Intelligent collective: some issues with collective cardinality
Van Du Nguyen\(^a\) and Mercedes G. Merayo\(^b\)
\(^a\)Department of Information Systems, Faculty of Computer Science and Management, Wroclaw University of Science and Technology, Wroclaw, Poland; \(^b\)Dept. Sistemas Informáticos y Computación, Universidad Complutense de Madrid, Madrid, Spain

**ABSTRACT**
In general, a common approach to enhance the diversity, an important criterion of intelligent collectives, is to enlarge the collective cardinality. The objective of this study was to analyse some issues with collective cardinality in the effectiveness of collective performance. For this aim, first, we investigate the impact of collective cardinality on collective performance. Subsequently, a comparison among collective performance measures is presented. The simulation results have qualified the positive impact of collective cardinality on the effectiveness of collective performance. Moreover, some issues on processing large collectives; some further challenges of the impact of diversity on collective prediction are discussed.

**ARTICLE HISTORY**
Received 10 December 2016
Accepted 5 April 2017

**KEYWORDS**
Collective prediction; wisdom of crowds; integration computing

**Introduction**
Recently, a collective of individuals has been proved to be superior to single individuals in solving a range of difficult judgment and prediction problems (Galton, 1907; Graefe & Armstrong, 2011; Surowiecki, 2005). Even though, in Lang, Bharadwaj, and Di Benedetto (2016), the experiment results have indicated that the collective predictions, which are determined on the basis of individuals predictions, are better than those generated by the traditional forecasting methods. In fact, from the practical point of view, one of the challenges of crowd-based applications is choosing either a small collective with a large amount of information or a large collective with a small amount of information (Sunstein, 2006; Surowiecki, 2005). The popular approach of using a small collective with more information is Delphi method, in which a panel of members who are more knowledgeable and well-informed are asked for giving their opinions on a given problem such as the outcome of a future event (Gaur, Kesavan, Raman, & Fisher, 2007; Graefe & Armstrong, 2011). However, there also exist some potential issues with this approach such as a small number of experts invited and costly issues. Furthermore, it violates the main criteria of which collectives to be intelligent, particularly the diversity of knowledge, perspectives and the independence in making decisions without influencing of others (Surowiecki, 2005). Conversely, the latter is often understood as collectives involving a large number of members that are uninformed or irrational. Several publications have appeared in
recent years documenting this capability of collectives as in Graefe and Armstrong (2011), Lang et al. (2016), Page (2007) and Surowiecki (2005).

In this paper, a collective is considered as a set of autonomous members (such as humans or systems). They are asked for giving their opinions (called predictions) on the same problem in the real world. Notice that these members may have different backgrounds, knowledge bases; therefore, their predictions can be different from each other. Since the proper value of the problem is not known before individual predictions made, then these predictions to some degree reflect the proper value. The collective prediction is determined on the basis of the individual predictions in the collective and considered as a representative for the collective as a whole. In general, the common approach to enhance the diversity level of a collective is to expand its cardinality. It is because as the collective cardinality increases, the diversity of backgrounds and opinions on the given problem will potentially increase. Moreover, it is useful in reducing the so-called correlated errors among individual opinions.

To determine the impact of collective cardinality on collective performance, we suppose that there exists a predefined threshold value of the maximal difference between the proper value and potential predictions of a given problem. For this aim, first, the impact of collective cardinality on collective performance that reported in Nguyen (2017) will be extensively investigated by taking into consideration two additional measures of collective performance. Subsequently, we investigate the problem of choosing a large collective with a small amount of information or a small collective with a large amount of information. The simulation experiments reveal that the larger collective cardinality we use, the higher collective performance we have. Moreover, the simulation results also emphasize the importance of choosing enough number of members for solving a given problem. Beyond the impact of collective cardinality, we also present some further challenges on the impact of diversity on collective performance. Furthermore, the problem of collective prediction determination in the case of large collectives is also investigated.

The structure of the remaining part is as follows. Some related works are briefly presented in the next section. Then we remind some basic notions used in the paper. Some criteria of intelligent collectives are introduced in the fourth section. After that, the simulation results on the impact of collective cardinality on collective performance are reported in the next section. The problem of determining collective predictions of large collectives is introduced in the sixth section. Some discussions on the impact of diversity on collective performance are presented in the seventh section. The paper ends with some conclusions and future work.

**Related works**

Although the problem of using a group of individuals for solving some difficult problems in the real world has been increasingly popular, it can be said that the experiment conducted by Galton in 1970 is the oldest one in the effectiveness of collective. In this experiment, 787 individuals are asked for estimating the weight of an ox (Galton, 1907). The collective guess was only nine pounds off the actual weight. The idea of the usefulness of this approach is that a collective may have additional knowledge that a single individual does not possess, and this knowledge may be relevant to the problem needed to be solved (Vroom, 2000). Similarly, in the case of ranking the weights of the objects,
introduced by Gordon (1924), the average of all individual rankings is 0.79 for collectives of 10; 0.86 for collectives of 20; and for collectives of 50 it is 0.94.

The statistical analysis from the game show ‘Who Wants to be a Millionaire?’, 91% of the answers given by the audience were on target, while in the case of experts only 65% answers were on target (Shermer, 2004). From this finding, it can be said that ‘A collective is better than single individuals.’ Later, Surowiecki (2005) has confirmed this statement by means of several experimental analyses. The most cited experiment is the one where a number of members were asked for giving their guesses on the number of beans in a jar. The collective guess is only 2.5% difference from the actual number of beans in the jar.

In Wagner and Suh (2014), 500 members were invited for solving a range of judgment problems such as forecasting the temperature of a region, and guessing the weights of specific amounts of coffee, milk, gasoline, air, and gold. The authors have found evidence that the larger collectives we use, the higher accuracy we get. The collective cardinalities used were 10, 20, 50, 100, and 200. The experiments conducted by Cui, Gallino, Moreno, and Zhang, (2015) have shown that a large collective has the positive impact on the accuracy of the collective forecast (collectives of 8 members and 18 members). Moreover, in Conradt (2011), the author has indicated that the large collective was positively associated with the accuracy of collective decision. In 2015, Arazy and colleagues utilized prediction markets for predicting weather events (Arazy, Halfon, & Malkinson, 2015). The findings have shown that collective predictions were superior to individuals’ predictions. Moreover, the difference between the prediction market and the meteorological model was not significant.

In the previous works (Nguyen & Nguyen, 2015a; 2016), we have preliminarily investigated the impact of collective cardinality on the accuracy of collective prediction. The simulation results have shown that the large collective positively affects the accuracy of its prediction. In this work, we aim at introducing some issues with collective cardinality in the effectiveness of collective performance. Moreover, some issues on processing large collectives; some further challenges of the impact of diversity on collective prediction are discussed.

**Preliminaries**

**Collective**

As mentioned already, a collective is considered as a set of members (such as humans or systems) who are invited for giving their opinions on a given problem. It is worth noting that collectives are different from traditional groups on the following characteristics. A group usually involves members that interact in person, whereas a collective can involve several members that may or may not interact. The members of a group often have the similar awareness and can give the similar solutions to the given problem that often causes the phenomenon of polarization (Janis, 1982) or representativeness fallacy (Argote, Devadas, & Melone, 1990). Conversely, the members of a collective do not necessarily have the similar viewpoints or opinions on the given problem.

**Collective of predictions**

Let $U$ be a set of objects representing the potential values referring to a problem needed to be solved. By symbol $\prod (U)$ we denote a set of all non-empty finite subsets with
repetitions of $U$. A collective of predictions $X \in \prod (U)$ involves the predictions given by collective members on the same problem and $X$ has the following form:

$$X = \{x_1, x_2, \ldots, x_n\},$$

where $n$ presents the number of predictions in collective $X$.

**Collective prediction**

From a collective of individual predictions, it is needed to determine a certain one that can be considered as the representative for the collective as a whole. For this aim, the integration algorithms proposed in Nguyen (2008a) may be used. In general, the most popular criteria that can be used for such integration tasks are as follows:

- **Criterion O₁**: if the sum of differences between the collective prediction and each prediction in collective $X$ to be minimal. That is,

  $$d(x^*, X) = \min_{y \in U} d(y, X).$$

- **Criterion O₂**: if the sum of squared difference between the collective prediction and each prediction in collective $X$ to be minimal. That is,

  $$d^2(x^*, X) = \min_{y \in U} d^2(y, X),$$

where $x^*$ presents the collective prediction of collective $X$, $d(x^*, X) = \sum_{i=1}^n d(x^*, x_i)$ and $d^2(x^*, X) = \sum_{i=1}^n d^2(x^*, x_i)$.

The collective prediction based on criterion $O₂$ has the following form:

$$x^* = (1/n) \sum_{i=1}^n x_i.$$ 

**Collective performance measures**

In crowd-based applications, beyond the accuracy of collective prediction, there exists another issue with the ability of a collective in comparison to those of its members in solving a given problem. Therefore, in this section, we will present three measures of collective performance. They are either based on the difference between collective prediction and the proper value or the ability of a collective outperforming its members.

**Based on the difference between collective prediction and proper value (Diff)**

The common approach for measuring the collective performance is based on the difference between the collective prediction and the proper value as in (Nguyen, 2008b).

**Definition 1.** The collective performance based on the difference between the collective prediction and the proper value is defined as follows:

$$\text{Diff}(X) = 1 - d(r, x^*),$$

where $x^*$ is the collective prediction of $X$, $r$ is the proper value. Notice that the values of $d(r, x^*)$ are normalized to $[0, 1]$. 
Based on win ratio

In the previous definition, the measure of collective performance only takes into account the accuracy of collective prediction. However, it can be argued that a collective is also called intelligent if its collective prediction is superior to individual predictions. By taking into account this capability, Kawamura and Ohuchi (2000) as well as Wagner and Vinaimont (2010) have defined a win ratio representing the number of times that the collective prediction is superior to the predictions in a collective.

Definition 2. The collective performance based on the win ratio is defined as follows:

\[ \text{WR}(X) = \frac{1}{n} \sum_{i=1}^{n} f(r, x^*, x_i), \]

where

\[ f(r, x^*, x_i) = \begin{cases} 1 & \text{if } d(r, x^*) < d(r, x_i), \\ 0 & \text{otherwise}. \end{cases} \]

Note that a win ratio (WR) of 75% means that 3–4 times the collective prediction is superior to individual predictions in a collective.

Based on the quotient of collective error and individual errors (QIC)

Another measure that also represents the capability of a collective to its members is based on the quotient of collective error and individual errors.

Definition 3. The collective performance based on the quotient of collective error and individual errors is defined as follows:

\[ \text{QIC}(X) = 1 - \frac{d^2(r, x^*)}{d^2(r, X)/n}. \]

According to Page (2007), the collective error does not exceed the individual errors, then QIC(X) \(\in [0, 1]\). It can be said that this measure is mainly suitable for O2 criterion.

The criteria of intelligent collectives

In the previous section, we have mentioned the superior capability of a collective to its members. For such capability, a collective is called intelligent if its collective prediction is not much different from the proper value (such as 5%); its collective prediction is better than individual predictions; the collective error is much smaller than the average of individual errors. In Surowiecki (2005), the author has opened up the underlying criteria of which a collective to be intelligent. They are diversity, independence, decentralization, and aggregation.

Diversity

In general, diversity can be considered as a variety of individual backgrounds or individual opinions on a given problem. It has often cited as one of the most important criteria of intelligent collectives (Page, 2007; Surowiecki, 2005). In Armstrong (2001), using heterogeneous
collectives have been recommended as an efficient way to make the collectives to be more intelligent. However, why is it important to have diversity in collectives? Intuitively speaking, diversity is not necessary for solving some common problems such as asking a number of high school students about the actual value of PI because most of them will answer it is 3.14. Conversely, with some difficult problems such as predicting the outcome of a future event or forecasting the demand for a new product, giving an accurate prediction seems not to be an easy task. Each prediction, in this case, can reflect a different aspect of the given problem. Then collectively the collective prediction is more accurate.

**Independence**

It can be argued that to ensure the existing of diversity in a collective; its members should make own predictions independently of others in the collective (Armstrong, 2006; Kahneman, 2011; Lorenz, Rauhut, Schweitzer, & Helbing, 2011; Page, 2007; Sunstein, 2006; Surowiecki, 2005). Independence is helpful in the effective of crowd-based applications because of correlated mistake reduction. Refer to Yin, Cui, and Huang (2011), if collective members have a minimal social influence of each other, then they can be independent of others in providing their opinions on a given problem.

**Decentralization**

According to Surowiecki (2005), a collective is called decentralization if its members are able to specialize and draw on local knowledge. It ensures members to act freely and independently of one another. It ensures members to act freely and independently of one another. Linux is known as the most popular decentralized systems. The solution to a concrete problem is often selected from many competing and diverse solutions worked out by independent programmers throughout the world.

**Aggregation**

In general, aggregation is a suitable mechanism aiming at combining individual predictions to produce a prediction. In Langlois and Roggman (1990), the composite face, which is combined using a computer, is more attractive than almost all of the individual faces. As mentioned earlier, two most popular aggregation methods to combine individual predictions are $O_1$ and $O_2$. In Nguyen (2008a), almost all of proposed algorithms are based on criterion $O_1$. In Galton (1907), criterion $O_1$ was used for determining collective guesses. Meanwhile, most experiments cited in Surowiecki (2005) used $O_2$ to determine collective guesses. In this work, criterion $O_2$ is used for such task.

**The impact of collective cardinality on collective performance**

**Simulation experiment 1**

As mentioned already, the common approach to enhance the diversity level of a collective is to expand its cardinality. In this section, we investigate to determine how expanding cardinality of a collective affects the effectiveness of a collective performance.
Simulation design 1
The simulation is similar to the problem of predicting the unknown outcome of a future event, given that there exists a maximal difference between the proper value and potential predictions of a given problem. In this paper, $U$ will be a set of integers belonging to the interval $[500, 1500]$. The proper value is assumed as 1000 for all experiment scenarios. The Manhattan distance is used for measuring the difference between predictions in a collective. The collective cardinality is expanded from 3 up to 1000. Each added element is randomly generated from the predefined set $U$. Therefore, each added prediction can be independent of other predictions in the collective.

Simulation results 1
Figure 1 presents the simulation results with respect to the previous simulation design. The values of Diff increase from 0.84 to 0.99. Similarly, the values of WR increase from 0.72 to 0.98 and the values of QIC increase from 0.61 to 0.99. Intuitively, we can state that the increase in collective performance can be caused by an increase in collective cardinality.

As can be seen from Figure 1, when the collective cardinality is large enough, its increase may not matter to collective performance. For further analysis, we computed the values of $R^2$ in the relationship between the log of collective cardinality and the collective performance.

From Table 1, the values of $R^2$ for Diff and WR increase with the collective cardinality increases up to 50. Meanwhile, $R^2$ for QIC increases only until the collective cardinality of 25 ($R^2 = 0.87$). To demonstrate such increases, we plotted the collective performance of selected collective cardinalities such as 3, 50, 75, 100, 250, 500, 750, and 1000 (see Figure 2).

Simulation experiment 2
From the previous section, it has been shown that the cardinality of a collective is positively associated with the collective performance. Moreover, when the number of
collective members is large enough, its increase does not matter to the collective performance. From the practical point of view, to solve some difficult problems, it often chooses between relying on a small number of members with a large amount of information or a large number of members with a small amount of information (Sunstein, 2006; Surowiecki, 2005). In this section, we present a comparison between the measures of collective performance by taking into account collective cardinality. For this aim, in the next simulation experiments, we will investigate two scenarios of collective members called novices and experts. The first scenario represents the case in which the members have a small amount of information. Meanwhile, the members in the second scenario will have a large amount of information. Meanwhile, the members in the second scenario will have a large amount of information.

**Simulation design 2**
In this simulation, we assume that the predictions given by novices will belong to [0, 2000]. Meanwhile, the predictions provided by experts will belong to [750, 1250]. Intuitively, the range of potential predictions given by novices is four times larger than that of potential predictions given by experts. Therefore, in the following experiments, the collective cardinalities in the case of novices are 20, 40, 100, and 200. Then, the collective cardinalities in the case of experts will be 5, 10, 25, and 50. We run 100 repetitions for each experiment.

**Simulation results 2**
In the following figures (from Figures 3–5), Diff_No presents Diff values of novices while Diff_Ex presents Diff values of experts (similar for WR and QIC).

As can be seen from Figures 3–5, the values of WR and QIC in the case of novices are higher than those in the case of experts. Conversely, the Diff values in the case of experts are always higher than those in the case of novices. It is because when the predictions in a collective are close to each other (in the case of experts), it is hard for the collective prediction to outperform individual predictions. Besides, the collective error seems to be close to the average of individual errors. Additionally, the measure of Diff only based on the collective prediction, while the others are based on both collective prediction and individual predictions. From this fact, it can be said that choosing the more knowledgeable members are often recommended because their predictions will be close to the proper value than those of less knowledgeable members. However, identifying this kind of members is a challenge because the proper value is not known when they are asked for giving predictions. From this finding, a raising question is how to measure the intelligence of a collective. Suppose that we have two collectives involves predictions as follows: \( X=\{1, 5, 9\} \), \( Y=\{4,5,6\} \). We also assume that the proper value is 5. In this case, we can see that the collective performances (in terms of Diff, WR, and QIC) of these collectives are identical. Based on these measures, we can state that these collectives have the same intelligence level. However, it can be said that collective \( Y \) is more intelligent than collective \( X \) because its

| Table 1. \( R^2 \) with different collective cardinalities. |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|
| Cardinality (from 3 up to) | 50  | 75  | 100 | 250 | 500 | 750 | 1000|
| \( R^2 \) (Diff)   | 0.96| 0.94| 0.91| 0.87| 0.86| 0.84| 0.82|
| \( R^2 \) (WR)     | 0.93| 0.92| 0.90| 0.87| 0.87| 0.85| 0.83|
| \( R^2 \) (QIC)    | 0.83| 0.78| 0.73| 0.61| 0.53| 0.47| 0.43|
predictions are closer to the proper value than those in collective $X$. This research problem should be the subject of the future work.

**Collective prediction of a large collective**

In some situation, for solving a given problem we need to refer opinions from a large number of members as in big data. The collectives, in this case, are often very large and a multi-stage approach can be useful in the process of collective prediction determination. However, we introduce a two-stage integration in which the collective prediction of a large collective is determined through two stages. With this approach, a large collective is clustered into smaller ones using an appropriate clustering method. Collective predictions of smaller collectives, which are determined in the first stage, will be treated as the

![Figure 2. Collective performances with selected cardinalities.](image)

![Figure 3. Comparison between WR of novices and WR of experts.](image)
predictions in a new collective. Subsequently, this new collective will be the subject for the second stage of the integration process. The detailed procedure of the two-stage integration is described in Figure 6.

In the previous work (Nguyen & Nguyen, 2015b), we have proposed an approach to reduce the difference between the collective predictions determined by single-stage and two-stage integrations. Moreover, we have also found the following dependencies between collective predictions determined by single-stage and two-stage integration process:

- If smaller collectives have the same number of members, then there is no difference between these collective predictions.
- If the quotient of the cardinality of each smaller collective and the total cardinality of the large collective is considered as the weight value for the collective prediction of the corresponding smaller collective, then there is also no difference between these collective predictions.

Figure 4. Comparison between QIC of novices and QIC of experts.

Figure 5. Comparison between Diff of novices and Diff of experts.
It can be concluded that the proposed method for collective prediction determination in the case of large collectives is helpful in reducing the execution time and computational complexity for some further computation tasks on large collectives such as the computation of diversity. This issue should be the subject of the future work.

Discussion

From the practical point of view, all research problems related to collective prediction have a common aim: To predict a future event with the highest accuracy. No research has been investigated so far for answering the question ‘How many people to ask to get high enough accuracy?’ The answer to this question is critical because it allows making the decision when to stop the prediction process with the certainty that the accuracy is high enough. From the previous research, it follows that the relationships between the cardinality of a collective and the accuracy of collective prediction exist, but its nature and characteristics are not known. It turned out that sociological and statistical tools are not enough for achieving robust and general results.

In Nguyen (2008a), the author has defined several functions that can be utilized to measure the diversity of collectives. We introduce two most used functions. The first one is based on the average of the differences between the predictions in a collective. A collective, which its predictions are close together, will have the small diversity level.

\[ c_3(X) = \begin{cases} 
\frac{1}{n(n-1)} \sum_{i=1}^{n} d(x_i, X) & \text{for } n > 1, \\
0 & \text{otherwise.} 
\end{cases} \]

where \( d(x_i, X) = \sum_{j=1}^{n} d(x_i, x_j) \), \( d(x_i, x_j) \) is a distance function measuring the difference between two predictions \( x_i \) and \( x_j \).

The second function is defined based on the minimum average of differences between a prediction of universe \( U \) and the predictions in a collective.

\[ c_5(X) = \min_{x^* \in U} d(x^*, X). \]

However, the biggest disadvantage is that their computation for large collectives is very time-consuming because it is needed to compute the differences between all pairs of predictions. Therefore, it is necessary to work out a more efficient algorithm for determining the diversity level for large collectives, which can consist of a vast number of members.

**Figure 6.** Two-stage integration of collective prediction determination.
Furthermore, diversity is an important criterion of which a collective to be intelligent. However, its impact on the accuracy is not formally proved. We consider the following example when a group of members is invited for giving their solutions on a given problem. For more detail, consider the following example.

**Example 1.** Let \( X = \{x_1, x_2, x_3\} \), \( Y = \{y_1, y_2, y_3\} \) be collectives of solutions for a given problem. Let \( x^* \) and \( y^* \) be the collective solutions of collectives \( X \) and \( Y \), respectively. Let \( r = (r_1, r_2) \) be the proper solution to the problem that these individual solutions reflect and the maximal difference between the proper solution and potential solutions is \( \alpha \). The difference between solutions in collectives is described in Figure 7.

According to Figure 7, the solutions in collective to the same degree reflect the proper solution. Then the diversity levels of collectives \( X, Y \) as follows:

\[
c_3(X) = \frac{\sum_{i=1}^{n} d(x_i, X)}{n(n-1)} = \frac{2 \times 3 \alpha \sqrt{3}}{3 \times 2} = a \sqrt{3},
\]

\[
c_3(Y) = \frac{\sum_{i=1}^{n} d(y_i, Y)}{n(n-1)} = \frac{2 \times 2(a \sqrt{3} + \alpha)}{3 \times 2} = \frac{2}{3} a \sqrt{3}.
\]

Intuitively, the diversity level of collective \( X \) is higher than that of collective \( Y \). In other words, the solutions in collective \( X \) are more diverse than those in collective \( Y \). From Example 1, we can state that in some sense the diversity has the positive impact on the collective performance. However, this statement should be formally proved and should be the subject of future work.

**Conclusion and future work**

In this paper, we have reported some issues with collective cardinality in the effective of collective performance. For this aim, we have simulated collectives with different cardinalities. The simulation results have revealed that the increase in collective performance is caused by an increase in collective cardinality. It also means the large collective has a positive impact on the collective performance. However, when the collective cardinality is large enough, its increase seems not to be effective to the

*Figure 7.* The positive impact of diversity.
collective performance. Moreover, the simulation results have also indicated the important role of choosing collective members for solving some difficult problems in the real world. In the future work, a mathematical model has been worked out to build collectives satisfying criteria of diversity, independence, decentralization, and aggregation. The methods for evaluating the independence level of a collective will also be proposed. Moreover, we also investigate the problem of determining the number of collective members needed to achieve a level accuracy of collective prediction in two cases: with a given degree of diversity level and without any assumption about the diversity degree.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by the Spanish MINECO/FEDER project DArDOS under Grants TIN2015-65845-C3-1-R and the Comunidad de Madrid project SICOMORo-CM under Grant S2013/ICE-3006.

Notes on contributors
Van Du Nguyen received the M.S. degree in Computer Science from Wroclaw University of Science and Technology in 2012. He is currently a Ph.D. student at Wroclaw University of Science and Technology, Poland. His research interests include knowledge integration, consensus theory, and collective intelligence.

Mercedes G. Merayo received her Ph.D. in Computer Science from Universidad Complutense de Madrid, Spain, in 2009. She holds an Associate Professor position in the Computer Systems and Computation Department at the same University. She has published more than 60 papers in refereed journals and international venues. She regularly serves in the Program Committee of conferences such as SEFM, ICTSS or QRS. Dr. Merayo has co-chaired QSIIC 2011, SEFM 2013, ICTSS 2014 and SAC-SVT 2017 among others. Her current research interests include model-based testing, distributed testing, asynchronous testing, mutation testing and timed extensions in formal testing.

References
Arazy, O., Halfon, N., & Malkinson, D. (2015). Forecasting rain events – Meteorological models or collective intelligence? Proceedings of EGU General Assembly Conference 17, Vienna, Austria (pp. 15611–15614).

Argote, L., Devadas, R., & Melone, N. (1990). The base-rate fallacy: Contrasting processes and outcomes of group and individual judgment. Organizational Behavior and Human Decision Processes, 46(2), 296–310.

Armstrong, J. S. (2001). Combining forecasts. In Principles of forecasting (pp. 417–439). New York, NY: Springer.

Armstrong, J. S. (2006). How to make better forecasts and decisions: Avoid face-to-face meetings. Retrieved from http://repository.upenn.edu/marketing_papers/44.

Conradt, L. (2011). Collective behaviour: When it pays to share decisions. Nature, 471(7336), 40–41.

Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2015). The operational value of social media information. Production and Operations Management. doi:10.2139/ssrn.2702151

Galton, F. (1907). Vox populi (the wisdom of crowds). Nature, 75, 450–451.
Gaur, V., Kesavan, S., Raman, A., & Fisher, M. L. (2007). Estimating demand uncertainty using judgmental forecasts. *Manufacturing & Service Operations Management, 9*(4), 480–491.

Gordon, K. (1924). Group judgments in the field of lifted weights. *Journal of Experimental Psychology, 7*(5), 398–400.

Graefe, A., & Armstrong, J. S. (2011). Comparing face-to-face meetings, nominal groups, delphi and prediction markets on an estimation task. *International Journal of Forecasting, 27*(1), 183–195.

Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes*. New York, NY: Houghton Mifflin.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kawamura, H., & Ohuchi, A. (2000). Evolutionary emergence of collective intelligence with artificial pheromone communication. Proceedings of 26th Annual Conference of the IEEE Industrial Electronics Society, Nagoya (pp. 2831–2836).

Lang, M., Bharadwaj, N., & Di Benedetto, C. A. (2016). How crowdsourcing improves prediction of market-oriented outcomes. *Journal of Business Research, 69*(10), 4168–4176.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1*(2), 115–121.

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences, 108*(22), 9020–9025.

Nguyen, N. T. (2008a). *Advanced methods for inconsistent knowledge management*. London: Springer-Verlag.

Nguyen, N. T. (2008b). Inconsistency of knowledge and collective intelligence. *Cybernetics and Systems, 39*(6), 542–562.

Nguyen, V. D. (2017). Collective knowledge: An enhanced analysis of the impact of collective cardinality. *Proc ACIIDS, 1*, 65–74.

Nguyen, V. D., & Nguyen, N. T. (2015a). A method for improving the quality of collective knowledge. *Proc ACIIDS, 1*, 75–84.

Nguyen, V. D., & Nguyen, N. T. (2015b). A two-stage consensus-based approach for determining collective knowledge. *Proc ICCSAMA, 1*, 301–310.

Nguyen, V. D., & Nguyen, N. T. (2016). An influence analysis of the inconsistency degree on the quality of collective knowledge for objective case. *Proc ACIIDS, 2*, 23–32.

Page, S. E. (2007). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton, NJ: Princeton University Press.

Shermer, M. (2004). *The science of good and evil*. New York, NY: Henry Holt.

Sunstein, C. R. (2006). *Infotopia: How many minds produce knowledge*. New York, NY: Oxford University Press.

Surowiecki, J. (2005). *The wisdom of crowds*. New York, NY: Doubleday/Anchor.

Vroom, V. H. (2000). Leadership and the decision-making process. *Organizational Dynamics, 28*(4), 82–94.

Wagner, C., & Suh, A. (2014). *The wisdom of crowds: Impact of collective size and expertise transfer on collective performance*. 47th Hawaii International Conference on System Sciences (HICSS), 2014, Waikoloa, HI, USA.

Wagner, C., & Vinaimont, T. (2010). Evaluating the wisdom of crowds. *Journal of Computer Information Systems, 11*(1), 724–732.

Yin, H., Cui, B., & Huang, Y. (2011). Finding a wise group of experts in social networks. *Proc ADMA, 1*, 381–394.