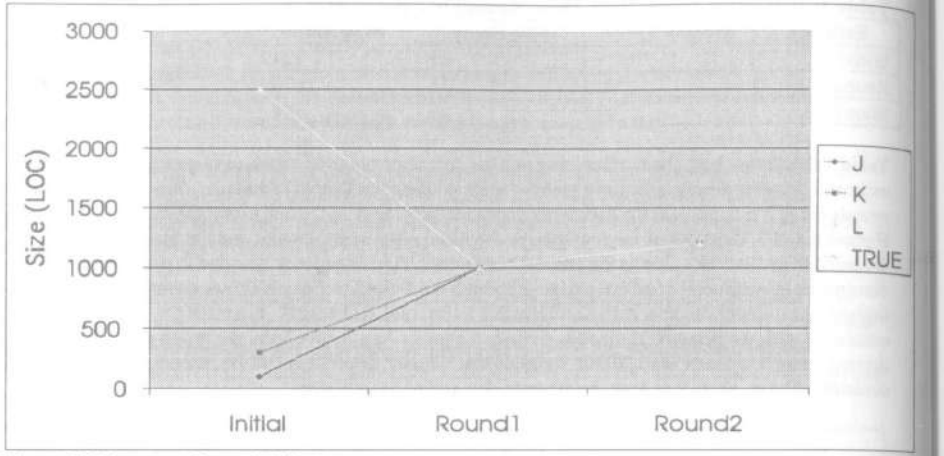

Figure 7: Divergent Group Estimates

Encouragingly, a similar observational study using postgraduate students at the University of Maryland yielded comparable results, in that successive rounds of the Delphi technique led to a substantial reduction in the range of estimates. In this case, the estimation task was a theoretical one, so the accuracy of the estimates could not be assessed. At Maryland, all students were working practitioners with considerable experience; they were enrolled in a Masters of Software Engineering program. The class of ten was divided into three teams of three or four students each. As part of a more general project involving requirements elicitation, analysis and estimation, the individuals on each team used two products, Data Salvage (an analogy tool developed at Bournemouth University) and Revic (a COCOMO-based tool developed for the US Department of Defense) to generate initial estimates for their team project. Then, after a lesson in how to use the Delphi method, the students were observed and tape-recorded as they worked through two twenty-minute Delphi rounds to converge on a team estimate. That is, each team was given the results of all the individual estimates, not just for their team but from everyone in the class. Each team was asked to record its confidence in the new estimate, and to document the assumptions it made in deriving the estimate. At the time of the study, the students were not aware that we were interested in the dynamics of their discussion, rather than the actual numbers they generated.

We noticed several important trends across the teams: First, the spread from smallest to largest estimate decreased dramatically over time. It went from 16239 at the beginning of the first round to 10181 at the beginning of the second, to 1520 at the end.

Second, there was a general growth in confidence as the students progressed through the Delphi discussions. That is, all students reported an increase in confidence with their estimate after the group discussions, regardless of their experience level. However, there was no evidence of any relationship between experience and confidence.

Because personality can play such a significant role in the Delphi discussions, we administered a Myers-Briggs test to each of the Maryland students. Myers-Briggs classifies each respondent on four scales: extroverted/introverted, sensing/intuition, thinking/feeling, and judging/perceiving. Thus, the results are reported as a four-letter combination, such as ISTJ, for introvert-sensing-thinking-judging, to describe the way in which the student typically responds to situations. The results for the second group were particularly interesting. Two members of this group were classified as ISTJ, a third was an ESTJ, and the fourth was an ISFJ. People with the ISTJ personality type tend to be precise and accurate, tending to follow rules and procedures and having excellent concentration. However, ISTJs do not like change and tend to be inflexible. ISFJs, like ISTJs, tend to be accurate, thorough and careful with details, as well as resistant to change. Interestingly, ISFJ personalities tend to underestimate their own value. By contrast, ESTJs tend to be practical and results-oriented. They focus on a goal, becoming impatient with inefficiency and with those that do not follow procedures.

Group 2's Delphi discussions focused on the details of the parameters of the Revic tool, rather than on more global issues about the project in general. Perhaps the detail-oriented personality types of these group members led them to look more at the tool rather than at their own experiences. As can be seen from Table 4, Group 2 was the least confident in its final estimate, perhaps because of its team member personality types.

Table 4: Maryland Group's Confidence in Final Estimate

Group	Number in Group	Median Confidence in Estimate	Range of Estimates
1	2	91	2
2	4	65	15
3	3	80	0
4	3	80	0

From a debriefing questionnaire that we administered after the study, it appears that the ten subjects had a favorable attitude towards Delphi, with all subjects reporting that the technique had increased their confidence in their estimate. When asked if each would consider using Delphi estimation techniques in a work situation, five said yes, four said maybe, and only one said no.

Table 5 lists the positive and negative issues identified by subjects in the debriefing questionnaire. They are sorted in decreasing order of frequency, so that the most commonly-mentioned issues appear first. Notice that problems related to a dominant individual and lack of expertise are the most frequently cited disadvantages, while the benefits of different perspectives and the value of group discussion are the most frequently cited advantages.

Table 5: Perceived Strengths and Weaknesses of the Delphi Technique

Weaknesses	Strengths
can wrongly influence an individual and the impact of a dominant individual	experts with different backgrounds/perspectives
depends upon knowledge/expertise of individuals	group discussion can correct mistakes
risk of erroneous assumptions	reconsideration
group discussion made little difference to the result (consensus group)	uses expert judgment
high variability in predictions	median better than mean
inappropriate target, should use for more detailed problems	provides comparison with other estimates
	anonymity/independence combined with group benefits

Lessons learned

A number of lessons emerge from the two studies of Delphi estimation techniques. The first is that the subjects had a broadly positive attitude to the technique. Moreover, from the Bournemouth study, it is clear that the technique led to improved estimates. And in both studies, the technique clearly reduced the spread of estimates, even though six of the ten Maryland students reported negative effects of the group discussions. As researchers such as Klein point out, there are other, indirect benefits, including education and the development of a common vocabulary.

A second clear lesson is that personalities can dominate the discussion, even when the dominant participant is not correct, much as Asch found several decades ago. Moreover, the individual assumptions that formed the basis of initial estimates (that is, the factors required as input by the tools used) were irrelevant in most of the subsequent group discussion. This result parallels findings reported in investigations of group meetings for reviews and inspections, where many individual findings are lost when not confirmed by others in the group. In particular, the focus of the Delphi discussions turns to the numbers themselves (rather than where those numbers come from), and to justifying gut feel, as Klein suggests. In other words, anchoring-and-adjustment is alive and well in the Delphi technique.

We often assume that those with the most experience will provide the strongest influence on group discussions, leading to more realistic estimates. But in our studies, even the most experienced group (Maryland's Group 3) looked to the median for reassurance. And all students reported an increase in confidence that had no relationship to the experience of the team members. Thus, personality may dominate experience in Delphi discussions, a situation that does not often lead to realistic or accurate final estimates.

Most of the existing research on estimation techniques has focused on the accuracy of individual estimates. However, most practitioners do their estimation in a group, either explicitly by relying on a technique such as Delphi, or implicitly by eliciting the opinions of their colleagues. For this reason, it is important that we acknowledge the role of group dynamics in the estimation process. We often assume that expertise and experience will dominate, that we have high-quality data in historical records of similar projects from which we can draw our analogies, and that we know how to tell when two projects are similar to one another. Unfortunately, many of these assumptions are wrong to some extent, and practitioners must rely on individual and group tools and techniques from which to generate an estimate. Although we claim that objectivity is essential to estimation, in fact we are forced to rely on subjective and often fuzzy data and techniques to make our estimation decisions. Observational studies such as ours, combined with a solid understanding of group dynamics, can help us to tailor our estimation processes to the realities of personality and process, and thus increase our confidence in the results.

References

- Solomon Asch, "Studies of independence and submission to group pressure," *Psychological Monographs*, 70, 1956.
- Barry W. Boehm, *Software Engineering Economics*, Prentice-Hall: Englewood Cliffs, N.J., 1981
- J. S. Busby and S. C. Barton, "Predicting the cost of engineering: does intuition help or hinder?," *Engineering Management Journal*, 1996, p. 177-182.
- Tom DeMarco, *Controlling Software Projects: Management, Measurement and Estimation.* Yourdon Press, New York, 1982.
- G. A. Forgionne, *Quantitative Decision Making*, Wadsworth, Belmont, California, 1986.
- H. C. Foushee, J. K. Lauber, M. M. Baetge and D. B. Acomb, "Crew performance as a function of exposure to high-density short-haul duty cycles," NASA Technical Memorandum 99322, Moffett Field, CA, NASA Ames Research Center, 1986.
- John C. Hershey, Howard C. Kunreuther and Paul Schoemaker, "Sources of bias in assessment procedures for utility functions," *Management Science*, 28, 1982, p. 936-954.
- Gary Klein, *Sources of Power*, MIT Press, 1998.
- Paul Kleindorfer, Howard Kunreuther and Paul Schoemaker, *Decision sciences: an integrative perspective*, Cambridge University Press, UK, 1993.
- James L. McKenney and Peter G. W. Keen, "How managers' minds work," *Harvard Business Review*, 52(2), May/June 1974.
- Sauer, Chris, *et al.*, 'The effectiveness of software development technical reviews: a behaviorally motivated program of research', *IEEE Transactions on Software Engineering* 26(1), pp1-14, 2000.
- Claude Steele, "Thin ice: stereotype threat and black college students," *Atlantic Monthly*, August 1999.
- Amos Tversky and Daniel Kahneman, "The framing of decisions and the psychology of choice," *Science*, 211, 1981, p. 453-458.

Sidebar: Delphi Techniques

The Delphi techniques was originally devised by Rand Corporation in the late 1940s as a method for structuring group communication processes in order to solve complex problems. For example, one application was to predict the future of the oil industry for the US Government. Delphi is intended to obtain informed judgment by keeping individual predictions anonymous and by iterating through several evaluations. Consensus is not essential. In fact, Delphi can also be used to educate its participants by drawing on multi-disciplinary or diverse inputs.

The technique was subsequently refined and popularized by Barry Boehm (1981) for the more specific task of software project estimation. The major steps in the Delphi process include:

1. A group of experts receives the specification plus an estimation form.
2. The experts discuss product and estimation issues.
3. The experts produce individual estimates.
4. The estimates are tabulated and returned to the experts.
5. An expert is made aware only of his or her own estimate; the sources of the remaining estimates remain anonymous.
6. The experts meet to discuss the results.
7. The estimates are revised.
8. The experts cycle through steps 1 to 7 until an acceptable degree of convergence is obtained.

While the technique is relatively well-known and is featured in many software engineering textbooks, there are few published experiences of using Delphi for software prediction.

Sources on Delphi Techniques:

H. Linstone and M. Turoff, *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, Massachusetts, 1975.

M. Turoff and S. R. Hiltz, "Computer Based Delphi Processes," in *Gazing Into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health*, M. Adler and E. Ziglio, Eds., Kingsley Publishers, London, 1995.