האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

THE BENEFIT OF ADDITIONAL OPINIONS

by

ILAN YANIV

Discussion Paper # 422 May 2006

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

The Benefit of Additional Opinions

Ilan Yaniv

Hebrew University of Jerusalem, Jerusalem, Israel

ABSTRACT

In daily decision making, people often solicit one another's opinions in the hope of improving their own judgment. According to both theory and empirical results, integrating even a few opinions is beneficial, with the accuracy gains diminishing as the bias of the judges or the correlation between their opinions increases. Decision makers using intuitive policies for integrating others' opinions rely on a variety of accuracy cues in weighting the opinions they receive. They tend to discount dissenters and to give greater weight to their own opinion than to other people's opinions.

Paper appeared in: *Current Directions in Psychological Science*, *13,* 75-78 (2004).

This research was supported by Grant No. 822/00 from the Israel Science Foundation. Address correspondence to Ilan Yaniv, Department of Psychology, Hebrew University, Jerusalem 91905, Israel; e-mail: ilan.yaniv@huji.ac.il.

It is common practice to solicit other people's opinions prior to making a decision. An editor solicits two or three qualified reviewers for their opinions on a manuscript; a patient seeks a second opinion regarding a medical condition; a manager considers several judgmental forecasts of the market before embarking on a new venture. All these situations involve the decision maker in the task of combining other people's opinions, mostly so as to improve the final decision.

People also seek advice when they feel strongly accountable for their decisions. An accountant performing a complex audit might solicit advice to help justify his or her decisions and share the responsibility for the outcome with others. One could justifiably argue, however, that even such reasons for seeking others' opinions are rooted in the belief that this process could improve decision making.

Two main questions arise in the research on combining opinions. One involves the statistical aspects of the combination task: Under what conditions does combining opinions improve decision quality? The other concerns the psychological process of combining judgments: How do judges utilize other people's opinions? These questions, which have been investigated by students of judgment and decision making, statistics, economics, and management, are intertwined, because the quality of the product is related to the way it is produced. In this review, I discuss what researchers have learned about the process and outcomes of combining opinions.

Our focus here is on situations in which a decision maker seeks quantitative estimates, judgments, and forecasts from people possessing the relevant knowledge. The opinions are then combined by the individual decision maker, not by a group (decision making in groups deserves a separate discussion, e.g., Hill, 1982). It is useful to distinguish between two ways in which expert judgments can be combined: (a) intuitively (subjectively) and (b) mechanically (formally), that is, by using a consistent formula, such as simple or weighted averaging.[1]

ACCURACY GAINS FROM AGGREGATION

Research has demonstrated repeatedly that both mechanical and intuitive methods of combining opinions improve accuracy. For example, in a study of inflation forecasts, the aggregate judgment created by averaging the forecasts of expert economists was more accurate than most of these individual forecasts , though not as good as the best ones (Zarnowitz, 1984). The best forecasts, however, could not be identified before the true value became known. Hence, taking the average was superior to selecting the judgment of any of the individuals (Armstrong,

2001; Clemen, 1989). Moreover, a small number of opinions (e.g., three to six) is typically sufficient to realize most of the accuracy gains obtainable by aggregation. These fundamental results have been demonstrated in diverse domains, ranging from perception (line lengths) and general-knowledge tasks (historical dates) to business and economics (sales or inflation forecasts), and are an important reason for the broad interest in research on combining estimates (Johnson, Budescu, & Wallsten, 2001; Sorkin, Hayes, & West, 2001; Yaniv & Kleinberger, 2000).

How Does Combining Opinions Improve Judgment?

The improvement in accuracy is grounded in statistical principles, as well as psychological facts. For quantitative estimates, a common measure of accuracy is the average distance of the prediction from the event predicted. In the special case of judgments made on an arbitrary rating scale (e.g., an interviewer's rating of a job candidate's capability on a 9-point scale), a common measure is the correlation between the judgments and some objective outcome (the candidate's actual success).

In the case of quantitative estimates, it can be outlined in simple terms why improvement is to be expected from combining estimates. A subjective estimate about an objective event can be viewed as the sum of three components: the “truth,” random error (random fluctuations in a judge’s performance), and constant bias (a consistent tendency to over- or underestimate the event). Statistical principles guarantee that judgments formed by averaging several sources have lower random error than the individual sources on which the averages are based. Therefore, if the bias is small or zero, the average judgment is expected to converge about the truth (Einhorn, Hogarth, & Klempner, 1977).

The case of categorical, binary judgments (e.g., a physician inspects a picture of a tumor and estimates whether it is benign or malignant) requires a special mention. Suppose a decision maker polls the judgments of N independent expert judges whose individual accuracy levels (chances of choosing the correct answer) are greater than 50% and then decides according to the majority. For example, three experts might judge whether or not a witness is lying, and the final decision would be the opinion supported by two or more experts. According to a well known 18th-century theorem (known as Condorcet's jury theorem), the accuracy of the majority increases rapidly toward 100% as N increases (e.g., Sorkin et al., 2001). Thus, the majority outperforms the individual judges. For instance, the majority choice of five independent experts who are each correct 65% of the time is expected to be correct approximately 76% of the time.

---

[1] More complex methods based on Bayes's theorem are less common in psychological research on combining opinions; hence, they are not treated here.

Conditions Under Which Accuracy Gains Are Observed

A central condition for obtaining optimal accuracy gains through aggregation is that the experts are independent (e.g., little gain is expected if judge B is essentially a replica of judge A). But gains of appreciable size can be observed even when there are low or moderate positive correlations between the judgments of the experts (Johnson et al., 2001). The gains from aggregating quantitative judgments are also determined by the bias and the random error of the estimates (the lower the better). If judgments are made on rating scales, then the accuracy gains are related directly to the validity of each judge (i.e., how the judge's ratings correlate with the objective value of what is rated) and indirectly to the correlations between different judges' ratings (Einhorn et al., 1977; Hogarth, 1978; Johnson et al., 2001).

Number of Opinions Needed

As already noted, as few as three to six judgments might suffice to achieve most of what can be gained from averaging a larger number of opinions. This puzzling result that adding opinions does not contribute much to accuracy is related to my previous comments. Some level of dependence among experts is present in almost any realistic situation (their opinions tend to have some degree of correlation for a variety of reasons--they may rely on similar information sources or have similar backgrounds, or simply consult one another; cf. Soll, 1999). Therefore, the benefits accrued from polling more experts diminish rapidly, with each additional one amounting to “more of the same.” Similarly, bias or low judge validity limits the potential accuracy gains and further diminishes the value of added opinions.

PSYCHOLOGICAL EFFECTS ON THE AGGREGATION OF OPINIONS

Consider generic scenarios involving intuitive methods of combining opinions: A moviegoer receives conflicting reviews about a movie, or an undergraduate student hears mixed evaluations from fellow students about an elective course. Although formal approaches deal with the conflict by assigning explicit weights to the various opinions, people often attempt to resolve the conflict by trying to form well-justified, coherent judgments, assessing the merit of each source and the arguments for or against each opinion and trying to explain away the differences. Specifically, several factors affect the weighting of opinions in intuitive decision making, including (a) cues for accuracy, (b) responses to dissension, and (c) self-versus-other effects.

Cues for Accuracy

A decision maker’s trust in a given opinion depends on his or her assessment of the accuracy of the source. How are expectations about this accuracy formed? How does trust develop? Studies suggest that a variety of cues serve as proxy measures of the actual accuracy of sources. These cues include expertise, confidence, and past performance.

First, people are sensitive to the expertise (or credibility) ascribed to various sources and assign weights to sources as a function of such attributions (Birnbaum & Stegner, 1979). Second, a frequent and immediate cue for accuracy is the judge's stated confidence about his or her opinion (Sniezek & Van Swol, 2001). Subjective statements such as "Trust me" or "I am 60% sure" are used as factors in weighting judgments. Such a policy is beneficial to the extent that confidence and accuracy are correlated (Yaniv, 1997). Finally, an expert's past performance serves as a cue to his or her accuracy. In studies in which the same experts give multiple opinions, participants form impressions about the accuracy of each expert and adjust their weights accordingly. Trust in experts is fragile, being "hard to gain, easy to lose," because negative experiences with a source have proportionally greater influence than positive ones (Yaniv & Kleinberger, 2000).

Ignoring Dissenters' Opinions

Certain configurations of opinions present particularly sharp dilemmas as to the appropriate weighting policy. Suppose that three out of four reviewers of a research proposal agree closely (consensus), but the fourth differs widely (dissension). A decision maker attempting to aggregate these opinions might rationalize the disagreement. Indeed, the need to form and maintain consonance, or harmony, is prominent in classical theories of social psychology (e.g., those of Heider and Festinger).

One mental process used to maintain consonance amounts simply to ignoring the dissonant pieces of information. Indeed, early studies of information integration (Anderson & Jacobson, 1965) and studies of judgments formed on the basis of numerical inputs of judgment (Slovic, 1966) have shown that people discount inconsistent inputs. Similarly, when intuitively combining a sample of opinions, people discount or completely ignore dissenters and assign greater weight to consensus opinions (Yaniv, 1997). Also, a dissenter's impact on a group's final decision declines as the discrepancy from the consensus increases (Davis, 1996).

On the one hand, decision makers who disregard divergent opinions could be ignoring good data because a dissenting estimate is not necessarily wrong. In general, the tendency to resolve inconsistencies by ignoring outlying views could reduce the quality of decision making. On the other hand, a policy of discounting outlying opinions might be justified if they tend to be wrong more often than consensus opinions. Certain structural aspects of the task might indicate when an outlying opinion is likely to be wrong. For example, suppose the distribution of opinions (in the population) is bell-shaped and thick-tailed. This implies that the prevalence of outlier opinions is larger than would be expected under a standard bell-shaped (normal) distribution. In such cases (assuming the bias is zero or small), an extreme opinion in a set is particularly likely to be wrong (see, e.g., DeGroot, 1986, for a discussion of the advantage of excluding outliers in estimating the center of a thick-tailed distribution). Therefore, discounting

dissenters is useful if one suspects that the distribution of opinions is thick-tailed, a situation not uncommon in behavioral studies (Yaniv, 1997).

Discounting dissenters might also be justified in scenarios in which one suspects exaggeration or manipulation. For example, in certain sports competitions, such as diving and gymnastics meets, performance is evaluated by several judges whose evaluations are then combined. Suppose a judge develops a liking for a certain performer and thus, consciously or unconsciously, produces an extreme, exaggerated evaluation that could unduly affect the aggregate opinion. A common practice in combining evaluations in such competitions involves dropping the most extreme evaluations (e.g., one on each end) and averaging the middle ones. Enacting a policy that discounts extreme judgments presumably dissuades judges from acting strategically and attenuates their influence if they do so (Yaniv, 1997).

Updating One's Own Opinion: Self Versus Other

Combining one's own opinion and an advisor's opinion is a special case that requires a separate discussion. Suppose you are responsible for hiring someone to fill a job, and you initially had a strongly favorable opinion about a candidate but are told that a colleague of yours has a lukewarm opinion of the same candidate. How might you revise your opinion in light of this conflict between your own and the other opinion? You could completely ignore the other opinion, make some adjustment of your own opinion toward the other, or completely adhere to the other opinion.

From a formal point of view, other things being equal, the two opinions (own and other) might be equally weighted. However, from your internal point of view, the two opinions are not on a par. Decision makers place more weight on beliefs for which they have more evidence. Because decision makers are privy to their own thoughts, but not to the reasons underlying an advisor's opinion, they place a higher weight on their own opinion than on an advisor's. Indeed, studies show that other things being equal, people discount others' opinions and prefer their own, with the weights split roughly 70% on self and 30% on other; this balance changes when differences in ability or knowledge between self and other are made salient (Harvey & Fischer, 1997; Yaniv, in press; Yaniv & Kleinberger, 2000). That individuals stick closely to their initial opinions is reminiscent of findings regarding attitude change–people favor their prior opinions even in the presence of contradictory evidence. But, despite the tendency to prefer one's own opinion over another person's opinion and the difficulty of assigning optimal weights to own versus other opinions, the benefit of utilizing others' estimates is appreciable. In one study (Yaniv, in press), respondents made initial estimates of the dates of historical events and final estimates after seeing other respondents' estimates, selected at random from a pool. Using just one other opinion reduced judgment errors by about 20%.

CONCLUDING COMMENTS

Students of reasoning, judgment, and decision making have traditionally underscored the importance of generating alternatives to one's current thoughts. Other people's opinions direct decision makers to additional alternatives or unintended consequences, as these opinions may provide a different framing of a problem, an alternative explanation, or disconfirming information. Soliciting opinions is therefore an adaptive process that helps improve decisions by compensating for a pervasive weakness of human thinking.

Two theoretical issues deserve attention. First, the view of opinions as alternatives is pertinent to opinions expressed in either numerical or verbal form. Although I have focused here on combining quantitative opinions, similar psychological processes might apply to verbal opinions (advice). Surprisingly, the use of advice in decision making has received little attention. Future research needs to consider how qualitative advice is elicited and used best.

Second, opinions about matters of fact (estimates or forecasts) differ from opinions about matters of taste (evaluations or attitudes). Theories about the benefit accrued from combining opinions about matters of fact are well developed. In contrast, simple aggregation of tastes (e.g., opinions about resorts or about types of music) for the purpose of individual decision making raises conceptual difficulties, because people are entitled to their different tastes. Nevertheless, other people's opinions about matters of taste could be used advantageously and constructively, challenging the decision maker's established preferences and inducing him or her to consider alternatives. Conceptual and empirical work is needed to clarify these issues.

REFERENCES

Anderson, N.H., & Jacobson, A. (1965). Effect of stimulus inconsistency and discounting instructions in personality impression formation. Journal of Personality and Social Psychology, 2, 531-539.

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

Birnbaum, M.H., & Stegner, S.E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. Journal of Personality and Social Psychology, 37, 48-74.

Clemen, R.T. (1989). Combining forecasts: A review and annotated bibliography. International Journal of Forecasting, 5, 559-583.

Davis, J.H. (1996). Group decision making and quantitative judgments: A consensus model. In E. Witte & J.H. Davis (Eds.), Understanding group behavior: Consensual action by small groups (pp. 35-59). Hillsdale, NJ: Erlbaum.

DeGroot, M.H. (1986). Probability and statistics (2nd ed.). Reading, MA: Addison-Wesley.

Einhorn, H.J., Hogarth, R.M., & Klempner, E. (1977). Quality of group judgment. Psychological Bulletin, 84, 158-172.

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment and sharing responsibility. Organizational Behavior and Human Decision Processes, 70, 117-133.

Hill, G.W. (1982). Group versus individual performance: Are N+1 heads better than one? Psychological Bulletin, 91, 517-539.

Hogarth, R.M. (1978). A note on aggregating opinions. Organizational Behavior and Human Performance, 21, 40-46.

Johnson, T.R., Budescu, D.V., & Wallsten, T.S. (2001). Averaging probability judgments: Monte Carlo analyses of asymptotic diagnostic value. Journal of Behavioral Decision Making, 14, 123-140.

Slovic, P. (1966). Cue-consistency and cue-utilization in judgment. The American Journal of Psychology, 79, 427-434.

Sniezek, J.A., & Van Swol, L.M. (2001). Trust, confidence, and expertise in a Judge-Advisor System. Organizational Behavior and Human Decision Processes, 84, 288-307.

Soll, J.B. (1999). Intuitive theories of information: Beliefs about the value of redundancy. Cognitive Psychology, 38, 317-346.

Sorkin, R.D., Hayes, C.J., & West, R. (2001). Signal detection analysis of group decision making. Psychological Review, 108, 183-203.

Yaniv, I. (1997). Weighting and trimming: Heuristics for aggregating judgments under uncertainty. Organizational Behavior and Human Decision Processes, 69, 237-249.

Yaniv, I. (in press). Receiving other people's advice: Influence and benefit. Organizational Behavior and Human Decision Processes.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. Organizational Behavior and Human Decision Processes, 83, 260-281.

Zarnowitz, V. (1984). The accuracy of individual and group forecasts from business and outlook surveys. Journal of Forecasting, 3, 11-26.