Scenarios as channels of forecast advice

Dilek Önkal [a,*], Kadire Zeynep Sayım [a], Mustafa Sinan Gönül [b]

[a] *Faculty of Business Administration, Bilkent University, 06800 Ankara, Turkey*
[b] *Department of Business Administration, Middle East Technical University, 06800 Ankara, Turkey*

ARTICLE INFO

ABSTRACT

Today's business environment provides tougher competition than ever before, stressing the important role played by information and forecasts in decision-making. The scenario method has been popular for focused organizational learning, decision making and strategic thinking in business contexts, and yet, its use in communicating forecast information and advice has received little research attention. This is surprising since scenarios may provide valuable tools for communication between forecast providers and users in organizations, offering efficient platforms for information exchange via structured storylines of plausible futures. In this paper, we aim to explore the effectiveness of using scenarios as channels of forecast advice. An experimental study is designed to investigate the effects of providing scenarios as forecast advice on individual and group-based judgmental predictions. Participants are given time series information and model forecasts, along with (i) best-case, (ii) worst-case, (iii) both, or (iv) no scenarios. Different forecasting formats are used (i.e., point forecast, best-case forecast, worst-case forecast, and surprise probability), and both individual predictions and consensus forecasts are requested. Forecasts made with and without scenarios are compared for each of these formats to explore the potential effects of providing scenarios as forecast advice. In addition, group effects are investigated via comparisons of composite versus consensus predictions. The paper concludes with a discussion of results and implications for future research on scenario use in forecasting.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Managerial decision making faces tremendous challenges in today's relentlessly fluctuating business environments. Staggering amounts of change in social, political, economic and cultural fronts make it excessively difficult to foresee the future for products and markets. There is no scarcity of data or 'tools' to turn it into meaningful information, particularly given the vast online resources and knowledge management support. Yet, this abundance of information does not mean that managers can predict the future easily and confidently, especially in face of persistently expanding uncertainties that make statistical models insufficient. As a result, forecasts represent one of the key judgments made in companies [1].

Although statistical techniques are widely used, human judgment plays a strategic role in business and economic forecasts (e.g., see [2–5]). In corporate practice, managers commonly use their judgment to adjust the statistical baseline forecasts [2,6,7], to the extent that there may actually be layers of adjustments by forecast users at different levels of decision making [8]. How the users accept or modify given forecasts (regardless of whether this information is presented through models or experts) presents an important issue for researchers investigating advice taking [9]. Extant work shows that advice could be sought to improve decision quality, to diffuse responsibility, or to ensure continued information flow [10,11]. However, advice is typically modified

---

* Corresponding author.
E-mail addresses: onkal@bilkent.edu.tr (D. Önkal), kzeynep@bilkent.edu.tr (K.Z. Sayım), msgonul@metu.edu.tr (M.S. Gönül).

in favor of users' own opinions [12–14]. The level of discounting may depend on multitude of factors including perceived expertise and difference between own judgment and given estimates [15]. One explanation could be that individuals can easily find justifications for their own predictions, whereas they do not have access to the forecast provider's rationale; thus forecast advice is not perceived as directly justifiable [12,15,16].

Given their capacity to supply additional information on alternative futures, scenarios may be used to expand forecast advice to more than 'dry' numbers. Hence, they may face less resistance than other forms of advice, and may act as antidotes to unnecessary and extreme adjustments that are commonly observed in business settings [e.g., 2,7]. Even though the scenario method has been popular for focused organizational learning, decision making, and strategic thinking in business contexts [17–24], its use in communicating forecast advice has received little research attention. This is surprising since scenarios offer open platforms for information exchange via structured storylines of plausible futures.

Requiring decision makers to construct their own scenarios provides an important approach to future-focused thinking [23,25]. This paper explores a complementary forecasting perspective whereby scenarios are not asked but are given as additional information to convey viable alternative futures. When used to provide forecast advice in this way, scenarios may represent powerful communication tools to assist a multi-faceted questioning of conceivable future values. Pursuing this perspective, the current study aims to examine the effects of providing scenarios on the users' responses to given forecast advice and their resulting judgmental adjustments. In the light of potential differences in how individuals and groups may react to forecast advice [26], we compare individual judgments with responses from small groups which are highly prevalent in organizations. Using a representative organizational task (i.e., forecasting of demand for a product line), we focus on the potential use of best-case and worst-case scenarios given solo or in tandem. Presentation of a single scenario may be imperative in domains like marketing, production and financial forecasts, where an expert advisor may want to emphasize a particular best-case or worst-case. However, providing information on both extremes could be critical from the users' perspective, as this may influence perceptions of incomplete information and bias. Hence, examining comparative impact of single versus multiple scenarios may be essential to gain a detailed understanding of forecast advice use, as explored in this paper.

Mobile telecommunications industry is chosen as the forecasting background as this industry is characterized by rapid change in trends and technology, fierce competition and fast pace accompanied by high levels of uncertainty. What is considered as long-term for such a highly volatile industry may be equivalent of short/medium-term for other industries where change occurs at relatively slower rates. In such rapidly fluctuating business situations, the value of scenarios in supporting forecasting processes may be even more pronounced, as explored in this paper.

Accordingly, the remaining sections are organized as follows: Section 2 provides a review of the relevant literature and identifies the research gaps that constitute the focus of our study. Methodology and details of the research procedure are presented in Section 3, followed by the results in Section 4. Section 5 offers a discussion of the findings, and the concluding comments are given in Section 6.

2. Literature review and research gap

Scenarios have been used in corporate strategy and planning since the early 1960s [27,28 — cf 2]. In the earlier variants, such as 'La Prospective' approach, the scenario method is defined as "… a powerful tool for constructing alternative futures … and exploring the pathways leading to them" [27:298]. Moreover, scenarios are "… flowing narratives rather than precise quantitative estimates" [29:405], and they "… should bound the range of plausible uncertainties and challenge managerial thinking" [30:549]. Therefore, scenarios do not tell *what* lies in the future; instead they emphasize *how* that future might evolve [31].

Presented as dynamic storylines incorporating possible futures, scenarios may be perceived as having more credibility by users in comparison to dull statistical projections, hence may have more impact on judgments [29,32]. For example, Anderson [33] concludes that the use of scenarios has a larger influence on users' judgments than that of statistical data. Gregory, Cialdini and Carpenter [34] find that users are more likely to assign higher likelihoods to those events contained in scenarios relative to other events that do not appear in scenarios. Advocates of scenarios claim that scenario analysis is valuable in providing alternative viable portrayals of the uncertainties, which may force decision makers to formulate strategies in response to any of these potential outcomes, thus decreasing confidence in a single prediction [35]. An associated drawback is that users, who are accustomed to dealing with a single projection of the future, might prefer to focus on a favored scenario and ignore the others [29,36]. For instance, when three scenarios are given, users may end up focusing on the midway scenario [37]. Thus, the presentation of alternative scenarios needs to be carefully planned.

Method-based statistical forecasting and forecasting with scenarios can be regarded as quite different, even contradictory approaches, in terms of their respective assumptions, purposes, and processes. A forecast is generally defined as a reflection of expert opinion based on probability assessments, i.e. '… the distillation of much expertise into one number or probability distribution' [30:551]. A scenario on the other hand is a conceptual description of a plausible future, which emphasizes the underlying reasoning as well as sources of uncertainty [31]. When a manager is given a forecast (point or interval prediction) made by an expert, the underlying assumptions are typically unknown to the user. Moreover, these assumptions may not always encompass the unexpected factors in the business environment, whereas they may represent the key starting points for scenarios. It is further claimed that a forecast is an efficient but 'impoverished' summary of describing the future that reduces rich information into a simplified form [38]. In contrast, scenarios, via the proposed chains of cause and effect that focus specifically on rich information, aim at triggering thinking processes and making further judgments. Consequently, there is significant support for using scenarios together with more traditional statistical forecasts in order to supplement the latter, especially when uncertainty and complexity in

the external environment is high, when organizational difficulties have been experienced in the past, and when the previous outcomes have been unfavorable [30]. Thinking with scenarios enables the decision makers to visualize different plausible futures and enhances their awareness and anticipation about what might be forthcoming. In particular, Bunn and Salo [39] maintain that scenarios may work well in conjunction with other forecasting techniques, facilitating the generation of forecasts and contributing positively to the process.

In a similar vein, it is argued that scenarios can assist with problems stemming from cognitive biases [24]. This is an important advantage because cognitive biases may hinder effective decision-making and planning [40]. Hence, there exists a large volume of work concentrating on remedial procedures for biases like anchoring (relying too heavily on a particular value), availability (focusing on the most salient and emotionally-charged information), framing (drawing different conclusions based on the identical information depending on its presentation format), recency (placing more weight on recent events), illusory correlation (inaccurately presuming relationships between events), overconfidence (having unwarranted confidence in one's knowledge and skills), confirmation bias (searching for information in a way that confirms one's preconceptions/expectations), and hindsight bias ("I-knew-it-all-along" effect), among a host of other cognitive biases (see [3,41–43] for reviews). It is argued that while the inappropriate use of scenarios may facilitate such biases, appropriate scenario construction techniques (e.g., incorporating stakeholder analysis to integrate wider perspectives; inviting 'remarkable people' to challenge the team members when constructing forecasts) may prove valuable in creating awareness and providing feedback to minimize their occurrence [44–46].

Given their strengths, scenarios have received a growing research interest in the forecasting domain (e.g., [31,44,47]). Focusing mainly on constructing scenarios, this work has encouraged employing multiple scenarios [44] and suggested using individuals with diverse views of the future to overcome biases such as frame blindness [47]. Overall, the importance of scenario thinking in forecasting has been heavily emphasized. Even with this emphasis, there has been surprisingly little empirical work on providing scenarios as forecast support, with few exceptions.

One of the early studies in this regard was carried out by Schnaars and Topol [29]. Providing 'optimistic', 'pessimistic', and 'middle-ground' scenarios to half of the participants (whereas the other half did not receive scenario information), the researchers found the scenarios to unwarrantedly increase participants' confidence, with no accompanying improvements in accuracy. As acknowledged by the authors, this work provides a partial test of scenario impact since the given scenarios were designed to exclude any exogenous factors and, hence, provided no stories for participants to expect any changes in the values. Alternatively, Kuhn and Sniezek [1] provided psychology students with one or two scenarios containing uni-directional or conflicting information and compared their results with students not receiving any scenarios. Their findings showed that giving any type of scenario information increased confidence (as measured on a confidence scale between 1 and 9), and also that providing multiple conflicting scenarios did not reduce confidence as compared to giving one scenario. Concluding that the key benefit of multiple scenarios in forecasting is therefore *not* attaining reduced levels of confidence but rather accepting a wider range of likely outcomes, the researchers argued that more work is needed to examine the relationships between acceptance/ rejection of scenarios and judgments about future, particularly via measures other than confidence in predictions. Finally, Schoemaker [48] used a different approach and asked graduate business students to write their own scenarios leading to different outcomes (i.e., one scenario resulting in an increase, the other leading to a decrease in the forecast variable). His findings showed that the participants gave wider prediction intervals after writing two contrasting scenarios, hence demonstrating lower confidence (as measured by interval widths).

Developing and writing scenarios are unquestionably important for organizations, entailing a combination of various techniques that are essential for futures studies [49]. However, the users of scenarios in organizations are generally not the same individuals who are responsible for their development, and their responses to scenarios are equally important for the resulting decisions [1]. Approaching from the users' perspective, the current study focuses on the effects of given scenarios on the forecast adjustments made by individuals and groups. While extant research has mostly examined scenarios as tools for constructing forecasts (thus studying forecast accuracy), our study uses scenarios as channels for communicating forecast advice (thus focusing on judgmental adjustments and acceptance of given forecast advice). Hence, the current work proposes a complementary perspective for examining the role of scenarios in providing additional information to forecast users.

In studying how forecast advice is used, we particularly concentrate on two-person groups (dyads) as they exemplify the decision-making groups commonly encountered in organizational settings [50]. Two-person groups are often used for problem solving, since the small group size minimizes the process losses (e.g., social loafing and groupthink), while maintaining other advantages of groups (e.g., broader range of ideas and skills) [51]. Almost half of managerial communication [52] and majority of negotiation processes [53] reportedly take place in dyads. Moreover, Leader Member Exchange (LMX) theory is a conceptualization of a dyadic relationship between a leader and each follower [54]. In short, dyads assume a special significance in a wide spectrum of managerial fields, such as communication, leadership, negotiation and decision making (e.g., [55–60]). For instance, past work has shown significant support for the positive influence of working in dyads (as opposed to working alone) on reducing overconfidence [61,62]. This may be explained by small group dynamics. In particular, when working in dyads, there is a need to justify one's estimate (e.g., [63]) and to convince the other party, with the negotiation process potentially affecting individual judgments and decisions. In this regard, dyadic interactions may be considered similar to taking advice for decision making; they enable the advice-taker to benefit from other's information to outweigh self-confirmation tendencies [15], thus leading to less biased judgments.

Given the importance of dyads in organizational contexts and the scarce empirical work on integrating dyads into structured forecasting studies, this paper examines forecast adjustments of individuals *and* dyads in response to forecast advice. In particular, we present an experimental study designed to investigate the effects of providing scenarios as forecast advice on individual and dyad-based judgmental predictions. Participants are given time series information and model forecasts, along with (i) best-case,

(ii) worst-case, (iii) both, or (iv) no scenarios. Different forecasting formats (i.e., point forecast, best-case forecast, worst-case forecast, and surprise probability) are used and the final forecasts of individuals, followed by consensus forecasts, are requested.

This work aims to fill research gaps in current literature as outlined above in terms of its synthesis in (1) incorporating scenarios into forecast advice, (2) examining effects of single *versus* multiple scenarios, and (3) comparing individuals' *versus* small groups' judgmental adjustments to given forecast advice. Using a representative organizational forecasting task (i.e., predicting demand for a product line) as opposed to general forecasting tasks (e.g., the mean world temperature or the number of nations in the next 50 years, as employed in previous work [1]), it provides an exploratory step towards integration of scenarios into forecast communication and advice taking.

3. Methodology

3.1. Participants and procedure

A total of 120 business students from Bilkent University and Middle East Technical University completed the experimental tasks. The study involved a paper-and-pencil format and started with initial instructions and background information about a (hypothetical) mobile telecommunications company producing a range of cellular phones (as given in Appendix A). Task requirements consisted of two phases involving individual forecasts followed by consensus forecasts, with a total of 120 minutes allocated to the study.

3.1.1. Phase 1 (individual forecasts)
In the first phase, each participant was provided with an "Individual Forecasts Form" that included task instructions and 18 time-series plots exhibiting historical demand for 18 products from the case company. Each time-series plot showed a particular product's demand over the past 20 months, and displayed a one-period-ahead model-based point forecast for the next month to serve as forecast advice. Scenarios appropriate to the experimental group (if any) were also on this form. Participants were then required to generate their point, best-case, and worst-case forecasts and to assess a surprise probability for their given forecast range (i.e., probability that the actual demand for the next period will turn out to be higher than their best-case forecast, or lower than their worst-case forecast). Afterwards, they provided ratings regarding the information value, usefulness, and the influence of the given scenarios (if any) in constructing their forecasts. This process was repeated for all 18 products (please see Appendix B for a sample "Individual Forecasts Form" given to G2 participants).

Once this stage was completed and the individual forecasts forms were collected, participants were assigned to their respective two-person groups (i.e., dyads) for the next phase of the study.

3.1.2. Phase 2 (consensus forecasts)
In the second phase, each participant was provided with a "Consensus Forecasts Form" that included the *same* 18 time-series plots, the model-based forecasts and the corresponding scenarios (if any). In this stage, the participants were requested to (i) discuss the given model-based forecasts, past demands, and scenarios (if any) as a dyad, and (ii) arrive at consensus forecasts in the form of point, best-case, and worst-case predictions, as well as their dyad's surprise probability assessments (i.e., assessments of predicted surprise index) for each of the 18 products. Upon completing the consensus forecasts for each product, they were asked to individually convey their level agreement with these consensus predictions (please see Appendix C for a sample "Consensus Forecasts Form" given to X2 participants receiving best-case scenarios only). After all the 18 products' forecasts were completed, they were requested to fill out an exit questionnaire.

3.2. Design

As outlined above, the study was designed to elicit the individual forecasts, followed by the consensus forecasts in two consecutive phases.

3.2.1. Phase 1 (individual forecasts)
In the first stage, the effects of providing best-case scenarios and/or worst-case scenarios on individual forecasts were investigated. Participants were randomly assigned to one of four experimental groups:

G1 — no scenarios given: Participants in this category received no scenarios and were presented with the time-series plots and model-based forecasts only. There were 36 individual participants in G1.
G2 — only best-case scenario given: In addition to the time-series plots and model-based forecasts, participants in this experimental group received best-case scenarios for the next period. There were 24 individual participants in G2.
G3 — only worst-case scenario given: In addition to the time-series plots and model-based forecasts, participants in this experimental group received worst-case scenarios for the next period. There were 24 individual participants in G3.
G4 — both best-case and worst-case scenarios given: Participants in this experimental group received both the best-case and the worst-case scenarios for the next period simultaneously as well as time-series plots and model-based forecasts. There were 36 individual participants in G4.

3.2.2. Phase 2 (consensus forecasts)

Phase 2 was designed to compare the consensus forecasts from experimental groups not given any scenarios with those from (a) groups where each participant receives a different scenario (i.e., one member is given a best-case scenario, while the other member is given a worst-case scenario), and (b) groups where both members are given both scenarios (i.e., both members receiving both a best-case and a worst-case scenario). Hence, there were three experimental groups in Phase 2 (where the participants were assigned to their dyads in accordance with their allocations in Phase 1):

X1 — no scenarios: the dyads in this group were entirely composed of participants from G1 in Phase 1, thus receiving no scenarios in either stage of the study. There were 18 dyads in X1.

X2 — One member received only best-case scenario — other member received only worst-case scenario: This group was composed of dyads where one member was from G2 (getting only the best-case scenarios), while the other member was from G3 (getting only the worst-case scenarios) in the individual forecasts stage. There were 24 dyads in X2.

X3 — Both members received both scenarios: Here, the dyads were composed of participants of G4 in Phase 1. Thus, both members received both scenarios simultaneously. There were 18 dyads in X3.

It is important to note that the participants received the identical scenario information on their forms in two phases of the study. The only additional information in Phase 2 was that provided by their discussion with the other member in their dyads.

3.2.3. Time series

Time series framed to represent the past demand values for the 18 products of the case company were constructed artificially to control for trend and uncertainty. The series incorporated three levels of trend (positive, negative, and stable) and two levels of uncertainty (low and high). By constructing three samples from these six combinations, a total of 18 series were employed. These series were presented to the participants in a random order. The procedure used for generating the artificial series was parallel to previous studies on judgmental forecasting (e.g. [8,64]). In particular:

$$y(t) = 125 + bt + \text{error}(t) \quad t = 0, 1, \ldots., 20$$

The trend coefficient b had a value of $+5$ for series with positive trends, -5 for negative trends, and 0 for those with stable trends. The error is distributed normally with zero mean and a standard deviation of either 10% (i.e., 0.1×125) for series with low uncertainty or 20% (i.e., 0.2×125) for those with high uncertainty. The model-based forecasts given to the participants were generated by using Holt's exponential smoothing method with error-minimizing parameters.

3.2.4. Scenarios

The scenarios were constructed following the generally accepted principles and guidelines proposed by previous work (e.g., [38,45,48]). First, a broad range of material was collected using publicly available data and by talking to experts in the mobile telecommunications sector. This information was summarized and a short list of industry trends and factors were generated. Then, the scenarios were constructed using the most significant variables relating to the industry and product, and the trends in the given series. It is important to note that both endogenous (e.g., product life cycle) and exogenous (e.g., changes in information technology, market forces, competitors) information were included so that the scenarios exemplified real-life forecasting environments. Hence, many key uncertainties were made transparent in the scenarios, unlike the model-based forecasts that cannot offer such extensive future-focused information. Also, within the information supplied by the scenarios, specific attention was paid to internal consistency, i.e., cause and effect arguments were included as consistently as possible. Lastly, feedback from industry experts was sought before the constructed scenarios were put to use in the experiment.

In line with the argument that at least two scenarios are needed to reflect the inherent environmental uncertainties (e.g., [38,65,66]), two extreme sets of scenarios (i.e., best-case vs worst-case) were constructed for each of the 18 time-series. Specifically, both sets of scenarios designed for each time-series were matched vis-à-vis length, content, language, and the type of presented information, based on the ratings of a separate group of forecast users.

3.2.5. Scenario framing

In addition to the original study detailed above, a *framing-manipulation* study was conducted to examine the sensitivity of forecast adjustments to the particular labeling of scenarios as "best-case" and "worst-case".[1] Thirty students participated in this manipulation study, where each participant received one of two unlabeled scenarios for each time series (previously labeled as best-case and worst-case scenarios — as in G2 and G3 of primary study). After completing their individual forecasts, the participants were paired so that the members of the dyad received different unlabeled scenarios (while each member retained the same scenario given in the individual forecasting stage — as in X2 of primary study). The experimental procedure and the task requirements were the same as in the primary study.

---

[1] We are grateful to two anonymous reviewers for emphasizing this point.

4. Results

Participants' assessments and perceptions of given scenarios constitute a key component of this study. If scenarios are not considered to be informative and/or useful, other findings relating to individual and consensus forecasts lose their meaning. Thus, we start by focusing on these findings and introducing the performance measures utilized. Results from individual and consensus forecasting phases are summarized, followed by a comparison of consensus versus composite predictions. Finally, the effects of scenario framing are examined.

4.1. Perceptions of scenarios

Participants' responses reveal that the scenarios were successful in achieving their intended purpose. The proportion of participants indicating that they found the scenarios to be informative, useful, and influential (via ratings of 3–4–5 on a scale of *1 = not at all informative/useful/influential* to *5 = extremely informative/useful/influential*) were around 93% for participants given only best-case scenarios, 90% for those given only worst-case scenarios, and 83% for those given both scenarios.

It is worth noting that no significant differences in any of the ratings given for best-case scenarios vs for worst-case scenarios could be found between the experimental groups. This suggests the 'equivalence' of the best-case and worst-case scenarios in terms of participants' perceptions of informativeness, usefulness, and influence of the scenarios on their final forecasts when they received only one of these scenarios as forecast advice.

On the other hand, participants provided with only the best-case scenarios (G2) rated these scenarios as being more informative ($t_{57} = 3.06$, $p = .003$), more useful ($t_{52} = 3.03$, $p = .004$), and more influential ($t_{47} = 3.47$, $p = .001$) for their final forecasts, when compared to participants who were presented with both the best-case and the worst-case scenarios (G4). Given that identical best-case scenarios were presented in both situations, it appears that best-case scenarios lose their 'perceived value' when accompanied by worst-case scenarios, whereas their appeal peaks when presented solo. This is very interesting since the reverse does not seem to take place — that is, there are no significant differences in how the worst-case scenarios were perceived whether they were given alone or alongside best-case scenarios (i.e., no significant differences in perceptions of G3 vs G4 participants).

Additionally, global evaluations of scenarios were requested in the exit questionnaire at the end of Phase 2. That is, the participants were asked to rate how they viewed *all* the given scenarios according to various dimensions. As portrayed in the radar graph in Fig. 1, these findings also provide evidence for the positive response to scenarios, as they were found to be realistic
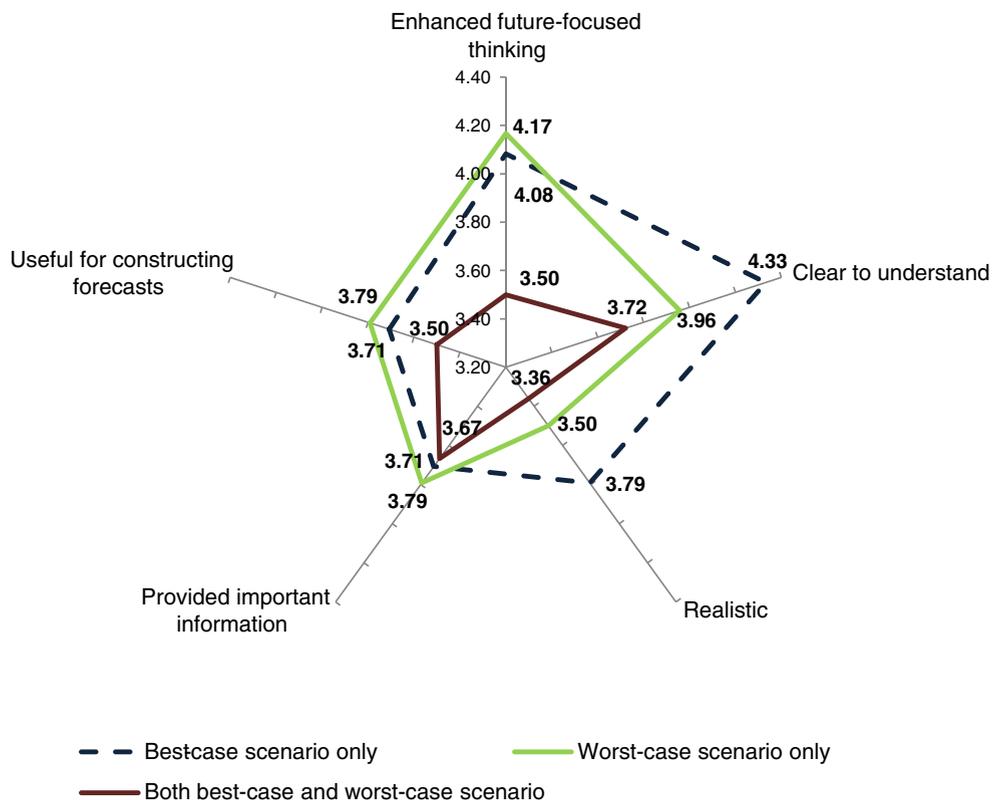

Fig. 1. Mean evaluation ratings of given scenarios (as given by participants in the exit questionnaire at the end of Phase 2).

and clear to understand, providing important information, enhancing future-focused thinking, as well as being useful for constructing the final forecasts (i.e., all ratings were significantly higher than 3.00 on scales of *1 = definitely disagree* to *5 = definitely agree* for each dimension; all $p < .05$). These overall evaluations again show that the participants in experimental groups receiving only best-case scenarios rated them as being clearer to understand than the participants receiving both best-case and worst-case scenarios ($t_{54} = 3.10$, $p = .003$). Furthermore, individuals receiving only the best-case or only the worst-case scenarios believed that the scenarios enhanced their future-focused thinking significantly more so than the individuals receiving both scenario sets ($t_{57} = 2.43$, $p = .018$ and $t_{57} = 3.00$, $p = .004$, respectively). Given that no other significant differences existed in the other dimensions, these findings may be viewed as suggesting a cognitive overload. When information is given in terms of one scenario, it may be easier to imagine a specific potential future; but when two possible extremes are given, this may lead to difficulties in thinking of alternate plausible futures concurrently.

4.2. Performance measures

Five performance measures are utilized to investigate the potential effects of scenarios on the final forecasts of individuals and dyads:

(i) percentage change of final *point forecasts* from the given model-based point forecasts,
(ii) percentage change of final *best-case forecasts* from the given model-based point forecasts,
(iii) percentage change of final *worst-case forecasts* from the given model-based point forecasts,
(iv) surprise index, and
(v) asymmetry ratio.

The first three measures capture the percentage difference between the participants' final forecasts (in the form of point, best-case and worst-case predictions) and the model-based point forecasts. That is, they offer indications of the extent to which given forecast advice is actually used. Last two measures provide information about the prediction intervals whose bounds are defined by best-case and worst-case forecasts, as detailed below. In particular, the surprise index conveys forecasters' confidence while the asymmetry ratio highlights forecasters' tendencies to lean towards best vs worst-case bounds.

Starting with measures of forecast advice use, percentage change scores allows for unit-free and scaled comparisons among experimental groups. Positive scores will reveal upward adjustment from the provided point prediction, whereas negative scores will indicate downward adjustment. A score of zero will represent no change (i.e., forecast advice is accepted as given with no adjustments). The formulas used to calculate these scores are as follows:

$$\text{Percent Change Point Frcst} = 100 \times \frac{(\text{generated point forecast} - \text{provided point forecast})}{\text{provided point forecast}}$$

$$\text{Percent Change BestCase Frcst} = 100 \times \frac{(\text{generated bestcase forecast} - \text{provided point forecast})}{\text{provided point forecast}}$$

$$\text{Percent Change WorstCase Frcst} = 100 \times \frac{(\text{generated worstcase forecast} - \text{provided point forecast})}{\text{provided point forecast}}$$

The fourth performance measure is the Surprise Index (SI). This is computed as the average value across all surprise probabilities that the realized values will fall outside the prediction limits (i.e., mean probability of actual values turning out to be higher than final best-case forecasts or lower than final worst-case forecasts). This index may be regarded as conveying the assessor's calibration expectations. Similarly, it may also be viewed as a proxy measure of the assessor's confidence. As stated by our participants in their exit interviews, lower surprise probabilities (leading to lower values of the surprise index) were used to express their expectation that the given forecast interval (as bounded by best/worst-case predictions) would capture the true value with a higher probability, hence pointing to greater confidence in their given interval.

The fifth performance measure is the *Asymmetry Ratio* (*ASR*), which provides an index of how symmetric/asymmetric an interval is. As suggested by O'Connor, Remus and Griggs [67], the ASR score is calculated via the following:

$$\text{ASR} = 100 \times \frac{(\text{point forecast} - \text{worstcase forecast})}{(\text{bestcase forecast} - \text{worstcase forecast})}$$

If the point forecast rests exactly in the middle of best-case and worst-case forecasts (i.e., the interval that is bound by best-case and worst-case predictions is symmetric with respect to the generated point prediction), then the corresponding ASR score will be 50%. When a best-case forecast is further away from the point prediction than a worst-case forecast, ASR will be less than 50%. On the other hand, when a worst-case prediction is further away from a point forecast as compared to its corresponding best-case prediction, ASR will be greater than 50%. Both these cases indicate asymmetry in intervals, highlighting forecaster tendencies to lean towards best vs worst-case bounds (i.e., tendencies to assess prediction bounds of unequal distances above and below their point forecasts).

4.3. Individual forecasts

Individual forecasts elicited in Phase 1 were examined for the five performance measures discussed above. Table 1 portrays these findings, with the scenario availability contrasts highlighted in Table 1A, and the comparisons among the four experimental groups summarized in Table 1B. Overall, when scenarios are not given, participants appear to make larger adjustments to given forecast advice. Specifically, their best-case forecasts reflect an average upward adjustment of 29% to the given forecast advice (whereas this value is 22% when scenarios are given, a statistically significant difference; $F_{1,118} = 5.75$, $p = .018$). This is also accompanied by relatively lower confidence levels, as signaled by higher surprise index values (i.e., when scenarios are not given, participants expect the real values to fall outside their specified intervals with a mean probability of 16% — this shrinks to 12% when scenarios are available; $F_{1,118} = 4.64$, $p = .033$). These findings appear to suggest that providing a scenario increases the plausibility of the given forecast advice, leading to less adjustment (to given model-based predictions) and increased confidence in the final forecasts. This may of course depend on the framing of the scenarios as well as the perceptions regarding the given forecast advice, as will be discussed later.

In addition, providing <u>best-case scenarios</u> appears to have a significant main effect across all performance measures ($F_{1,116} = 30.32$, $p < .0001$; $F_{1,116} = 5.51$, $p = .021$; $F_{1,116} = 20.52$, $p < .0001$; $F_{1,116} = 9.48$, $p = .003$; and $F_{1,116} = 4.91$, $p = .029$ for the five performance measures, respectively). When only best-case scenarios are given, individuals appear to (a) make larger adjustments to the initially provided model-based forecasts to arrive at their final point forecasts, (b) be more confident, and (c) give the most symmetric intervals as compared to the other experimental groups.

On the other hand, the presence of <u>worst-case scenarios</u> shows a main effect on forecast use, as indexed by percentage change of point/best-case/worst-case forecasts from the initially provided model-based forecasts ($F_{1,116} = 64.00$, $p < .0001$; $F_{1,116} = 32.55$, $p < .0001$; $F_{1,116} = 17.22$, $p < .0001$); but with no corresponding effects on interval performance, as measured by SI and ASR. Participants receiving only the worst-case scenarios tend to (a) lower the given model-based forecast advice the most among the four groups (to arrive at their final point forecasts), (b) make best-case predictions that are closest to the forecast advice, and (c) give worst-case forecasts that are adjusted more than the other experimental groups.

4.4. Consensus forecasts

Mean performance scores for consensus forecasts are given in Table 2. Table 2A summarizes the performance scores for groups based on scenario availability, while Table 2B provides a breakdown for each of the experimental groups. Interestingly, the scores appear to be quite similar across groups with and without scenarios, significant differences appearing only in the best-case forecasts' positioning relative to the initial model-based forecasts and the surprise index. In particular, when scenarios are given, consensus best-case forecasts end up being significantly closer to the model-based forecasts than those of experimental groups receiving no scenarios ($F_{1,58} = 14.06$, $p < .0001$), accompanied by increased confidence ($F_{1,58} = 17.62$, $p < .0001$). That is, when a best-case and/or a worst-case scenario is/are given to the dyads, their expected likelihood of being caught in a surprise is assessed as being significantly lower. These findings are similar to results with individual surprise probabilities.

No significant differences in any of the performance measures could be found between groups where members receive different scenarios (i.e., X2 with one member given best-case scenario, while the other member is given worst-case scenario) and groups where both members receive both scenarios (X3) (all $p > .10$). This shows that group discussion was equally effective in communicating the differential scenario information given to members of dyads in X2, when compared with X3 dyad members who received identical information on *both* scenarios in their individual forms.

The participants were also asked their levels of agreement with the consensus forecasts reached in their dyads; these responses are summarized in Fig. 2. Given that such small group sizes are found to negate biases resulting from groupthink [51],

Table 1
Mean performance scores for individual forecasts (number of individuals within each category is given in parenthesis).

	% Change of point forecasts from model-based point forecasts	% Change of best-case forecasts from model-based point forecasts	% Change of worst-case forecasts from model-based point forecasts	Surprise Index	Asymmetry Ratio
A. Mean performance scores for experimental groups with and without scenarios					
No scenarios given (G1)	2.67%	29.31%	−23.22%	16.25%	55.17%
	(36)	(36)	(36)	(36)	(36)
Scenarios given (G2, G3 & G4)	0.86%	22.06%	−19.93%	12.27%	53.11%
	(84)	(84)	(84)	(84)	(84)
B. Mean performance score breakdowns for each experimental group					
No scenarios given (G1)	2.67%	29.31%	−23.22%	16.25%	55.17%
	(36)	(36)	(36)	(36)	(36)
Only best-case scenario given (G2)	10.65%	33.41%	−10.05%	9.42%	50.89%
	(24)	(24)	(24)	(24)	(24)
Only worst-case scenario given (G3)	−6.80%	12.77%	−25.93%	15.60%	54.68%
	(24)	(24)	(24)	(24)	(24)
Both best-case and worst-case scenarios given (G4)	−0.56%	20.70%	−22.52%	11.94%	53.54%
	(36)	(36)	(36)	(36)	(36)

Table 2
Mean performance scores for consensus forecasts (number of dyads within each category is given in parenthesis).

	% Change of point forecasts from model-based point forecasts	% Change of best-case forecasts from model-based point forecasts	% Change of worst-case forecasts from model-based point forecasts	Surprise Index	Asymmetry Ratio
A. Mean performance scores for experimental groups with and without scenarios					
Experimental groups not given any scenarios (X1)	2.97% (18)	29.93% (18)	−21.92% (18)	16.91% (18)	53.31% (18)
Experimental groups given scenarios (X2 & X3)	0.51% (42)	19.03% (42)	−19.49% (42)	8.31% (42)	53.98% (42)
B. Mean performance score breakdowns for each experimental group					
No scenarios (X1)	2.97% (18)	29.93% (18)	−21.92% (18)	16.91% (18)	53.31% (18)
One member received only best-case scenario; other member received only worst-case scenario (X2)	−0.35% (24)	16.97% (24)	−19.33% (24)	8.01% (24)	53.99% (24)
Both members received both scenarios (X3)	1.66% (18)	21.78% (18)	−19.71% (18)	8.72% (18)	53.96% (18)

significantly high levels of agreement displayed in Fig. 2 may be viewed to represent true indicators of their agreement with the consensus forecasts reached. It is worth noting that no significant differences in agreement ratings among the experimental groups could be found for any of the forecast formats (i.e., point, best-case or worst-case forecasts) (all $p > .10$).

4.5. Composite Forecasts

A composite forecast is the simple average of the independent forecasts generated by individual participants. Composite forecasts (arrived via arithmetical calculations) constitute a benchmark for evaluating consensus forecasts (arrived via group discussions). Following the notion of *wisdom of crowds* [68], composite forecasts offer a convenient alternative to structured groups, although their effectiveness appears to be dependent on task characteristics [69]. For the current study, composite point forecast can be calculated as:

$$\text{Composite point forecast for group } i = \frac{\text{point forecast}_{member1} + \text{point forecast}_{member2}}{2}$$
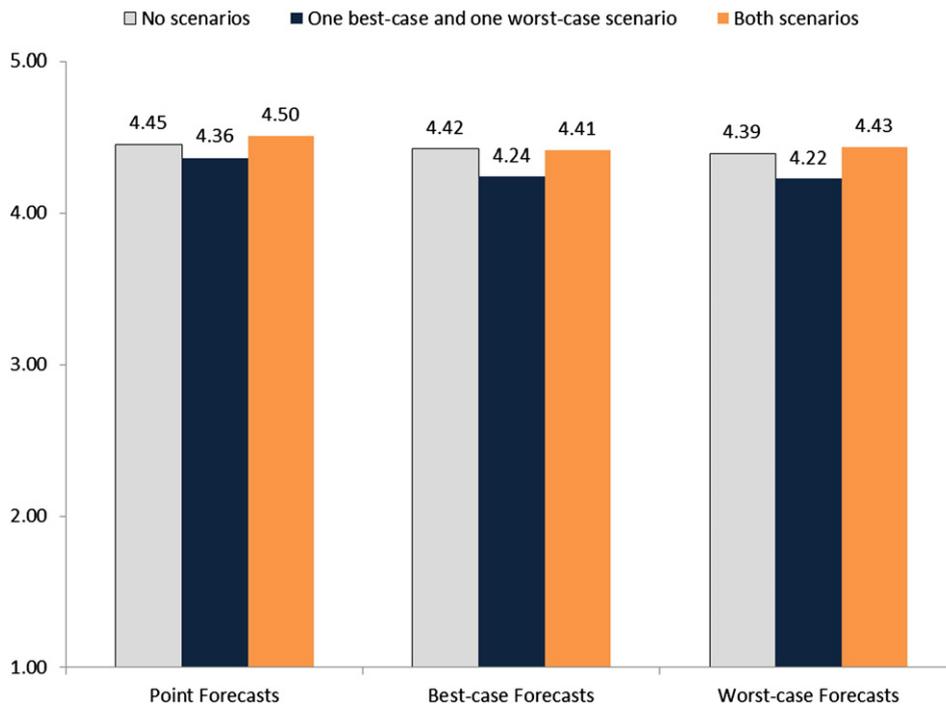

Fig. 2. Mean levels of agreement with consensus forecasts.

Table 3
Mean performance scores for composite forecasts (number of dyads within each category is given in parenthesis).

	% Change of point forecasts from model-based point forecasts	% Change of best-case forecasts from model-based point forecasts	% Change of worst-case forecasts from model-based point forecasts	Surprise Index	Asymmetry Ratio
A. Mean performance scores for experimental groups with and without scenarios					
Experimental groups not given any scenarios (X1)	2.67% (18)	29.31% (18)	−23.22% (18)	16.24% (18)	55.56% (18)
Experimental groups given scenarios (X2 & X3)	0.86% (42)	22.06% (42)	−19.93% (42)	12.27% (42)	53.03% (42)
B. Mean performance score breakdowns for each experimental group					
No scenarios (X1)	2.67% (18)	29.31% (18)	−23.22% (18)	16.24% (18)	55.56% (18)
One member received only best-case scenario; other member received only worst-case scenario (X2)	1.93% (24)	23.09% (24)	−17.99% (24)	12.51% (24)	52.73% (24)
Both members received both scenarios (X3)	−0.56% (18)	20.70% (18)	−22.52% (18)	11.94% (18)	53.43% (18)

The formulas used to calculate composite best-case forecast, composite worst-case forecast and composite surprise probability are similar.

Performance scores for the composite forecasts are displayed in Table 3, with Table 3A showing the performance scores for groups given *vs* not given scenarios, while the breakdown for each of the experimental groups displayed in Table 3B. Findings are similar to those of consensus forecasts in that the only significant differences are found for the surprise index and the best-case forecasts' positioning relative to the initial model-based forecasts. Specifically, giving scenarios appears to boost confidence ($F_{1,58} = 4.43$, $p = .04$). Also, when scenarios are not given, composite best-case forecasts end up being significantly further away from the model-based forecasts than those of experimental groups given scenarios ($F_{1,58} = 6.91$, $p = .011$).

4.6. Consensus vs composite forecasts

In this study, composite forecasts could be directly compared to the consensus forecasts, since the participants providing independent individual forecasts in Phase 1 are asked to work together to arrive at consensus forecasts in Phase 2. Although the patterns in the composite forecasts seem to be very similar to those of the consensus predictions, a further analysis of comparing consensus forecasts directly with the composite forecasts may provide insights into the effects (if any) of group discussions. Such direct comparisons could be made via *percentage difference* scores of consensus forecasts from the respective composite predictions for each of the five performance measures. For instance, in point forecasts:

$$\%\text{Difference Consensus Composite} = 100 \times \frac{(\text{Consensus point forecast} - \text{Composite point forecast})}{\text{Composite point forecast}}$$

Other performance measures will have similar computations for respective percentage difference scores.

Table 4 displays the consensus vs composite comparisons as analyzed via these percentage difference scores. The comparisons of groups with and without scenarios are displayed in Table 4A, while the detailed breakdowns for each experimental group are given in Table 4B. Significant differences among experimental groups with and without scenarios only appear in percentage difference scores for SI and ASR ($F_{1,58} = 11.66$, $p = .001$ and $F_{1,58} = 4.34$, $p = .042$, respectively). In terms of the Surprise Index, this suggests that the presence of scenarios leads to higher confidence (lower mean values of surprise index) expressed in consensus assessments when compared with the relatively lower confidence (higher mean values of surprise index) in statistical averaging of individual forecasts (i.e., composite predictions). There is also a significant effect of working in a group (i.e., consensus predictions) on the asymmetry of intervals (as bound by the best-case and worst-case forecasts) with respect to the position of the point forecasts. When no scenarios are provided, intervals appear to be similarly symmetric for consensus and composite forecasts. However, when scenarios are given, level of asymmetry in consensus intervals is escalated as compared to intervals computed by statistical averaging of individual predictions.

Furthermore, when members receive divergent scenarios (i.e., X2 participants with one member receiving best-case while the other getting worst-case scenario), consensus forecasts appear to be smaller than composite forecasts for their point and best-case predictions, while being almost identical for worst-case forecasts, as can be seen from Table 4B. This is totally the opposite when members receive identical sets of scenarios (i.e., X3 participants with *both* members receiving *both* best-case and worst-case scenarios). In these cases, consensus forecasts appear to be higher than the composite predictions, signaling larger adjustments from initial forecast advice in consensus as opposed to averaged individual forecasts. These differences between dyads with members receiving different versus identical scenarios are all statistically significant ($t_{39} = -3.58$, $p = .001$ for point forecasts; $t_{35} = -3.20$, $p = .003$ for best-case forecasts; and $t_{38} = -2.24$, $p = .031$ for worst-case forecasts).

Table 4
Mean percentage difference scores between consensus and composite forecasts (number of dyads within each category is given in parenthesis).

	% Difference between consensus and composite point forecasts	% Difference between consensus and composite best-case forecasts	% Difference between consensus and composite worst-case forecasts	% Difference between consensus and composite Surprise Index	% Difference between consensus and composite Asymmetry Ratio
A. Mean percentage difference scores for experimental groups with and without scenarios					
Experimental groups not given any scenarios (X1)	0.82% (18)	1.02% (18)	2.65% (18)	13.41% (18)	1.21% (18)
Experimental groups given scenarios (X2 & X3)	0.94% (42)	− 1.17% (42)	2.40% (42)	− 14.96% (42)	12.73% (42)
B. Mean percentage difference score breakdowns for each experimental group					
No scenarios (X1)	0.82% (18)	1.02% (18)	2.65% (18)	13.41% (18)	1.21% (18)
One member received only best-case scenario; other member received only worst-case scenario (X2)	− 1.24% (24)	− 3.37% (24)	− 0.01% (24)	− 19.37% (24)	15.37% (24)
Both members received both scenarios (X3)	3.84% (18)	1.77% (18)	5.62% (18)	− 9.09% (18)	9.22% (18)

4.7. Effects of scenario framing

The *framing-manipulation* study focused on examining how sensitive the above results would be to the particular labeling of scenarios as *best-case* and *worst-case*. The findings showed that the labels attached to scenarios made no significant difference in participants' perceptions of scenario informativeness, usefulness, and influence (all $p > .10$), which were all rated highly again. Even though the participants' stated assessments of the scenarios were the same, those receiving unlabeled scenarios reacted differently to this information as compared to the participants who were given the best-case and worst-case labels.

In particular, the participants given *best-case scenarios* set their individual worst-case forecasts closer to the model-based point predictions as compared to those participants given the same scenarios without labeling (-10% vs -20% change from given point prediction, respectively; $F_{1,37} = 11.84$, $p = .001$). Similarly, the participants given *worst-case scenarios* set their individual best-case forecasts closer to the given model-based predictions as compared to those participants given the same scenarios without labeling ($+13\%$ vs $+35\%$ change from given point prediction, respectively; $F_{1,37} = 6.67$, $p = .014$). These participants also adjusted their point forecasts more (and in a negative direction) than the individuals given unlabelled scenarios (-7% vs $+2\%$ change from given point prediction, respectively; $F_{1,37} = 5.41$, $p = .026$).

Consensus forecasts reflected similar effects. The dyads given labeled scenarios (where one member received a 'best-case' while the other received a 'worst-case' scenario) made consensus best-case forecasts closer to the given model-based predictions as compared to those dyads given the same scenarios without labels ($+17\%$ vs $+32\%$ change from given point prediction adjustment, respectively; $F_{1,37} = 6.80$, $p = .013$). Interestingly, those dyads given labeled scenarios adjusted their point forecasts much less than the dyads given unlabeled scenarios (0% vs $+5\%$ change from given point prediction, respectively; $F_{1,37} = 4.67$, $p = .037$).

Comparisons with participants who did not receive any scenarios showed no significant differences in forecast adjustments for either the individuals or the groups. However, being given a scenario (even though unlabelled) appeared to increase confidence (for individual forecasts: mean SI values of 9% vs 16%, for unlabeled scenarios vs no scenarios, respectively; $F_{1,64} = 6.09$, $p = .016$; for consensus forecasts: mean SI values of 7% vs 17%, respectively; $F_{1,31} = 9.71$, $p = .004$). These findings provide evidence for framing effects. In particular, unlabeled scenarios do not seem to enhance the acceptance of forecast advice (unlike the labeled scenarios which raise advice acceptance, as shown by smaller adjustments to given forecasts), while still boosting confidence. Such results highlight the need for further work on structuring the scenarios so as to challenge the users' assumptions and increase their risk-awareness while deterring from overly optimistic planning and confirmation thinking.

5. Discussion

This study presented a two-phased experimental setting to explore the potential impact of integrating scenarios into forecast advice. In addition to examining the effects of giving single versus multiple scenarios to forecast users, the study focused on investigating the differences between the forecast adjustments made by small-groups and the forecast adjustments made independently by individuals.

Our findings reveal that, overall, individual users tend to make larger adjustments to given model-based forecasts when scenarios are not available. When compared to providing only forecast values (with no accompanying scenarios), incorporating scenarios into forecast information (with best-case and worst-case labels) appears to increase the acceptance of forecast advice while strengthening the users' confidence in their prediction intervals. This could be due to the priming effect of the labels

attached to scenarios, or it could be because the plausible futures communicated by scenarios may be effective in increasing the salience of their predictions.

For individual forecasts, providing a best-case or worst-case scenario seems to have a significant effect on the forecasts at the other extreme. When a best-case scenario is presented, the final worst-case forecasts are significantly closer to the model-based point forecasts than the cases when no best-case scenarios are presented. Similarly, individuals receiving worst-case scenarios appear to generate best-case forecasts that are significantly closer to the model-based point forecasts than those not given worst-case scenarios. It looks as if though the given best-case and worst-case scenarios impose an adjustment limit on the opposite predictions (i.e., worst-case and best-case forecasts, respectively), creating a balancing effect. One possible explanation could be that participants receiving a particular scenario spend most of their time debating the implications of this scenario, and thus, not devoting enough time to think through the opposite case. An alternative explanation may be that providing one of the extreme case scenarios may prompt different biases (as opposed to giving scenarios for both extremes). In either case, this finding might have important ramifications for organizations suffering from optimism [70–73] or desirability bias [74–77] in their forecasting processes. In these organizations, giving worst-case scenarios as part of forecast advice to overly-optimistic decision makers may actually prevent excessively high predictions. In fact, any organization tackling with the implications of unwarrantedly low/high forecasts may benefit from such forecast support structured to alleviate optimistic/pessimistic tendencies that can easily be triggered by communication choices.

Another significant impact of scenarios is evident in the forecast users' confidence as measured by their surprise probabilities, which also reveal differences according to the type of scenarios provided. Those given only best-case scenarios appear to have the highest confidence, closely followed by those who received both scenarios. When scenarios are given without labels, they are still effective in boosting confidence, confirming previous findings that any type of scenario information increases confidence in forecasts [1]. One potential explanation could be that the presence of scenarios in addition to mere numbers as forecasts encourages participants to commit more to their final forecasts, which in turn translates to higher confidence. This would support the findings from sense-making research in that stories (as provided by scenarios) may promote commitment [78], potentially affecting confidence [79]. An alternative explanation could be that the format used to elicit forecasts (best-case forecasts and worst-case forecasts in addition to point forecasts) helps elevate such tendencies. When coupled with providing scenarios labelled as best/worst-case, the choice of forecasting format may further augment the commonly encountered overconfidence. Further research on alternative formats both for eliciting predictions and for giving forecast advice will prove crucial in designing improved communication structures to enhance the decision-making value of forecasts. Similarly, future work could explore possible sequencing effects of forecast and scenario presentation, as well as potential confirmation bias and anchoring effects by asking participants to give their initial expectations prior to giving any forecast advice. Results could provide significant input into designing more efficient forecasting processes in organizations.

Turning to consensus forecasts, we find support that dyads also utilize scenarios as forecast advice. An overall comparison of dyads with and without scenarios shows significant differences in group adjustment patterns for best-case forecasts as well as in confidence (as indexed via SI). When scenarios are given, best-case forecasts are arrived at via smaller adjustments to model-based forecasts while confidence shows a significant increase. That is, when scenarios are not given, consensus best-case forecasts turn out to be much higher than the initial model-based forecasts accompanied by lower confidence. Also, there are no significant differences between dyads where each member receives a different scenario versus where each member receives both scenarios. This finding may be interpreted as demonstrating the effectiveness of group discussions in communicating the scenario information even in cases where group members receive disparate scenarios.

Comparisons of consensus forecasts made by dyads with the averages of individual group member's forecasts (i.e., composite forecast) reveal additional insights. When dyad members with deviating (i.e., one with best-case and the other with worst-case) scenarios work together, the group process seems to result in consensus forecasts where the model-based forecasts are adjusted less than those of the average of independent individual forecasts. This finding may be interpreted as indicating active information exchange between the dyad members. That is, when working on their own and not knowing the other scenario, individuals appear to make larger adjustments to model-based forecasts to arrive at their final forecasts. After learning about each other's scenarios, they tend to rely more on the given model-based forecasts, hence the significant reductions in the adjustments.

Interestingly, identical sets of information available to dyad members at the outset (i.e., *both* have *both* best- and worst-case scenarios) result in a completely opposite state, where consensus forecasts are adjusted more than the composite predictions. This signals that forming a small group where the members already possess identical sets of information on two extremes (i.e., best vs worst case) leads to a decreased usage of scenario-based advice, resulting in more extreme consensus forecasts. While individuals might make lower adjustments to given model forecasts, group discussion focused on both sets of scenarios appears to result in much more liberally adjusted forecasts as a dyad. While individual forecasts place total responsibility on one person, consensus forecasts involving joint work could lead to a diffusion of responsibility with less individual accountability, resulting in increased motivation to make more radical adjustments. Alternatively, when both members receive both scenarios, group discussions may intensify the visualization of more extreme cases, leading to aggravated frames as compared to more restricted individual frames. This is also supported by the finding that the presence of scenarios leads to higher confidence in consensus forecasts than the relatively lower confidence levels in the composite forecasts.

The findings also suggest that the way the scenarios are phrased and presented (e.g., by the "best-case" and "worst-case" labels) may play an important role in how they are perceived and used as forecast advice, confirming the pervasiveness of such 'framing' effects shown in a variety of decision situations (e.g., [80–82]). As discussed above, our results indicate that giving

best-case scenarios may induce a strong 'optimism' bias for individuals; while a 'pessimism' bias may be provoked via worst-case scenarios. However, findings from the manipulation study show that when labels are removed from the same scenarios, evidence towards optimistic and pessimistic tendencies seem to disappear. Hence, the particular wording used in communicating scenarios as forecast aids appears to be a critical choice that may in turn be used to challenge mental frames and tunnel vision in managerial decision making [24,83], potentially preventing decision makers from drawing unjustified conclusions based on the scenarios. Future work on using scenarios as forecast advice would benefit from studying such framing effects along with their potential ramifications as debiasing tools against hindsight bias and confirmation bias [25], overconfidence [24] and illusion of control [20]. Similarly, multiple scenarios could be effectively employed to counteract 'future myopia' [84], as well as to overcome organizational deficiencies in utilizing internal *vs* external information [85] and 'retrospective sensemaking' [86]. Our findings provide encouraging suggestions for further work in these arenas to help design effective forecast management systems in company settings.

6. Conclusions

Scenario thinking is highly critical for organizational learning as well as for helping managers plan for the future [25,87]. This paper presents an exploratory attempt at incorporating scenarios into forecast communication and advice taking processes via giving scenarios to users as additional forecast information. Our findings suggest evidence of the important role played by scenarios in encouraging forecasts users to consider alternative outcomes, thus strengthening the forecast message.

The current study offers an initial step in examining how scenarios could be used as channels of forecast advice. It needs to be extended to assess the forecasting performance of practitioners in diverse organizational contexts with different group sizes and group processes. Furthermore, comparisons of situations where scenarios are constructed by the forecasters themselves (either individually or in groups) with cases where the scenarios are generated by domain experts and supplied as forecast advice would be highly informative. Such work would also benefit from studying the interactions between different facets of scenario construction and communication processes (e.g., increasing transparency in constructing scenarios, including consistency checks to alleviate biases, enhancing credibility of scenarios, fostering trust via scenario-sharing, etc.). Overall, the lens of storytelling promises to provide a rich platform to examine the prolific role that scenarios can play in improving forecast communication and predictive accuracy. Hence, future research exploring the dynamic interplays of scenario use with forecast expectations, feedback systems, group composition, and organizational culture will prove essential in enhancing forecast management and process improvement in organizations.

Acknowledgements

Appendix A. Background information on the case company

De Facto Company

De Facto Company manufactures mobile phones, which are sold through specialized retailers as well as large supermarket chains. *De Facto* has various models in four main product categories: tablets, smart phones with android operating systems, mobile phones with many complex features (e.g. WiFi, GPS, premium cameras, etc.), and basic mobiles. Competition in this industry is very harsh in general, particularly in the first three categories. One of the major costs is related to research and development (R&D) for new products and innovations. As new technologies are being constantly introduced to the market, speed is essential for remaining competitive in today's digital era. Moreover, as new products are introduced into the market very quickly, existing ones have a very short span of life. If not sold, such 'aged' products increase the overall costs drastically. These factors are not that significant for the basic mobiles category, as in this market, products do not change quickly. It is a 'niche' market, where main customers are elderly people and parents of young children who want to give their children a basic mobile. Both groups' needs are the same: a simple mobile with no fancy features but only to be used for mobile contact when necessary. Finally, for all product groups, there are other potential costs to be minimized: firstly, the cost of over-production and the associated storage costs; and secondly, the cost of lost sales for to running out of stock as a result of under-production. Combined together, these factors amplify the importance of forecasting the demand correctly (with minimal error) for *De Facto Company*'s various products.
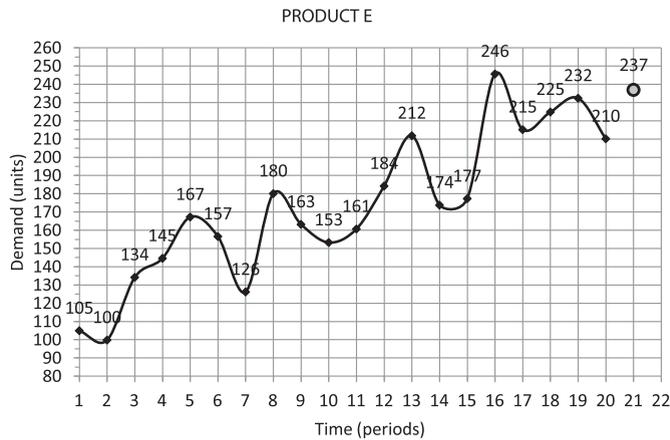
In order to improve the accuracy of its forecasts, *De Facto* employs a combination approach that involves using the scenario method. One-month-ahead demand forecasts are generated at the beginning of each month for the following month (e.g., September demand forecasts for each of the products are constructed at the beginning of August). This process works as follows:

1. Statistical forecasts are produced using the computerized forecasting system
2. "Worst case" and "best case" scenarios for each product's demand in the following period are developed by a team of experts in the company using all the information available. The basic rule in preparing and using these scenarios is that every scenario is equally plausible — these scenarios are constructed to depict 'best' and 'worst' plausible futures that need to be considered and prepared for regarding each product's potential demand for the next period.
3. The scenarios (generated by the expert team) and model forecasts (constructed by statistical models) are given to executives in charge of making forecasts for individual consideration, followed by group discussion. These people firstly produce their

individual forecasts by looking at the scenarios and model forecasts. Then they form groups of two and come up with consensus forecasts for typical, best and worst cases after discussions in light of the scenario(s) and the model forecast provided for each product. Your task in this study replicates this process.

Appendix B. Sample form for individual forecasts — only best-case scenario given (G2)

PRODUCT E

Model-based forecast for period 21: **237**

Best-case scenario:

This is one of our best performing mobile phones with complex features. Its built-in camera is a premium one, plus it has got all of the other usual features (e.g. GPS, music player, Bluetooth, etc.). It was a hit product when first introduced, and kept its upward trend ever since. With new features added from time to time (for instance, new colors, improved camera, etc.), its popularity has been kept above a certain level. With a price very reasonably positioned among that of its major competitors, Product E is reported to be a desirability object particularly for the teenager consumer group. The future looks very bright for this product. It is expected to carry on its upward trend, even to the levels that look overly optimistic at the moment.

YOUR FORECASTS:

Please provide your *point forecast* for period 21: ………………
Please provide your *best-case forecast*(highest value you predict) for period 21: ………………
Please provide your *worst-case forecast*(lowest value you predict) for period 21: ………………
What is the probability that the *actual demand for period 21* will be a **surprise** to you?
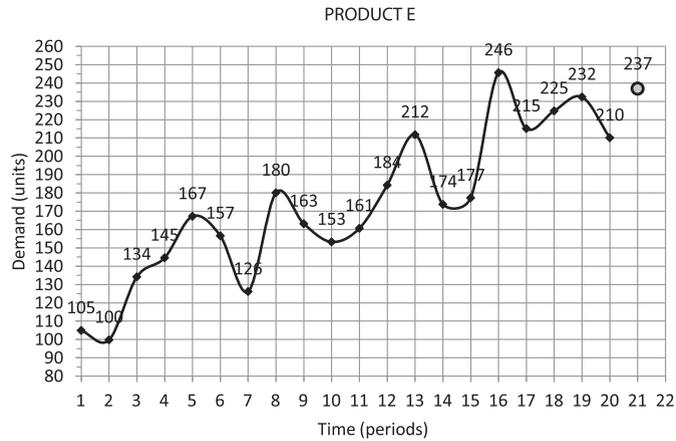(What is the probability that the *actual demand for period 21* will be **higher** than your best-case forecast or **lower** than your worst-case forecast?):
………………%

In *constructing my forecasts*, the given **best-case scenario** was

Not informative at all (provided no information at all)				**Extremely informative** (provided very significant information)
1	2	3	4	5
Not useful at all (was not helpful at all in making my forecasts)				**Extremely useful** (was extremely helpful in making my forecasts)
1	2	3	4	5
Not influential at all (had no effect whatsoever on my forecasts)				**Extremely influential** (had a very significant effect on my forecasts)
1	2	3	4	5

Appendix C. Sample form for consensus forecasts ––member received only best-case scenario (X2)

PRODUCT E

PRODUCT E

Model-based forecast for period 21: **237**

Best-case scenario:

This is one of our best performing mobile phones with complex features. Its built-in camera is a premium one, plus it has got all of the other usual features (e.g. GPS, music player, Bluetooth, etc.). It was a hit product when first introduced, and kept its upward trend ever since. With new features added from time to time (for instance, new colors, improved camera, etc.), its popularity has been kept above a certain level. With a price very reasonably positioned among that of its major competitors, Product E is reported to be a desirability object particularly for the teenager consumer group. The future looks very bright for this product. It is expected to carry on its upward trend, even to the levels that look overly optimistic at the moment.

CONSENSUS FORECASTS FOR YOUR GROUP

Please provide your group's **consensus** *point forecast* for period 21: ....
Please provide your group's **consensus** *best-case forecast(highest value predicted)* for period 21: ....
Please provide your group's **consensus** *worst-case forecast(lowest value predicted)* for period 21: ....
What is the probability that the *actual demand for period 21* will be a **surprise** to your group?
(What is the probability that the *actual demand for period 21* will be **higher** than your group's consensus best-case forecast or **lower** than your group's consensus worst-case forecast?):...............%

Please provide your level of agreement with your group's **consensus POINT forecast**

Totally disagree				Totally agree
1	2	3	4	5

Please provide your level of agreement with your group's **consensus BEST-CASE forecast**

Totally disagree				Totally agree
1	2	3	4	5

Please provide your level of agreement with your group's **consensus WORST-CASE forecast**

Totally disagree				Totally agree
1	2	3	4	5

References

[1] K.M. Kuhn, J.A. Sniezek, Confidence and uncertainty in judgmental forecasting: differential effects of scenario presentation, J. Behav. Decis. Mak. 9 (1996) 231–247.
[2] R. Fildes, P. Goodwin, M. Lawrence, K. Nikolopoulos, Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning, Int. J. Forecast. 25 (2009) 3–23.
[3] P. Goodwin, D. Önkal, M. Lawrence, Improving the Role of Judgment in Economic Forecasting, in: M.P. Clements, D.F. Hendry (Eds.), The Oxford Handbook of Economic Forecasting, Oxford University Press, Oxford, 2011, pp. 163–189.
[4] M. Lawrence, P. Goodwin, M. O'Connor, D. Önkal, Judgemental forecasting: a review of progress over the last 25 years, Int. J. Forecast. 22 (2006) 493–518.
[5] N.R. Sanders, K.B. Manrodt, The efficacy of using judgmental versus quantitative forecasting methods in practice, Omega Int. J. Manag. Sci. 31 (2003) 511–522.
[6] D.W. Bunn, G. Wright, Interaction of judgemental and statistical forecasting methods: issues and analysis, Manag. Sci. 32 (1991) 501–518.
[7] M.S. Gönül, D. Önkal, P. Goodwin, Expectations, use and judgmental adjustment of external financial and economic forecasts: an empirical investigation, J. Forecast. 28 (2009) 19–37.
[8] D. Önkal, M.S. Gönül, M. Lawrence, Judgmental adjustments of previously-adjusted forecasts, Decis. Sci. 39 (2008) 213–238.
[9] D. Önkal, P. Goodwin, M. Thomson, M.S. Gönül, A. Pollock, The relative influence of advice from human experts and statistical methods on forecast adjustments, J. Behav. Decis. Mak. 22 (2009) 390–409.
[10] N. Harvey, I. Fischer, Accepting help, improving judgment, and sharing responsibility, Organ. Behav. Hum. Decis. Process. 70 (2) (1997) 117–133.
[11] N. Harvey, C. Harries, I. Fischer, Using advice and assessing its quality, Organ. Behav. Hum. Decis. Process. 81 (2) (2000) 252–273.
[12] I. Yaniv, E. Kleinberger, Advice taking in decision making: egocentric discounting and reputation formation, Organ. Behav. Hum. Decis. Process. 83 (2) (2000) 260–281.
[13] N. Harvey, C. Harries, Effects of judges' forecasting on their later combination of forecasts for the same outcomes, Int. J. Forecast. 20 (2004) 391–409.
[14] S. Bonaccio, R.S. Dalal, Advice taking and decision-making: an integrative literature review, and implications for the organizational sciences, Organ. Behav. Hum. Decis. Process. 101 (2) (2006) 127–151.
[15] I. Yaniv, Receiving other people's advice: influence and benefit, Organ. Behav. Hum. Decis. Process. 93 (2004) 1–13.
[16] I. Yaniv, The benefit of additional opinions, Curr. Dir. Psychol. Sci. 13 (2) (2004) 75–78.
[17] T. Comes, N. Wijngaards, J. Maule, D. Allen, F. Schultmann, Scenario reliability assessment to support decision makers in situations of severe uncertainty, in: IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, 2012, pp. 30–37.
[18] S. French, A.J. Maule, G. Mythen, Soft modelling in risk communication and management: examples in handling food risk, J. Oper. Res. Soc. 56 (2004) 879–888.
[19] P. Goodwin, G. Wright, Enhancing strategy evaluation in scenario planning: a role for decision analysis, J. Manag. Stud. 38 (2001) 1–16.
[20] G.P. Hodgkinson, N.J. Bown, A.J. Maule, K.W. Glaister, A.D. Pearman, Breaking the frame: an analysis of strategic cognition and decision making under uncertainty, Strateg. Manag. J. 20 (1999) 977–985.
[21] G.P. Hodgkinson, A.J. Maule, N.J. Bown, A.D. Pearman, K.W. Glaister, Further reflections on the elimination of framing bias in strategic decision making, Strateg. Manag. J. 23 (2002) 1069–1076.
[22] W. Walker, P. Harremoës, J. Rotmans, J.P. van der Sluijs, M.B.A. van Asselt, P. Janssen, M.P. Krayer von Krauss, Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support, Integr. Assess. 4 (2003) 5–17.
[23] G. Wright, P. Goodwin, Future-focussed thinking: combining scenario planning with decision analysis, J. Multi-Criteria Decis. Anal. 8 (1999) 311–321.
[24] G. Wright, P. Goodwin, Decision making and planning under low levels of predictability: enhancing the scenario method, Int. J. Forecast. 25 (2009) 813–825.
[25] G. Wright, G. Cairns, Scenario Thinking: Practical Approaches to the Future, Palgrave Macmillan, Hampshire, 2011.
[26] D. Önkal, K.Z. Sayım, M. Lawrence, Wisdom of group forecasts: does role-playing play a role? Omega Int. J. Manag. Sci. 40 (2012) 693–702.
[27] M. Godet, From forecasting to 'La Prospective' a new way of looking at futures, J. Forecast. 1 (1982) 293–301.
[28] H. Kahn, A.J. Weiner, The Year 2000, Macmillan, New York, 1967.
[29] S.P. Schnaars, M.T. Topol, The use of multiple scenarios in sales forecasting. An empirical test, Int. J. Forecast. 3 (3–4) (1987) 405–419.
[30] P.J.H. Schoemaker, When and how to use scenario planning: a heuristic approach with illustration, J. Forecast. 10 (6) (1991) 549–564.
[31] P. Goodwin, G. Wright, The limits of forecasting methods in anticipating rare events, Technol. Forecast. Soc. Chang. 77 (3) (2010) 355–368.
[32] S.E. Taylor, S.C. Thompson, Stalking the elusive vividness effect, Psychol. Rev. 89 (1982) 155–181.
[33] C.A. Anderson, Abstract and concrete data in the perseverance of social theories: when weak data lead to unshakeable beliefs, J. Exp. Soc. Psychol. 19 (1983) 93–108.
[34] W.L. Gregory, R.B. Cialdini, K. Carpenter, Self-relevant scenarios as mediators of likelihood estimates and compliance: does imagining make it so? J. Personal. Soc. Psychol. 43 (1982) 89–99.
[35] R.D. Zentner, Scenarios: past, present and future, Long Range Plan. 15 (1982) 12–20.
[36] R.D. Zentner, Scenarios in forecasting, Chem. Eng. News (October 1975) 22–34.
[37] I.H. Wilson, Scenarios, in: J. Fowles (Ed.), Handbook of Futures Research, Greenwold Press, Westport, CT, 1978, pp. 22–47.
[38] K. van der Heijden, Scenarios: The Art of Strategic Conversion, Wiley, Chichester, 1996.
[39] D.W. Bunn, A.A. Salo, Forecasting with scenarios, Eur. J. Oper. Res. 68 (3) (1993) 291–303.
[40] A. Tversky, D. Kahneman, Judgment under uncertainty: heuristics and biases, Science 185 (1974) 1124–1131.
[41] In: D. Kahneman, P. Slovic, A. Tversky (Eds.), Judgment Under Uncertainty: Heuristics and Biases, Cambridge University Press, New York, 1982.
[42] J. Baron, Thinking and Deciding, 4th ed. Cambridge University Press, New York, 2007.
[43] T. Gilovich, D. Griffin, Heuristics and Biases: Then and Now, in: T. Gilovich, D. Griffin, D. Kahneman (Eds.), Heuristics and Biases: The Psychology of Intuitive Judgment, Cambridge University Press, Cambridge, 2002, pp. 1–4.
[44] W.L. Gregory, A. Duran, Scenarios and Acceptance of Forecasts, in: J.S. Armstrong (Ed.), Principles of Forecasting: A Handbook for Researchers and Practitioners, Kluwer Academic Publishers, Boston, 2001, pp. 519–540.
[45] P.J.H. Schoemaker, Forecasting and Scenario Planning: The Challenges of Uncertainty and Complexity, in: D.J. Koehler, N. Harvey (Eds.), Blackwell Handbook of Judgment and Decision Making, Blackwell, Oxford, 2004, pp. 247–296.
[46] G. Wright, G. Cairns, P. Goodwin, Teaching scenario planning: lessons from practice in academe and business, Eur. J. Oper. Res. 194 (2009) 323–335.
[47] G. Wright, G. Rowe, Group-based judgmental forecasting: an integration of extant knowledge and the development of priorities for a new research agenda, Int. J. Forecast. 27 (2011) 1–13.
[48] P.J.H. Schoemaker, Multiple scenario development: its conceptual and behavioral foundation, Strateg. Manag. J. 14 (1993) 193–213.
[49] P. Bishop, A. Hines, T. Collins, The current state of scenario development: an overview of techniques, Futures 9 (1) (2007) 5–25.
[50] J. Gooty, F.J. Yammarino, Dyads in organizational research: conceptual issues and multilevel analyses, Organ. Res. Methods 14 (3) (2011) 456–483.
[51] I. Steiner, Group Processes and Productivity, Academic Press, New York, 1972.
[52] R. Panko, Managerial communication patterns, J. Organ. Comput. (1992) 95–122.
[53] H. Topi, J.S. Valacich, M.T. Rao, The effects of personality and media differences on the performance of dyads addressing a cognitive conflict task, Small Group Res. 33 (6) (2002) 667–701.
[54] C.A. Schriesheim, S.L. Castro, X. Zhou, F.J. Yammarino, The folly of theorizing "A" but testing "B". A selective level-of-analysis review of the field and detailed leader-member exchange illustration, Leadersh. Q. 12 (2001) 515–551.
[55] T.L. Boles, R.T.A. Croson, J.K. Murnighan, Deception and retribution in repeated ultimatum bargaining, Organ. Behav. Hum. Decis. Process. 83 (2000) 235–259.

[56] H.H. Blumberg, On taking too much: a point accumulation procedure for comparing mutual agreements with controlled individual initiatives, Small Group Res. 28 (1997) 171–193.
[57] B.H. Drake, D.J. Moberg, Communicating influence attempts in dyads: linguistic sedatives and palliatives, Acad. Manag. Rev. 11 (1986) 567–584.
[58] S.N. Fraidin, When one head is better than two? Interdependent information in group decision making, Organ. Behav. Hum. Decis. Process. 93 (2004) 102–113.
[59] N.S. Hill, K.A. Bartol, P.E. Tesluk, G.A. Langa, Organizational context and face-to-face interaction: influences in the development of trust and collaborative behaviours in computer-mediated groups, Organ. Behav. Hum. Decis. Process. 108 (2009) 187–201.
[60] L. Zhou, D. Zhang, A comparison of deception behaviour in dyad and triadic group decision making in synchronous computer-mediated communication, Small Group Research 37 (2) (2006) 140–164.
[61] C.M. Allwood, P.A. Granhag, Realism in confidence judgments as a function of working in dyads or alone, Organ. Behav. Hum. Decis. Process. 66 (1996) 277–289.
[62] A. Koriat, Alleviating inflation of conditional predictions, Organ. Behav. Hum. Decis. Process. 106 (2008) 61–76.
[63] P.E. Tetlock, The Impact of Accountability on Judgment and Choice: Toward a Social Contingency Model, in: M.Zanna(Ed.), Advances in Experimental Social Psychology (Vol. 25), Academic Press, New York, 1992, p.331–376.
[64] M.S. Gönül, D. Önkal, M. Lawrence, The effects of structural characteristics of explanations on use of a DSS, Decis. Support. Syst. 42 (2006) 1481–1493.
[65] V.A. Bañuls, M. Turoff, Scenario construction via Delphi and cross-impact analysis, Technol. Forecast. Soc. Chang. 78 (2011) 1579–1602.
[66] T.J. Chermack, Studying scenario planning: theory, research suggestions, and hypotheses, Technol. Forecast. Soc. Chang. 72 (2005) 59–73.
[67] M. O'Connor, W. Remus, K. Griggs, The asymmetry of judgemental confidence intervals in time series forecasting, Int. J. Forecast. 17 (4) (2001) 623–633.
[68] J. Surowiecki, The Wisdom of Crowds, Anchor Books, New York, 2005.
[69] D. Önkal, M. Lawrence, K.Z. Sayım, Influence of differentiated roles on group forecasting accuracy, Int. J. Forecast. 27 (2011) 50–68.
[70] D. Lovallo, D. Kahneman, Delusions of success: how optimism undermines executives' decisions, Harv. Bus. Rev. (July 2003) 56–63.
[71] G. Tichy, The over-optimism among experts in assessment and foresight, Technol. Forecast. Soc. Chang. 71 (4) (2004) 341–363.
[72] N.D. Weinstein, Unrealistic optimism about future life events, J. Personal. Soc. Psychol. 39 (5) (1980) 806–820.
[73] T. Sharot, The Optimism Bias, Constable & Robinson, London, 2012.
[74] D.V. Budescu, M. Bruderman, The relationship between the illusion of control and the desirability bias, J. Behav. Decis. Mak. 8 (2) (1995) 109–125.
[75] P. Ecken, T. Gnatzy, H.A. von der Gracht, Desirability bias in foresight: consequences for decision quality based on Delphi results, Technol. Forecast. Soc. Chang. 78 (2011) 1654–1670.
[76] R.A. Olsen, Desirability bias among professional investment managers: some evidence from experts, J. Behav. Decis. Mak. 10 (1) (1997) 65–72.
[77] P.D. Windschitl, A.R. Smith, J.P. Rose, Z. Krizan, The desirability bias in predictions: going optimistic without leaving realism, Organ. Behav. Hum. Decis. Process. 111 (1) (2010) 33–47.
[78] K.E. Weick, Sensemaking in Organizations, Sage Publications, California, 1995.
[79] C.R. Schwenk, Information, cognitive biases and commitment to a course of action, Acad. Manag. Rev. 11 (1986) 298–310.
[80] D. Kahneman, A. Tversky, Choices, values, and frames, Am. Psychol. 39 (4) (1984) 341–350.
[81] B.J. McNeil, S.G. Pauker, H.C. Sox Jr., A. Tversky, On the elicitation of preferences for alternative therapies, N. Engl. J. Med. 306 (21) (1982) 1259–1262.
[82] R.H. Thaler, Toward a positive theory of consumer choice, J. Econ. Behav. Organ. 1 (1980) 39–60.
[83] S.S. Posavac, F.R. Kardes, J.J. Brakus, Focus induced tunnel vision in managerial judgment and decision making: the peril and the antidote, Organ. Behav. Hum. Decis. Process. 113 (2010) 102–111.
[84] F.A. O'Brien, Scenario planning––lessons for practice from teaching and learning, Eur. J. Oper. Res. 152 (3) (2004) 709–722.
[85] R. Durand, Predicting a firm's forecasting ability: the roles of organizational illusion of control and organizational attention, Strateg. Manag. J. 24 (2003) 821–838.
[86] K.E. Weick, K.M. Sutcliffe, D. Obstfeld, Organizing and the process of sensemaking, Organ. Sci. 16 (2005) 409–421.
[87] K. van der Heijden, R. Bradfield, G. Burt, G. Cairns, G. Wright, The Sixth Sense: Accelerating Organizational Learning with Scenarios, Wiley, Chichester, 2002.

Dilek Önkal is Professor of Decision Sciences and Dean of the Faculty of Economics, Administrative and Social Sciences at Bilkent University, Turkey. She received a Ph.D. in Decision Sciences from the University of Minnesota. Her research focuses on judgmental forecasting, judgment and decision making, forecasting/decision support systems, risk perception and risk communication with a strong emphasis on multi-disciplinary interactions. Her work has appeared in journals such as *Journal of Behavioral Decision Making, Organizational Behavior and Human Decision Processes, Decision Sciences, Risk Analysis, Decision Support Systems, International Journal of Forecasting*, and *Omega: The International Journal of Management Science*.

Kadire Zeynep Sayım is Assistant Professor in the Faculty of Business Administration at Bilkent University, Ankara, Turkey. Her main research interests lie in the areas of employment relations in multinationals, group dynamics and role playing in decision making, and gender diversity in management. Her work has appeared in such journals as *Human Relations, International Journal of Human Resource Management, International Journal of Forecasting*, and *Omega: The International Journal of Management Science*.

M. Sinan Gönül is an Associate Professor in the Department of Business Administration at Middle East Technical University, Turkey. His research in judgmental forecasting and forecasting support systems has been published in various journals including *Journal of Behavioral Decision Making, Decision Sciences, Decision Support Systems* and the *Journal of Forecasting.*