UNIVERSITY OF
GLOUCESTERSHIRE

This is a peer-reviewed, final published version of the following document and is licensed under All Rights Reserved license:

Gear, Tony and Endress, Tobias (2013) e-Delphi Pilot Experiment of Quality of Equity Predictions in Online Groups. International Journal of Management Cases, 15 (4). pp. 74-89.

EPrint URI: http://eprints.glos.ac.uk/id/eprint/1321

PLEASE SCROLL DOWN FOR TEXT.

International Journal of Management Cases

Special Issue

10th International CIRCLE Conference

Hosted by IPVC, Portugal
3rd - 6th April 2013

Volume 15 Issue 4

e-Delphi Pilot Experiment of Quality of Equity Predictions in Online Groups

Tobias Endress
University of Gloucestershire, UK

Tony Gear
University of Gloucestershire, UK

Abstract

This paper presents the results of a pilot experiment of stock price predictions by online groups, including the research process and a summary of the results. The overall objectives of the planned research study are three-fold: to assess the effect of individual and remote group decision-making approaches to stock price predictions; to assess whether a learning effect exists through the feedback loop of an e-Delphi process; and to identify the underlying key mechanisms of the individual and of the group that influence the decision-making process. The pilot run was performed with a small group (11 participants) and three financial analysts to benchmark the group over five e-Delphi cycles (five weeks). Each participant in the pilot was asked to provide an estimation of the movement (up or down) over a one-week and three-month period of four shares as well as enter a stock price prediction for a three-month period. The pilot run has provided some indications that, in certain situations and with careful group design, stock price predictions can be superior to the predictions of experts.

Key Words: e-Delphi, Online community, group decision-making processes, collective intelligence, equity predictions, stock-trading

Introduction

The purpose of this study is to inform understanding of the group decision making process of Internet communities which focus on stock-trading, based on predicting share prices. Equity research is a topic that is relevant for academic research as well as for business purposes. The work of academics that focus on financial markets and of business financial analysts has a special significance for brokers and investment banks, but it is also true that almost every financial newspaper, stock market journal or TV programme that deals with financial topics, reverts to these putative experts (Stanzel, 2007). Studies have already been conducted to assess the quality of the resulting forecasts from financial analysts (Bolliger, 2004; Clement, 1999; Fleischer, 2005; Stanzel, 2007). Generally, these studies show little evidence that it is possible to generate predictions that create, in the long run, and after transaction costs, profits higher than the market average (Malkiel, 2007; Stanzel, 2007).

The lack of reliable predictions appears to be one of the reasons why the investment community is still looking for new approaches to conducting traditional equity research and predicting future share prices. One of the alternative approaches to conducting equity research, generating investment ideas and creating stock market forecasts is the group decision approach (eg. Kaplan, 2001), which is used by several special interest (stock trading) communities on the Internet. This approach follows the proposal that a group-based decision may be able to outperform the decision of an individual (Page,

2008; Sunstein, 2008; Surowiecki, 2005).

Other authors doubt that groups can decide better than an expert; for example, essayist Henry David Thoreau, stated that "the mass never comes up to the standard of its best member but on the contrary degrades itself to a level with the lowest member" (as cited in Menschel, 2002, p.51). The philosopher Friedrich Nietzsche (1989), wrote that madness is rare in individuals, but he regarded it as the rule in groups, and Gustave Le Bon regarded crowds as "organisms", but argued that they can never attain a high degree of intelligence (2009). These two contrasting, but equally compelling views—regarding groups as "smarter" or groups as unintelligent—arguably rest on how the respective author views the "operation" of the group, and lead to an examination of the issues that influence group processes.

A pilot test using a small sample was conducted of the operation of the online process for the proposed research. The purpose of the study, following the pilot, was to gain an understanding of the group decision-making process used by Internet communities, focusing on stock trading based on the prediction of share prices. The design and methodology, as well as some preliminary results based only on the one week predictions of this pilot study have already been presented at the Second Annual Doctoral Colloquium in Berlin on July 14th (Endress, 2012). This paper presents the results of the pilot experiment, including an analysis of the three-month predictions as well as the price recommendations.

Pilot Stage Experiment Design and Data Analysis

To test and refine the process, the questions and the group design, a pilot run was performed with a small group (11 participants) and three financial analysts to benchmark the group over five e-Delphi cycles (five weeks).

The field experiment was conducted following an e-Delphi (Dalkey & Helmer-Hirschberg, 1962; Lindqvist & Nordänger, 2007) approach. Each e-Delphi cycle in this experiment consisted of a first stage for data collection of predictions. These data were compiled and distributed back to the group. In a second round, participants were able to provide different responses.
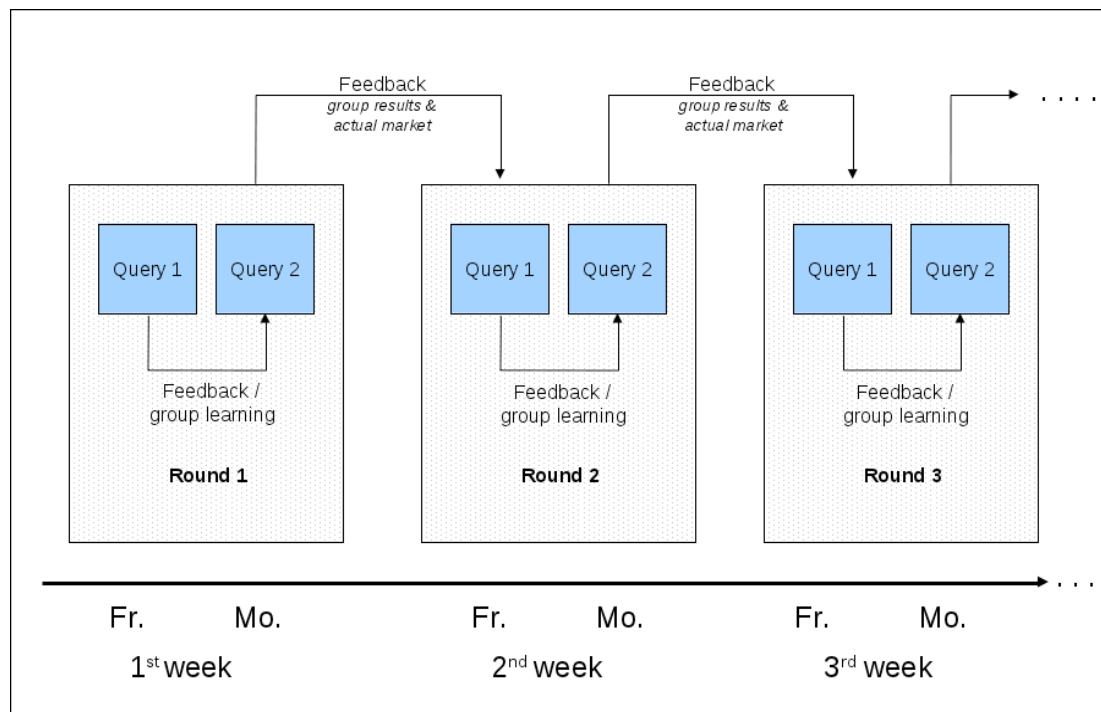
Figure 1. Group process within one round and between rounds from group as well as actual market results (Fr = Friday and Mo = the following Monday).

The shares were selected from four different companies in four different sectors: consumer goods (Adidas), chemicals (BASF), utilities (RWE) and industry (ThyssenKrupp). Each participant in the pilot was asked to provide an estimation of the movement (up or down) over a one-week and three-month period for every share as well as to enter a stock price prediction for a three-month period.

One-Week Predictions Pilot Stage

The pilot run of the group decision-making experiment demonstrated that a mixed-method approach (Creswell, 2009; Johnson & Onwuegbuzie, 2004; Tashakkori, 2010) works in this context. Handling the e-Delphi survey is possible given the set-up, software (Limesurvey) and Internet infrastructure that were used. The feedback from most participants was that the set up was simple to use and the questions were easy to understand. The analysis of the short-term predictions (one week) were analyzed and preliminary results and findings have been presented at the Second Annual Doctoral Colloquium in Berlin on July 14 (Endress, 2012).

The examination of the first estimations (for one week) showed that the group of lay people was slightly better in predicting stock price movements than the experts (see Table 1). From 40 predictions (m=40), the group had 22 (59.5%) correct predictions, the expert group had 16 (40%) correct predictions and the single expert had 18 (45%) correct predictions. In three rounds, the lay group came up with no recommendation (meaning that exactly 50% of the participants voted up and 50% voted down); these undecided rounds have been excluded from the analysis. The group's performance was better during weeks when the stock price was declining. From 17 predictions (m=17), the group had 10 (71.4%) correct predictions (three undecided rounds have been excluded), the expert group had six (35.3%) correct predictions and the single expert had nine (52.9%) correct predictions.

Table 1. Aggregated One-Week Pilot Run Predictions

	Single Expert		Expert Group		Lay Group			Measurements
	right	wrong	right	wrong	right	wrong	excluded	
Adidas	6	4	3	7	6	4	0	10
BASF	3	7	4	6	4	5	1	10
RWE	4	6	4	6	6	2	2	10
ThyssenKrupp	5	5	5	5	6	4	0	10
Sum	18	22	16	24	22	15	3	40

The group's overall decisions did not change from the first to the second e-Delphi round (see Table 2), Almost all group members stated in the interviews that they were not influenced by the group feedback from the e-Delphi rounds. The group may have a tendency towards conforming , in particular with price predictions. Additional data would be helpful to provide more knowledge of process.

Table 2. Pilot Run One Week Predictions in e-Delphi Round 1 and Round 2

	Single Expert		Expert Group		Lay Group			Measurements
e-Delphi	right	wrong	right	wrong	right	wrong	excluded	
Round 1	9	11	10	10	11	7	2	20
Round 2	9	11	6	14	11	8	1	20

Table 3 shows the performance of the individual members of the lay group and their self-estimated knowledge of the stock market (scale 1-10, from 1=no knowledge to 10=expert).

Table 3. Pilot Run One-Week Predictions of Lay Participants

	right	wrong	Measurements	Success Rate	Skill (Self Est.)
Proband 1	17	15	32	53,1%	3
Proband 2	20	16	36	55,6%	3-4
Proband 3	20	12	32	62,5%	2-3
Proband 4	22	14	36	61,1%	6
Proband 5	23	13	36	63,9%	1
Proband 6	22	18	40	55,0%	2
Proband 7	22	18	40	55,0%	7
Proband 8	26	10	36	72,2%	2
Proband 9	17	14	31	54,8%	7-8
Proband 10	17	15	32	53,1%	5
Proband 11	14	18	32	43,8%	2
Ø				57.3%	

All participants of the pilot were interviewed. The questions (see Appendix: Interview Questionnaire) were intended to gain a deeper understanding of the decision-making process and improve the design of the planned experiment. All participants agreed that the questions were easy to understand and all felt that they were able to make estimations or at least enter 'guesses' as to whether stock prices were going up or down. One participant felt uncomfortable giving a forecast of the stock price over a three-month period. He stated that he did not know the current stock price and, therefore, was not able to provide a forecast in terms of a concrete price target. In the interviews, some

participants asked why the survey did not ask for a one-week price target. Accordingly, asking for one-week and three-month price targets might be interesting, but not as mandatory fields in the online survey; it may be left to the participants to enter a concrete price target.

Three-Month Predictions Pilot Stage

The three-month predictions consisted of two components: an estimation of whether the share would go up or down, and an actual target price estimation for a 3-month period. Each participant had to enter both components for the four stocks in the pilot experiment independently.

Accuracy of Individual Predictions of e-Delphi Group Members

Table 4. Results Overview: Three-Month Predictions of Lay Participants

3 Months	Right	Wrong	Measurements	Success Rate	Skill (Self-Est.)
Proband 1	20	12	32	62.5%	3
Proband 2	23	13	36	63.9%	3–4
Proband 3	22	10	32	68.8%	2–3
Proband 4	10	26	36	27.8%	6
Proband 5	16	20	36	44.4%	1
Proband 6	20	20	40	50.0%	2
Proband 7	26	6	32	81.3%	7
Proband 8	22	14	36	61.1%	2
Proband 9	20	12	32	62.5%	7–8
Proband 10	12	20	32	37.5%	5
Proband 11	16	16	32	50.0%	2
Ø				55.4%	

The analyses of the individual results showed that 8 of 11 participants had a success rate of higher than 50% of the predictions (Table 4). Most participants missed one or two of the 10 e-Delphi rounds (= 5 x 2 rounds), but there was no "drop out" in terms of a participant leaving the panel during the five weeks without returning. All participants had been interviewed in parallel to the e-Delphi rounds. In the interviews, all participants were asked to provide self-assessment of their investment expertise on a scale from 1 to 10 (1 = no knowledge; 10 = expert). It may be hypothesized that there is a high correlation between success rate and self-estimated skill. An interesting observation is that this could not be confirmed by the results of the pilot experiment. Contrary to this hypothesis, for three-month predictions, there was a correlation of 0.12 and even a slightly negative correlation of self-estimated skill and success rate for the one-week predictions (-0.20). Table 6 shows the accuracy of the individual predictions of the experts (professional financial analysts) for the three-month estimates.

Comparison of Three-month Predictions of Each Share

In the examination of the longer-term estimations (for three-month) the group of lay people was again better in predicting stock price movement than were the experts (see Table 5). From 40 predictions (m = 40), the group had 17 right. In four rounds, the lay group came up with no recommendation (that is, exactly 50% of the participants voted up

and 50% down); these predictions have been excluded from the analysis. The expert group had 10 correct predictions (25%) and the single experts had 15 right (37.5%).

Table 5. Comparison of Three-Month Predictions Per Share

Three-Month	Single Expert		Expert Group		Lay Group			Measurements
	correct	wrong	correct	wrong	correct	wrong	excluded	
Adidas	5	5	4	6	6	4	0	10
BASF	4	6	4	6	0	8	2	10
RWE	6	4	2	8	4	5	1	10
ThyssenKrupp	0	10	0	10	7	2	1	10
Sum	15	25	10	30	17	19	4	40

The comparison of the three-month predictions of each share (see Table 5) shows that the lay group had more correct predictions than the expert group and also slightly more than the single expert within their narrow field of expertise. This result contradicts the hypothesis that while lay people might 'guess' the price movement more correctly in the short term, while over a period longer than a week, the expert opinion (based on rational valuation models and market insight) would outperform the lay group. The pilot experiment did not deliver any evidence for such an advantage on the part of the experts. Actually, even the best individual analysts did not perform better than the lay group (see Table 6).

Table 6. Results Overview: Three-Month Predictions of Experts

3-Month	Correct	Wrong	Measurements	Success Rate
Expert 1	9	31	40	22.5%
Expert 2	9	31	40	22.5%
Expert 3	19	17	36	52.8%
Ø				32.6%

Comparison of Three-Month Predictions from e-Delphi Rounds 1 and 2

From the initial experiments at RAND with the Delphi method. Dalkey and colleagues (1969) and Dakley and Helmer-Hirschberg (1962) concluded that there was a convergence of answers and an improvement in the second round. Dalkey (1969) stated "that without feedback there is either no improvement or degradation. The same groups showed definite improvement with feedback" (Dalkey, 1969, p. 66). Since that time, the Delphi method has become popular and has been used many times in a wide range of applications (Chen & Yang, 2004; Lindqvist & Nordänger, 2007; Linstone & Turoff, 2002). Nevertheless, now, some decades later, the application of the Delphi method in scientific research is not without criticism (van de Ven & Delbecq, 1974; Fischer, 1978; Linstone & Turoff, 2002). Despite all the controversy about the correct application and the value of the method, in the literature there is still a consensus that there is generally an improvement from the first round to the second round and that there is a tendency toward conforming with group opinion in the second Delphi round (Fischer, 1978; Linstone & Turoff, 2002; Rowe & Wright, 1999; Rowe, Wright, & McColl, 2005).

Table 7. Comparison of 3-Month Predictions from Rounds 1 and 2

3-Month	Single Expert		Expert Group		Lay Group			Measurements
e-Delphi	Correct	Wrong	Correct	Wrong	Correct	Wrong	Excluded	
Round 1	6	14	5	15	9	8	3	20
Round 2	9	11	5	15	8	11	1	20

It is relevant to note that many studies using the Delphi method have no stringent follow ups, and it is often unclear whether the predictions made with the Delphi panel turn out correct or not (e.g., Cole, 2008; Hsu, 2005; Kuhn, 2004). The results of the e-Delphi pilot experiment (see Table 7) involved a follow up, and even though it was only three months later, it is possible to assess whether or not the predictions were correct. The results of the pilot experiment are contrary to the opinion that there is an improvement with the second Delphi round. This might be attributed to the research design and feedback loop. Since participants received information about share prices and company development not only from the Delphi group but also from other sources, it might be possible that they relied more on the information from outside the group. In interviews with the group participants, some also said that they had not read the feedback before they made their second predictions. Some participants also mentioned that they did not trust the group because they did not know the degree of expertise of the group participants, or their rationales for their predictions. Linstone and Turoff have pointed out that "poor techniques of summarising and presenting the group response and ensuring common interpretations of the evaluation scales utilised in the exercise" (2002, p. 6) is a common weakness in Delphi surveys. Accordingly, it might be interesting to conduct a follow-up with variations in the feedback loop for the group.

Group Learning During the Pilot Run

Tables 8 and 9 show the correct and incorrect answers in the first half and second halves of the experiment. The short-term predictions improved considerably, but with the longer-term predictions, only the lay group's improved. The single experts and the expert group did less well in the second half of the pilot experiment.

Table 8. Comparison of One-Week Predictions from Weeks 1-5 and Weeks 6-10

One-Week	Single Expert		Expert Group		Lay Group			Measurements
e-Delphi	Correct	Wrong	Correct	Wrong	Correct	Wrong	Excluded	
Weeks 1-5	8	12	6	4	9	9	2	20
Weeks 6-10	10	10	10	10	13	6	1	20

Table 9. Comparison of Three-Month Predictions from Weeks 1-5 and Weeks 6-10

Three-Month	Single Expert		Expert Group		Lay Group			Measurements
e-Delphi	Correct	Wrong	Correct	Wrong	Correct	Wrong	Excluded	
Weeks 1-5	11	9	7	13	7	10	3	20
Weeks 6-10	4	16	3	17	10	9	1	20

Recommended Changes of Lay Group Participants

An analysis of the change behaviour of the participants shows that thay did not change

recommendations very often; however, when they did change, it was more often to a correct result than to an incorrect result. Overall, there where 56 actual changes of prediction during the pilot (see Table 10). That means only 14.9% of change options (N = 376) were used by the participants. In particular, the short-term predictions were better after the change: of 29 changes, 18 turned out to be correct and 11 wrong. This might be partly attributed to the shorter prediction period because the second round was only Monday to Friday, while the first round of predictions was from Saturday to Monday. This difference was needed to administer the e-Delphi experiment and organize the feedback loops. The changes of three-month predictions did not bring about such a big improvement: of 27 changes, 14 were correct and 13 wrong. This means for the three-month predictions, there was an overall improvement of only one recommendation.

Table 10. Changes of Recommendation of Lay Group Participants

	1st e-Delphi Week		2nd e-Delphi Week		3rd e-Delphi Week		4st e-Delphi Week		5st e-Delphi Week	
Adidas	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
1 W	1	1	1			1			1	
3 M		1	1					1	1	
BASF										
1 W	2	1	3			1			1	1
3 M				1		2		1	2	2
RWE										
1 W	2		1	1			1		2	
3 M			1			2		1	1	2
Thyssen Krupp										
1 W			2	1	1			1	3	
3 M		2		1	1	1		1		2
1 W Changes	5	2	7	2	1	2	1	1	7	1
3 M Changes	0	3	1	3	1	5	0	4	4	6

## Accuracy of 3-Month Price Predictions

The analysis of the accuracy of 3-month price predictions (see Table 3 in the Appendix) shows that there was not a big difference in prediction accuracy overall between the lay group and expert group. The price estimate of the lay group averaged 17.58% off the target from actual market prices, the single expert 17.41%, and the expert group 17.63%. For individual shares, there were some big differences: the single expert was better for RWE three-month price estimates, and the group outperformed the experts in the case of ThyssenKrupp (see Table 3). This finding supports a hypothesis and observation that the lay group performed well in comparison with experts, especially in the case of falling stock prices. ThyssenKrupp lost about 40% of market value (see Table 2), by far the highest loss of all shares in the pilot experiment.

Price Movement Changes

An analysis of the change events during the one-week predictions showed that there were nine changes of direction (in terms of movement change from up/down) during the five weeks of the pilot. Adidas changed price movement direction four times, Bayer changed direction four times, ThyssenKrupp changed twice, and RWE changed twice. These nine changes where correctly predicted by the group of experts six times, by the single experts six times and by the lay group five times.

An analysis of the change events during the three-month predictions showed that there where only six changes of direction during the five weeks. Bayer and ThyssenKrupp did not change, but continuously went down. RWE changed once and Adidas changed direction five times. These six changes where correctly predicted by the group of experts three times, by the single experts three times and by the lay group two times. Overall, the experts did slightly better than the lay group in the analysis of predictions of change events only, but,– due to the small data set of the pilot run,– it has to be noted that only one correct prediction for each prediction period made this difference.

Figure 2. Group comparison: three-month performance.
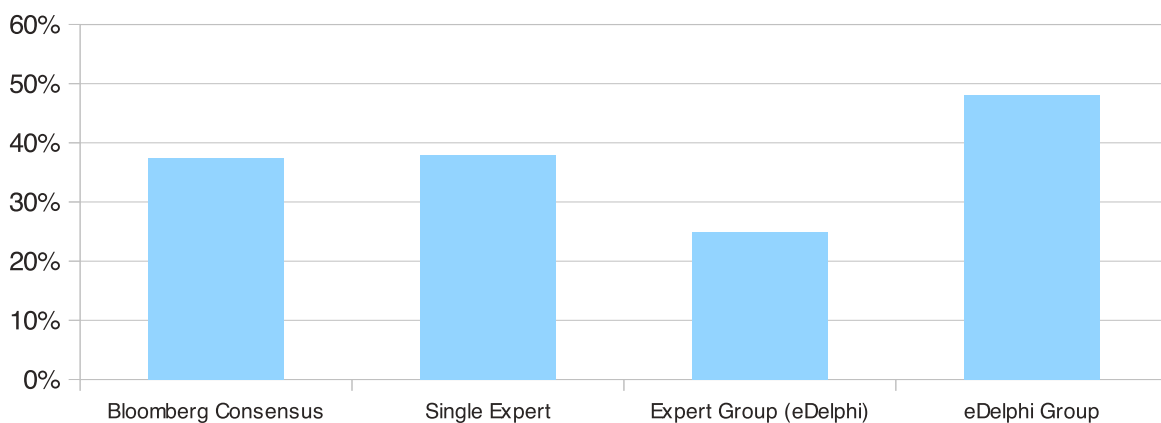
Value and Knowledge Contribution of the Pilot Run

The pilot run of the proposed experiment provided a few indications for an online group: in certain situations and with careful group design, predictions that are superior to predictions of experts are possible. In particular, the pilot run helped to identify the basic procedures for the individuals' decision-making approaches. These preliminary results are the basis for the later survey design. The results have indicated that some potential exists to improve the survey design and adjust the structure and process. In general, the pilot experiment demonstrated the feasibility of the experiment and showed that the tools and set up can be used to conduct the proposed experiment.

The pilot experiment was aimed at gaining a deeper understanding of the planned research. The overall research objectives of the planned research was to assess the impact of individual and remote group decision-making approaches to stock price predictions; to assess whether there was a learning effect through the feedback loop of

an e-Delphi process; and to identify the underlying key mechanisms of the individual and of the group that would influence the decision-making process. The three-month results generally confirmed the results from an examination of the one-week predictions (Endress, 2012). The pilot run of the group decision-making experiment demonstrated that a mixed-method approach works in this context, but it also showed some weaknesses and pitfalls of the planned research design. The pilot also provided valuable insight for improving the planned research approach; in particular, the e-Delphi survey. Reflective development of the research design is an iterative process during the research journey. Different ideas come up often, and old ideas were redefined accordingly. One interesting idea as a follow up might be to test a group with a stronger feedback loop, such as by having a short conversation involving group participants between Rounds 1 and 2. The pilot run of the proposed experiment also provided some indications that it might be possible for an online group to create (in certain situations and with careful group design) predictions that are superior to the predictions of experts.

Key Learnings From the Pilot Experiment

The pilot experiment generally demonstrated the feasibility of the research project to address the research questions with the research design. The pilot experiment provided some indication, of how the research design might be improved, as follows:
- Adjusting the group design and feedback loop
- Assessing the participants
- Enhancing the online questionnaire

Adjusting the Group Design and Feedback Loop

The group size of the pilot experiment (n=11) turned out to be quite appropriate in terms of manageability and explanatory power. However, it might be true that more data points and the coverage of more market phases (bull market and bear market) could help to increase the quality of the experiment. Accordingly, the main experiment should run longer than the five weeks of the pilot.

Another finding of the pilot was that the people did not change their predictions very often after they received the group feedback with the e-Delphi method. The literature suggests that there are more changes, and a stronger convergence in the group decisions (Dalkey, 1969; Dalkey & Helmer-Hirschberg, 1962). Therefore, it might be interesting to test the effect of the feedback loop more carefully. The literature suggests that one reason might be that the feedback loop is not strong enough. An interesting experiment might be to implement a stronger feedback loop for one group. This stronger feedback loop would be facilitated with an audio conference (with Skype) between e-Delphi round one and two. A second control group would be set up with no feedback from the group at all. With these three groups (regular e-Delphi Group, Interactive-/Conference call Group, and No-Feedback Group), it may be possible to more clearly determine the effect of the feedback on the groups' decision-making.

Assessment of the Participants

To understand more about the group decision making process it might be helpful to understand more about the decision-making process of the individual group participants as well. In order to gain more understanding of the individual decision-making process an

individual assessment of the participants should be conducted for all participants of the main experiment. This assessment should include age, gender, education level, profession and decision making type. While the questions about age, gender, education level and profession are quite easy to answer the question about decision-making type might not be. An approach to address this question was developed by Cornelia Betsch (2004; Schunk & Betsch, 2006; Traufetter, 2009). She created and thoroughly tested a questionnaire to determine the preference for intuition and/or deliberation of persons. An assessment of all participants might help to understand the reasons for particular predictions and to ensure that the three groups are equally diverse in terms of the assessed criteria.

Enhancements of the Online Questionnaire

The analysis of the procedure and the results of the pilot experiment also provided some indications in terms of how to improve the online questionnaire. The questions about the share movement (up or down) turned out to be useful and easy to understand, but not many participants provided information about their decision-making processes in the free text field on the online form. Nevertheless, the interviews with participants during the pilot run indicated some clusters of different types and sources for the decision-making process (see section below). In order to simplify the answer options and to obtain more information, these types will be provided as a tick-a-box field for the share estimations group of participants; that might make them more likely to provide more information about the reasoning behind their decision-making at the very moment they actually put their predictions into the online form. One participant of the pilot study did not feel comfortable providing a prediction of an actual price target for the three-month period; accordingly, it might be a good idea to change the question from a concrete stock price to a price movement in per-cent for this period. Additionally, this question should be changed to an optional question, for those who might still feel uncomfortable answering. Another change might be to introduce a question about 'conviction levels' for predictions (from not at all to absolutely sure, 1-5). Even though it might be interesting to include a few more questions, it also has to kept in mind that some participants indicated that they would not be willing to fill in a much longer questionnaire twice a week. In order to minimize the drop out rate, this needs to be taken seriously and the questionnaire should remain simple to answer and to understand.

Participant Interviews

All participants of the pilot were interviewed. The questions were intended to gain a deeper understanding of the decision-making process and to improve the design of the planned experiment. All participants agreed that the questions were easy to understand and all felt able to make estimation or at least enter guesses as to whether a stock price was going up or down. One participant felt uncomfortable making a forecast of the stock price over a three-month period. He stated that he did not know the current stock price and, therefore, was not able to provide a forecast in terms of a concrete price target. In the interviews, a few other participants asked why the survey did not ask for a one-week price target. Accordingly, asking for one-week and three-month price targets might be interesting, but should not be mandatory fields in the online survey. This should be left to the participants to decide whether they feel able to and comfortable in reference to entering concrete price targets with their predictions.

The interviews of the pilot experiment participants indicated different reasons for the

individual decisions In particular to the questions of the semi-structured interview: "How did you make your decision?", "Did you prepare for the survey rounds? If yes, how?" and, "Did you use external sources for the experiment? If yes, which ones?" The answers grouped in nine clusters of different decision-making influences in Table 11.

Table 11. Clusters of different decision-making fundamentals/influences

Company	Products, brands, customers, innovations, company development
Experts	Financial analysts and other expert opinions
Financial ratios	Market cap, P/E, dividend yields etc.
Fundamental analysis	Discounted cash flow, dividend discount model, peer group analysis etc.
Group results	Feedback from the e-Delphi group (last week or first round)
Intuition	Like gut feeling, instinct, guess
Market sentiment	General market situation and market outlook
News	Including daily press, Internet, business- and finance news
Technical analysis	Chart development, price-movement, momentum etc.

These clusters need to be presented as easy to understand options for the lay participants of the main experiment. They would be asked to tick a box or add a comment if they used something not mentioned there.

Synopsis and Conclusion

The pilot experiment was aimed at gaining a deeper understanding of the planned research. The overall research objectives of the planned research were to assess the impact of individual and remote group decision-making approaches to stock price predictions, and assess whether there was a learning effect through the feedback loop of an e-Delphi process, and identify the underlying key mechanisms of the individual and of the group that would influence the decision-making process. The three-month results generally confirmed the results from an examination of the one-week predictions (Endress, 2012). The pilot run of the group decision-making experiment demonstrated that a mixed-method approach works in this context, but also showed some weaknesses and pitfalls of the planned research design. The pilot also provided valuable insight to improve the planned research approach, in particular, the e-Delphi survey. Reflective development of the research design is an iterative process during the research journey. Different ideas often come up, and old ideas need to be redefined accordingly. An interesting idea as a follow up might be to test a group with a stronger feedback loop, such as by holding a short conversation among group participants between Rounds 1 and 2. The pilot run of the proposed experiment also provided some indications that it might be possible for an online group to create (in certain situations and with careful group design) predictions that are superior to the predictions of experts.

References

Betsch, C. (2004). Präferenz für Intuition und Deliberation (PID) [Preference for Intuition and Deliberation (PID): An Inventory for Assessing Affect- and Cognition-Based Decision-Making]. Zeitschrift für Differentielle und Diagnostische Psychologie, 25(4), 179–197. doi: 10.1024/0170-1789.25.4.179

Bolliger, G. (2004). The characteristics of individual analysts' forecasts in Europe. Journal of Banking & Finance, 28(9), 2283-2309.

Le Bon, G. (2009). Psychologie der Massen. Hamburg: Nikol.

Chen, S.-H., & Yang, C.-C. (2004). Applications of Web-QFD and E-Delphi method in the higher education system. HUMAN SYSTEMS MANAGEMENT, 23, 245–256.

Clement, M. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics, 27(3), 285-303. doi:10.1016/S0165-4101(99)00013-0.

Cole, M. (2008). A critical assessment of professional skills and knowledge in supplier diversity: A Delphi study. Antioch University. Retrieved from http://etd.ohiolink.edu/send-pdf.cgi/Cole%20Mia.pdf?antioch1223305907

Creswell, J. (2009). Research design: Qualitative, quantitative, and mixed method approaches (3th ed.). Thousand Oaks: SAGE Publications.

Dalkey, N. C. (1969). The Delphi method: An experimental study of group opinion. Santa Monica: Rand Corp.; Distributed by Clearinghouse for Federal Scientific and Technical Information, U.S. Dept. of Commerce, National Bureau of Standards, Institute for Applied Technology. Retrieved from http://www.rand.org/pubs/research_memoranda/RM5888

Dalkey, N. C., Brown, B. B., & Cochran, S. W. (1969). The Delphi Method, III: Use of self-ratings to improve group estimates. Santa Monica: Rand corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/research_memoranda/2006/RM6115.pdf

Dalkey, N. C., & Helmer-Hirschberg, O. (1962). An experimental application of the Delphi method to the use of experts. Santa Monica: Rand Corp.

Endress, T. (2012). A Pilot Experiment on Group Decision-Making in the Case of Online Group Stock Price Predictions. Proceedings of the 2nd DBA Annual Doctoral Colloquium (pp. 33–46). Presented at the University of Gloucestershire - Second DBA Annual Doctoral Colloquium, Berlin: University of Gloucestershire.

Fischer, R. G. (1978). The Delphi Method: A description, review, and criticism. Journal of Academic Librarianship, 4(2), 64–70. doi:Article

Fleischer, J. (2005). Die Performance von Analystenempfehlungen (1st ed.). Wiesbaden: Gabler.

Hsu, C. C. (2005). Identification of intangible resources essential to agri-tourism enterprises in Taiwan: a delphi study. The Ohio State University. Retrieved from http://ttracanada-torc.ca/torc/downs1/agritourism%20thesis.pdf

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. Educational researcher, 33(7), 14–26.

Kaplan, C. A. (2001). Collective intelligence: A new approach to stock price forecasting. IEEE INTERNATIONAL CONFERENCE ON SYSTEMS MAN AND CYBERNETICS (Vol. 5,

pp. 2893–2898).

Kuhn, J. (2004). Kommerzielle Nutzung mobiler Anwendungen-Ergebnisse der Delphi-Studie" Mobile Business." Retrieved from http://epub.uni-regensburg.de/10163/1/Kommerzielle_Nutzung_mobiler_Anwendungen.pdf

Lindqvist, P., & Nordänger, U. K. (2007). (Mis-?) using the E-Delphi Method: An attempt to articulate the practical knowledge of teaching. Journal of Research Methods and Methodological Issues, Volume 1(issue 1, 2007). Retrieved from http://www.scientificjournals.org/journals2007/articles/1222.pdf

Linstone, H. A., & Turoff, M. (Eds.). (2002). The Delphi Method: Techniques and Applications. Addison-Wesley. Retrieved from http://is.njit.edu/pubs/delphibook

Malkiel, B. (2007). A random walk down Wall Street: The time-tested strategy for successful investing ([9th ed.].). New York: W.W. Norton.

Menschel, R. (2002). Markets, mobs & mayhem a modern look at the madness of crowds. Retrieved from http://www.netlibrary.com/urlapi.asp?action=summary&v=1&bookid=74081

Nietzsche, F. (1989). Beyond good and evil: Prelude to a philosophy of the future. New York: Vintage.

Page, S. E. (2008). The difference: How the power of diversity creates better groups, firms, schools, and societies. Princeton N.J.; Woodstock: Princeton University Press.

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. International journal of forecasting, 15(4), 353–375.

Rowe, G., Wright, G., & McColl, A. (2005). Judgment change during Delphi-like procedures: The role of majority influence, expertise, and confidence. Technological Forecasting and Social Change, 72(4), 377–399. doi:10.1016/j.techfore.2004.03.004

Schunk, D., & Betsch, C. (2006). Explaining heterogeneity in utility functions by individual differences in decision modes. Journal of Economic Psychology, 27(3), 386–401. doi:10.1016/j.joep.2005.08.003

Stanzel, M. (2007). Qualität des Aktienresearch von Finanzanalysten: Eine theoretische und empirische Untersuchung der Gewinnprognosen und Aktienempfehlungen am deutschen Kapitalmarkt (1st ed.). Wiesbaden: Dt. Univ.-Verl.

Sunstein, C. (2008). Infotopia: How many minds produce knowledge. New York; Oxford: Oxford University Press.

Surowiecki, J. (2005). The wisdom of crowds (1st ed.). New York: Anchor Books.

Tashakkori, A. (2010). Sage handbook of mixed methods in social & behavioral research (2nd ed.). Los Angeles: SAGE Publications.

Traufetter, G. (2009). Intuition die Weisheit der Gefühle. Reinbek bei Hamburg: Rowohlt-

Taschenbuch-Verl.

van de Ven, A. H., & Delbecq, A. L. (1974). The effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes. Academy of Management Journal, 17(4), 605–621.

Appendix

Interview Questionnaire (Pilot Experiment)
Topics and sample questions to be explored in the semi-structured interviews with all participants of the pilot experiment:

Question	German Version	English Translation	Topics
1	Hattest Du das Gefühl, dass die Fragen leicht zu beantworten sind?	Did you feel it was easy to answer the questions?	General / Initial Situation
2	Musstest Du Dein Internet verhalten ändern, um an der Umfrage teilzunehmen?	Did you need to change your Internet usage in order to participate at the survey?	General / Initial Situation
3	Wie würdest Du Deine Kenntnisse zum Aktienmarkt selbst einschätzen? (1 gar keine; 10 Experte)	How would you self-assess your knowledge about the stock market? (1, no knowledge; 10, expert)	General / Initial Situation
4	Was war die Grundlage für Deine Entscheidungen?	How did you make your decision?	Decision-Making Process
5	Hast Du Dich auf die Umfrage-Runden vorbereitet? Wenn ja, wie?	Did you prepare for the survey rounds? If yes, how?	Decision-Making Process
6	Hast Du für das Experiment auf externe Quellen zugegriffen? Wenn ja, welche?	Did you use external sources for the experiment? If yes, which ones?	Decision-Making Process
7	Hast Du selbst Aktien gekauft? Auch welche die in dem Experiment vorkommen?	Did you ever buy shares? Did you buy some used in the experiment?	Personal Impact
8	Denkst Du das e-Delphi-Experiment bzw. die Gruppenergebnisse hat Deine Entscheidungen beeinflusst?	Do you think the e-Delphi-experiment / the group results influenced your decisions?	Personal Impact
9	Du hast Deine Entscheidung [X mal von Y to Z] in Runde 2 geändert, warum?	You changed your decision [X times from Y to Z] in round 2; why?	Personal Impact
10	Denkst Du, dass Du neue Expertise oder Erkenntnisse hinzugewonnen hast?	Do you think you gained new expertise or knowledge during this experiment?	Personal Impact

11	Achtest Du jetzt mehr auf Nachrichten, insbesondere zu den Unternehmen der Umfrage-Runden?	Do you care more about news now, in particular news of the companies of the survey?	Personal Impact
12	Was denkst Du über das Web-Umfrage-Tool?	What do you think about the usability of the web survey tool?	Survey Structure / Web Tool
13	Was würdest Du an der Umfrage verbessern?	What would you like be changed for the survey?	Survey Structure / Web Tool
14	Hast Du weitere Kommentare oder Anregungen?	Any further comments or suggestions?	General Issues